Politechnika Śląska

Wydział Automatyki, Elektroniki i Informatyki



Inteligentne nadrzędne sterowanie zespołem sprężarek minimalizujące zużycie energii elektrycznej przy zapewnieniu stałości dostaw sprężonego powietrza w kontekście założeń Przemysłu 4.0

Autor:
mgr inż. Kamil Kasprzyk

Promotor:
dr hab. inż. Adam Gałuszka

Opiekun pomocniczy:
dr inż. Łukasz Antczak

Rozprawa doktorska
w dyscyplinie

Automatyka, elektronika, elektrotechnika i technologie kosmiczne

Gliwice 2026

Streszczenie

Tytuł: Inteligentne nadrzędne sterowanie zespołem sprężarek minimalizujące zużycie energii elektrycznej przy zapewnieniu stałości dostaw sprężonego powietrza w kontekście założeń Przemysłu 4.0

Słowa kluczowe: BiLSTM, sprężarki śrubowe, predykcja obciążenia, dedykowana funkcja straty, sterowanie nadrzędne, efektywność energetyczna, asymetryczny koszt, Przemysł 4.0

Sprężone powietrze jest jednym z podstawowych mediów energetycznych w przemyśle, często stanowiąc istotny koszt operacyjny zakładu produkcyjnego. Dyrektywa Unii Europejskiej 2023/1791 nakłada obowiązek ograniczenia zużycia energii elektrycznej, której sprężarki śrubowe są istotnym konsumentem. Jednocześnie Przemysł 4.0 zachęca do rozwoju oraz wykorzystania zbieranych i archiwizowanych danych procesowych w nowy sposób, jakim może być wzbogacenie istniejących systemów sterowania o warstwę predykcyjną, której celem w rozprawie jest minimalizacja zużycia energii elektrycznej w układach sprężonego powietrza.

W celu synchronizacji pracy wielu sprężarek śrubowych zakłady produkcyjne używają sterownika nadrzędnego, który podejmuje decyzję o sekwencyjnym załączaniu lub wyłączaniu maszyn, dostosowując ich liczbę do aktualnego zapotrzebowania na sprężone powietrze. Systemy te, często bazujące na dedykowanych zamkniętych sterownikach lub programowalnych sterownikach logicznych, nie mają wiedzy o cykliczności lub wzorcach zapotrzebowania, jakie występują w danym zakładzie. Skutkiem tego oraz dużej zmienności zapotrzebowania, może być potrzeba utrzymywania maszyn w stanie biegu jałowego, w którym pobierają energię elektryczną jednocześnie nie wytwarzając sprężonego powietrza.

Rozprawa proponuje warstwę predykcyjną systemu sterowania jako wartość dodaną istniejących rozwiązań sterowania nadrzędnego, bazując na dwukierunkowej sieci rekurencyjnej BiLSTM, przewidując stany obciążenia poszczególnych maszyn w zadanym horyzoncie na podstawie przebiegu ciśnienia i stanów historycznych. Zaprojektowane rozwiązanie nie wymaga dodatkowego opomiarowania i korzysta z już dostępnych sygnałów. Praca wprowadza czteroskładnikową funkcję straty, którą można zaadaptować do różnych charakterystyk poboru, z naciskiem na wykrywanie błędów krytycznych

(mogących zachwiać stabilność dostaw sprężonego powietrza), które stanowią kryterium walidacji modelu produkcyjnego. Do oceny przyjęte zostały wskaźniki, które zamiast bazować na uśrednionych miarach odchyłek, skupiają się na ocenie decyzji wspierających wyłączenie maszyny przy pojawieniu się stanu odciążenia.

Metodyka została zweryfikowana na 12 obiektach przemysłowych należących do trzech popularnych klas o zróżnicowanej strukturze oraz zapotrzebowaniu. Wdrożenie produkcyjne zostało wykonane na ośmiu obiektach, spośród których cztery nie wchodziły w skład badanych wcześniej. Studium ablacji wykazało, że dwukierunkowość przetwarzania jest warunkiem koniecznym, ponieważ architektury jednokierunkowe mogą degenerować się do przewidywania tylko jednego stanu. Wdrożenia przyniosły redukcję czasu pracy na biegu jałowym sięgającą 83%, przy dwóch obiektach bez poprawy statystycznej. Wskazano również granice stosowania metody oraz wskaźniki pozwalające je rozpoznać. Wdrożenie zaowocowało również wskaźnikiem nadzoru operacyjnego pozwalającego wykryć degradację modelu.

Wkładem pracy jest funkcja straty wiążąca koszt procesowy z wyborem modelu, zestaw wskaźników oceny odniesionych do zdarzeń decyzyjnych, a także kryteria kwalifikacji obiektu i mechanizm nadzoru operacyjnego.

Abstract

Title: Intelligent supervisory control of a compressor system minimizing electricity consumption while ensuring a stable compressed air supply, in the context of Industry 4.0 principles.

Keywords: BiLSTM, screw compressors, load prediction, custom loss function, supervisory control, energy efficiency, asymmetric cost, Industry 4.0

Compressed air is one of the principal energy carriers in industry and often represents a considerable share of a manufacturing plant's operating costs. Directive (EU) 2023/1791 imposes an obligation to reduce electricity consumption, of which screw compressors are a substantial consumer. At the same time, Industry 4.0 encourages new ways of exploiting the process data that plants already collect and archive. One such use is to extend an existing control system with a predictive layer, whose purpose in this dissertation is to minimize the electricity consumed by compressed air systems.

To coordinate the operation of several screw compressors, manufacturing plants employ a supervisory controller that sequentially starts and stops machines, matching their number to the current demand for compressed air. These systems, commonly built on proprietary embedded controllers or on programmable logic controllers, have no knowledge of the cyclical behaviour or demand patterns specific to a given plant. As a consequence, and because demand varies considerably, machines may have to be kept running unloaded, drawing electrical power while producing no compressed air.

The dissertation proposes a predictive control layer that adds value to existing supervisory control solutions. It is based on a bidirectional recurrent network (BiLSTM) and predicts the load states of individual machines over a given horizon from the pressure trajectory and historical states. The proposed solution requires no additional instrumentation and relies solely on signals that are already available. The work introduces a four-component loss function that can be adapted to different demand characteristics, with emphasis on detecting critical errors, that is, those capable of compromising the stability of the compressed air supply, which serve as the criterion for validating the production model. The evaluation adopts indicators that, rather than relying on averaged deviation measures, assess the decisions that support switching a machine off once an unloaded state occurs.

The methodology was verified on twelve industrial installations belonging to three

common classes of differing structure and demand. Production deployment was carried out on eight installations, four of which were not part of the previously studied set. An ablation study showed that bidirectional processing is a necessary condition, since unidirectional architectures may degenerate into predicting a single state. The deployments reduced unloaded running time by up to 83%, with two installations showing no statistically significant improvement. The limits of applicability of the method were also identified, together with the indicators that allow them to be recognized. The deployment further yielded an operational supervision indicator capable of detecting model degradation.

The contribution of this work comprises a loss function that links process cost to model selection, a set of evaluation indicators referred to decision events, together with criteria for qualifying an installation and a mechanism for operational supervision.

Mojej żonie Asi i córce Matyldzie

Spis treści

Streszczenie **iii**

Abstract **v**

Spis rysunków **xv**

Spis tabel **xvii**

Spis skrótów **xxi**

Spis symboli **xxiii**

1 Wprowadzenie **1**

1.1 Kontekst badań: Przemysł 4.0 i optymalizacja energetyczna 1
1.2 Problematyka układów sprężonego powietrza 2
1.3 Motywacja i uzasadnienie badań 3
1.4 Problem badawczy 5
1.5 Cele pracy 11
1.6 Tezy badawcze 11
1.7 Zakres i ograniczenia pracy 12
1.8 Struktura rozprawy 13

2 Przegląd literatury **15**

2.1 Układy sprężonego powietrza w przemyśle 15
2.1.1 System sprężonego powietrza 15
2.1.2 Dobór właściwych maszyn i ich ograniczenia 16
2.1.3 Struktura sprężarkowni 19
2.1.4 Problemy eksploatacyjne i kierunki optymalizacji 21
2.1.5 Tradycyjne metody optymalizacji 22
2.2 Uczenie głębokie w sterowaniu przemysłowym 23
2.2.1 Kontekst historyczny — od klasycznej automatyki do Przemysłu 4.0 23
2.2.2 Sieci rekurencyjne (LSTM, BiLSTM, GRU) 28
2.2.3 Zastosowania predykcji w energetyce 33

2.3 Uczenie wrażliwe na koszt i asymetryczne koszty . . . 36
2.4 Funkcje straty w zadaniach z asymetrycznym kosztem . . . 36
2.4.1 Funkcje straty w uczeniu maszynowym . . . 36
2.4.2 Funkcja straty ogniskowej i warianty asymetryczne . . . 38
2.4.3 Asymetria w prognozowaniu szeregów czasowych . . . 39
2.5 Ocena modeli przy asymetrycznych kosztach . . . 40
2.6 Wybór modelu i strojenie hiperparametrów . . . 41
2.7 Dryft koncepcyjny i uczenie w eksploatacji . . . 42
2.8 Identyfikacja luki badawczej . . . 43

3 Metodyka badań 49
3.1 Sformułowanie problemu . . . 49
3.1.1 Definicja i założenia . . . 49
3.1.2 Dane wejściowe . . . 50
3.1.3 Schemat metodyki . . . 52
3.2 Autorska asymetryczna funkcja straty . . . 53
3.2.1 Koncepcja i uzasadnienie . . . 53
3.2.2 Składowe funkcji straty . . . 53
3.2.3 Wariant bazowy mse_bce . . . 56
3.3 Architektura sieci BiLSTM . . . 56
3.3.1 Uzasadnienie wyboru architektury . . . 56
3.3.2 Konfiguracja sieci . . . 57
3.3.3 Kierunek przetwarzania danych . . . 57
3.4 Metodyka optymalizacji hiperparametrów . . . 58
3.4.1 Dwuetapowe podejście . . . 58
3.4.2 Siatka kombinacji . . . 59
3.4.3 Optymalizacja bayesowska (Optuna) . . . 59
3.5 Metodyka ewaluacji . . . 59
3.5.1 Okno decyzyjne . . . 59
3.5.2 Metryki ewaluacyjne i bezpieczeństwa . . . 61
3.5.3 Ranking modeli . . . 63
3.5.4 Walidacja wieloobiektowa . . . 63
3.6 Środowisko i narzędzia . . . 64
3.7 Podsumowanie . . . 64

4 Eksperymenty i wyniki 65
4.1 Charakterystyka obiektów badawczych . . . 65
4.2 Ocena zmienności obiektów . . . 66
4.3 Założenia eksperymentów . . . 70
4.4 Studium ablacji . . . 71

4.5 Wyniki dla grupy A — obiekty bez maszyn z falownikiem 75
4.5.1 Obiekt A1 76
4.5.2 Obiekt A2 77
4.5.3 Obiekt A3 79
4.5.4 Obiekt A4 80
4.6 Wyniki dla grupy B — obiekty z jedną maszyną falownikową 82
4.6.1 Obiekt B1 83
4.6.2 Obiekt B2 84
4.6.3 Obiekt B3 86
4.6.4 Obiekt B4 88
4.7 Wyniki dla grupy C — obiekty z samymi maszynami VFD 90
4.7.1 Obiekt C1 90
4.7.2 Obiekt C2 92
4.7.3 Obiekt C3 93
4.7.4 Obiekt C4 95
4.8 Analiza wpływu horyzontu predykcji 96
4.9 Podsumowanie 98

5 Wdrożenie produkcyjne **99**
5.1 Architektura modułu SI 99
5.2 Zabezpieczenia i przebieg wdrożenia 102
5.2.1 Parametry zabezpieczeń 102
5.2.2 Opóźnienie realizacji decyzji 103
5.2.3 Weryfikacja decyzji 103
5.2.4 Schemat przygotowania danych 104
5.2.5 Dopuszczenie obiektu do wdrożenia 105
5.3 Charakterystyka wdrożeń 105
5.4 Metryki modeli produkcyjnych 106
5.5 Nadzór operacyjny — wskaźnik R_{FN} 108
5.5.1 Definicja 108
5.5.2 Wyniki eksploatacyjne oraz wpływ warstwy zabezpieczeń 110
5.6 Oszczędność energii elektrycznej 112
5.6.1 Potencjał oszczędnościowy 112
5.6.2 Wynik 806 dni wdrożenia 113
5.7 Zauważone mechanizmy degradacji 116
5.7.1 Czynniki nieprzewidywalne — obiekt A3 116
5.7.2 Dryft koncepcyjny — zmiana profilu produkcji — obiekt D3 . . 116
5.7.3 Zanikanie efektywności mimo dotrenowania — obiekt A1 117
5.7.4 Brak potencjału oszczędnościowego — obiekt C1 118

5.8 Rola dotrenowania modelu podczas eksploatacji . . . 119
5.9 Studium przypadku — obiekt D3 . . . 122
5.9.1 Konfiguracja startowa . . . 122
5.9.2 Gradient ciśnienia . . . 123
5.9.3 Znaczenie okresu próbkowania . . . 124
5.9.4 Zmiana okresu próbkowania . . . 126
5.9.5 Optymalizacja hiperparametrów . . . 128
5.9.6 Opóźnienie predykcji . . . 128
5.9.7 Opóźnienie realizacji decyzji sekwencyjnych . . . 129
5.9.8 Eksploatacja i zakończenie pracy . . . 129
5.9.9 Wnioski ze studium przypadku . . . 130
5.10 Podsumowanie . . . 131

6 Dyskusja **135**
6.1 Interpretacja wyników badawczych . . . 135
6.1.1 Rola funkcji straty . . . 135
6.1.2 Wpływ architektury sieci . . . 136
6.1.3 Znaczenie horyzontu predykcji . . . 137
6.2 Wnioski z wdrożenia . . . 138
6.2.1 Rozbieżność wskaźników oceny offline a wskaźnikiem oceny operacyjnej . . . 138
6.2.2 Znaczenie nadzoru operacyjnego . . . 139
6.2.3 Wierność odwzorowania procesu . . . 139
6.2.4 Parametry czasowe realizacji decyzji . . . 140
6.2.5 Oszczędność energii . . . 141
6.3 Kiedy zastosowanie uczenia maszynowego/głębokiego jest uzasadnione . 142
6.4 Porównanie z literaturą i wskazanie wkładu własnego . . . 145
6.5 Ograniczenia badań . . . 147
6.6 Znaczenie w kontekście założeń Przemysłu 4.0 . . . 149

7 Podsumowanie **151**
7.1 Podsumowanie przeprowadzonych prac . . . 151
7.2 Weryfikacja tez badawczych . . . 152
7.3 Oryginalny wkład autora . . . 156
7.4 Zalecenia praktyczne . . . 157
7.5 Kierunki dalszych badań . . . 159

Bibliografia **163**

A Opis implementacji kluczowych modułów **173**
A.1 Funkcje straty . . . 173
A.1.1 Wariant bazowy . . . 173
A.1.2 Wariant asymetryczny . . . 173
A.2 Architektura sieci . . . 175
A.3 Wyznaczanie metryk bezpieczeństwa . . . 176
A.4 Weryfikacja następcza decyzji w eksploatacji . . . 177

B Konfiguracja systemu **179**
B.1 Konfiguracja procesu uczenia . . . 179
B.2 Konfiguracja wdrożeniowa . . . 180
B.2.1 Komunikacja ze sterownikiem . . . 180
B.2.2 Mapowanie sygnałów . . . 181
B.2.3 Parametry sprężarek . . . 181
B.2.4 Parametry sterowania i bezpieczeństwa . . . 182
B.2.5 Dotrenowanie i monitorowanie . . . 182

Spis rysunków

1.1 Sprężarkownia firmy Marani Sp. z o.o. składająca się z pięciu sprężarek śrubowych. 6
1.2 Szafka sterownika nadrzędnego bazująca na sterowniku PLC. 7
1.3 Struktura nadrzędnego sekwencyjnego sterowania zespołem sprężarek. Wartością zadaną jest zakres ciśnienia $[p_{\min}, p_{\max}]$, a komparator z histerezą generuje komendy dwupołożeniowe $d(k) \in \{-1, 0, +1\}$, bez aktywnej regulacji w zakresie granic. 10
1.4 Przebieg ciśnienia i stanów pracy sprężarek rzeczywistej instalacji. Pominięto wartości ciśnienia oraz oznaczenia czasu. 10

3.1 Schemat blokowy proponowanej metodologii predykcji stanów obciążenia sprężarek śrubowych. Postępowanie rozpoczyna się od wariantu bazowego funkcji straty z wagami domyślnymi; wariant rozszerzony wraz ze strojeniem wag wprowadzany jest dopiero wtedy, gdy ocena wskaźników bezpieczeństwa okaże się niewystarczająca. 52

5.1 Schemat blokowy architektury wdrożeniowej. 102

Spis tabel

2.1 Porównanie wybranych prac z niniejszą rozprawą. ✓ — cecha występuje, - — nie występuje. 44
2.2 Publikacje własne związane z realizacją rozprawy. 46

3.1 Parametry czasowe modelu i ich rola w ocenie dla przykładowego okresu próbkowania $T_s = 20$ s. 60

4.1 Charakterystyka obiektów badawczych. N — łączna liczba sprężarek, N_{VFD} — liczba sprężarek z falownikiem, T_s — okres próbkowania. . . . 66
4.2 Charakterystyka zmienności cykli pracy obiektów badawczych. CV — współczynnik zmienności (odchylenie standardowe odniesione do wartości średniej), γ — współczynnik skośności rozkładu długości okresów biegu jałowego, $\mu/\tilde{x}$ — stosunek wartości średniej do mediany, $\tau_{<h}$ — udział okresów biegu jałowego krótszych od horyzontu predykcji. 68
4.3 Charakterystyka zmienności cykli pracy obiektów wdrożeniowych. Oznaczenia jak w tabeli 4.2. Dla obiektu D3 zestawiono dodatkowo wartości uzyskane przed zmianą okresu próbkowania (D3, $T_s = 20$ s) oraz po niej (D3', $T_s = 5$ s). 68
4.4 Hiperparametry modelu BiLSTM, jednakowe dla wszystkich obiektów badawczych. 70
4.5 Wyniki studium ablacji dla obiektów A3, B1 i C3, horyzont 2 min. Wskaźniki policzono łącznie dla wszystkich sprężarek obiektu. 71
4.6 Konfiguracja wariantów funkcji straty w studium ablacji. Wagi składowych nie są normalizowane. We wszystkich wariantach ze składową $\mathcal{L}_{reload}$ mnożnik kary za FP wynosi $\beta = 1$. 72
4.7 Liczby zdarzeń w studium ablacji dla obiektu A3, horyzont 2 min, suma dla wszystkich sprężarek. N_T — liczba tranzycji. 73
4.8 Liczby zdarzeń w studium ablacji dla obiektu B1, horyzont 2 min, suma dla wszystkich sprężarek. N_T — liczba tranzycji. 74
4.9 Liczby zdarzeń w studium ablacji dla obiektu C3, horyzont 2 min, suma dla wszystkich sprężarek. N_T — liczba tranzycji. 75
4.10 Zdarzenia w oknach decyzyjnych dla obiektu A1, horyzont 2 min, `mse_bce`. N_T — liczba tranzycji. 76

4.11 Wskaźniki bezpieczeństwa dla obiektu A1, horyzont 2 min, `mse_bce`. Symbol „—” oznacza brak zdarzeń, dla których można wyznaczyć wskaźnik. 76
4.12 Wyniki dla obiektu A1 dla różnych horyzontów predykcji, `mse_bce`. . . 76
4.13 Zdarzenia w oknach decyzyjnych dla obiektu A2, horyzont 2 min, `mse_bce`. N_T — liczba tranzycji. 77
4.14 Wskaźniki bezpieczeństwa dla obiektu A2, horyzont 2 min, `mse_bce`. Symbol „—” oznacza brak zdarzeń, dla których można wyznaczyć wskaźnik. 78
4.15 Wyniki dla obiektu A2 dla różnych horyzontów predykcji, `mse_bce`. . . 78
4.16 Zdarzenia w oknach decyzyjnych dla obiektu A3, horyzont 2 min, `mse_bce`. N_T — liczba tranzycji. 79
4.17 Wskaźniki bezpieczeństwa dla obiektu A3, horyzont 2 min, `mse_bce`. Symbol „—” oznacza brak zdarzeń, dla których można wyznaczyć wskaźnik. 79
4.18 Wyniki dla obiektu A3 dla różnych horyzontów predykcji, `mse_bce`. . . 80
4.19 Zdarzenia w oknach decyzyjnych dla obiektu A4, horyzont 2 min, `mse_bce`. N_T — liczba tranzycji. 81
4.20 Wskaźniki bezpieczeństwa dla obiektu A4, horyzont 2 min, `mse_bce`. Symbol „—” oznacza brak zdarzeń, dla których można wyznaczyć wskaźnik. 81
4.21 Wyniki dla obiektu A4 dla różnych horyzontów predykcji, `mse_bce`. . . 81
4.22 Zdarzenia w oknach decyzyjnych dla obiektu B1, horyzont 2 min, `mse_bce`. N_T — liczba tranzycji. 83
4.23 Wskaźniki bezpieczeństwa dla obiektu B1, horyzont 2 min, `mse_bce`. Symbol „—” oznacza brak zdarzeń, dla których można wyznaczyć wskaźnik. 83
4.24 Wyniki dla obiektu B1 dla różnych horyzontów predykcji, `mse_bce`. . . 84
4.25 Zdarzenia w oknach decyzyjnych dla obiektu B2, horyzont 2 min, `mse_bce`. N_T — liczba tranzycji. 85
4.26 Wskaźniki bezpieczeństwa dla obiektu B2, horyzont 2 min, `mse_bce`. Symbol „—” oznacza brak zdarzeń, dla których można wyznaczyć wskaźnik. 85
4.27 Wyniki dla obiektu B2 dla różnych horyzontów predykcji, `mse_bce`. . . 85
4.28 Zdarzenia w oknach decyzyjnych dla obiektu B3, horyzont 2 min, `mse_bce`. N_T — liczba tranzycji. 86
4.29 Wskaźniki bezpieczeństwa dla obiektu B3, horyzont 2 min, `mse_bce`. Symbol „—” oznacza brak zdarzeń, dla których można wyznaczyć wskaźnik. 87

4.30 Wyniki dla obiektu B3 dla różnych horyzontów predykcji, `mse_bce`. . . 87
4.31 Zdarzenia w oknach decyzyjnych dla obiektu B4, horyzont 2 min, `mse_bce`. N_T — liczba tranzycji. 88
4.32 Wskaźniki bezpieczeństwa dla obiektu B4, horyzont 2 min, `mse_bce`. Symbol „—” oznacza brak zdarzeń, dla których można wyznaczyć wskaźnik. 88
4.33 Wyniki dla obiektu B4 dla różnych horyzontów predykcji, `mse_bce`. . . 89
4.34 Zdarzenia w oknach decyzyjnych dla obiektu C1, horyzont 2 min, `mse_bce`. N_T — liczba tranzycji. 90
4.35 Wskaźniki bezpieczeństwa dla obiektu C1, horyzont 2 min, `mse_bce`. Symbol „—” oznacza brak zdarzeń, dla których można wyznaczyć wskaźnik. 91
4.36 Wyniki dla obiektu C1 dla różnych horyzontów predykcji, `mse_bce`. . . 91
4.37 Zdarzenia w oknach decyzyjnych dla obiektu C2, horyzont 2 min, `mse_bce`. N_T — liczba tranzycji. 92
4.38 Wskaźniki bezpieczeństwa dla obiektu C2, horyzont 2 min, `mse_bce`. Symbol „—” oznacza brak zdarzeń, dla których można wyznaczyć wskaźnik. 92
4.39 Wyniki dla obiektu C2 dla różnych horyzontów predykcji, `mse_bce`. . . 92
4.40 Zdarzenia w oknach decyzyjnych dla obiektu C3, horyzont 2 min, `mse_bce`. N_T — liczba tranzycji. 93
4.41 Wskaźniki bezpieczeństwa dla obiektu C3, horyzont 2 min, `mse_bce`. Symbol „—” oznacza brak zdarzeń, dla których można wyznaczyć wskaźnik. 94
4.42 Wyniki dla obiektu C3 dla różnych horyzontów predykcji, `mse_bce`. . . 94
4.43 Zdarzenia w oknach decyzyjnych dla obiektu C4, horyzont 2 min, `mse_bce`. N_T — liczba tranzycji. 95
4.44 Wskaźniki bezpieczeństwa dla obiektu C4, horyzont 2 min, `mse_bce`. Symbol „—” oznacza brak zdarzeń, dla których można wyznaczyć wskaźnik. 95
4.45 Wyniki dla obiektu C4 dla różnych horyzontów predykcji, `mse_bce`. . . 96
4.46 Wskaźnik S_{FN} [%] dla wszystkich obiektów i horyzontów predykcji, `mse_bce`. Pogrubiono najniższą wartość dla każdego obiektu. 97

5.1 Charakterystyka wdrożeń produkcyjnych. N — liczba sprężarek w układzie, h — horyzont predykcji, strojenie — optymalizacja bayesowska hiperparametrów. 106

5.2 Wskaźniki bezpieczeństwa modeli produkcyjnych (ocena OOT). N_W — liczba okien decyzyjnych, N_R — liczba powrotów obciążenia ($TP+FN$), N_O — liczba okien bez powrotu ($TN+FP$). Dla D3 ($T_s = 5$ s) oceniono dobę z tygodnia uczącego (sekcja 5.9). 107
5.3 Wyniki nadzoru operacyjnego. N_D — decyzje bez powrotu obciążenia, N_B — zablokowane decyzje (w nawiasie udział w sumie $N_B + N_D$), FN_{op} — błędne decyzje, D/dz — średnia liczba decyzji na dobę. S_{FN} z tabeli 5.2 liczono w oknie 2 min, a R_{FN} w oknie 30 s. 110
5.4 Procentowy udział biegu jałowego w czasie pracy sprężarek. 113
5.5 Udział biegu jałowego w okresie bazowym i po wdrożeniu oraz jego względna zmiana. 114
5.6 Tygodniowy przebieg wskaźnika R_{FN} dla obiektu D3 w pierwszych czterech pełnych tygodniach pracy. W pozostałych dniach nie wystąpiły błędy FN_{op}. 117
5.7 Konfiguracja mechanizmu dotrenowania, nadzór operacyjny oraz oszczędności energii (średnia liczona z pominięciem miesiąca M_0). 120
5.8 Metryki bezpieczeństwa najlepszego modelu obiektu D3 w konfiguracji bazowej ($T_s = 20$ s, horyzont 2 min, funkcja `mse_bce`). 123
5.9 Mediany czasu obciążenia i biegu jałowego, liczba okresów biegu jałowego na dobę oraz wskaźnik $\tau_{<h}$ obiektu D3 dla danych 5 s, tych samych danych przeskalowanych do 20 s (sym.) oraz danych archiwalnych 20 s z wcześniejszego okresu (arch.). Symbolem * oznaczono doby, w których SP2 była sprężarką wiodącą. 125
5.10 Rozkład długości okresów obciążenia i biegu luzem [%] obiektu D3 (dane 5 s z tygodnia uczącego). Pogrubiono przedziały o największym udziale. 127
5.11 Konfiguracja modelu produkcyjnego obiektu D3 po optymalizacji bayesowskiej. 128

Spis skrótów

ARIMA	autoregresyjny zintegrowany model średniej ruchomej (ang. *AutoRegressive Integrat*
BCE	binarna entropia krzyżowa (ang. *Binary Cross-Entropy*)
BiLSTM	dwukierunkowa sieć z komórkami pamięci krótkotrwałej (ang. *Bidirectional Long Short-Term Memory*)
CUSUM	skumulowana suma odchyleń (ang. *Cumulative Sum*)
CPS	system cyber-fizyczny (ang. *Cyber-Physical Systems*)
DCS	rozproszony system sterowania (ang. *Distributed Control System*)
FN	wynik fałszywie negatywny (ang. *False Negative*)
FP	wynik fałszywie pozytywny (ang. *False Positive*)
FL	funkcja straty ogniskowej (ang. *Focal Loss*)
GRU	bramkowana jednostka rekurencyjna (ang. *Gated Recurrent Unit*)
HMI	interfejs operatorski (ang. *Human-Machine Interface*)
IT	warstwa informatyczna (ang. *Information Technology*)
LIME	lokalne wyjaśnienia modelu niezależne od jego typu (ang. *Local Interpretable Model-agnostic Explanations*)
LSTM	sieć z komórkami pamięci długo—krótkotrwałej (ang. *Long Short-Term Memory*)
ML	uczenie maszynowe (ang. *Machine Learning*)
MSE	błąd średniokwadratowy (ang. *Mean Squared Error*)
OT	warstwa operacyjna (ang. *Operational Technology*)
OOT	walidacja na danych z okresu późniejszego (ang. *out-of-time*)
PLC	programowalny sterownik logiczny (ang. *Programmable Logic Controller*)
PI	regulator proporcjonalno-całkujący
PID	regulator proporcjonalno-całkująco-różniczkujący
RMSE	pierwiastek błędu średniokwadratowego (ang. *Root Mean Squared Error*)
RNN	rekurencyjna sieć neuronowa (ang. *Recurrent Neural Network*)
RTU	zdalna jednostka końcowa (ang. *Remote Terminal Unit*)
SCADA	nadrzędny system sterowania i przetwarzania danych pomiarowych

	(ang. *Supervisory Control and Data Acquisition*)
SI	sztuczna inteligencja (ang. *Artificial Intelligence*)
SHAP	addytywne wyjaśnienia oparte na wartościach Shapleya (ang. *SHapley Additive exPlanations*)
SPn	oznaczenie n-tej sprężarki w układzie
TN	wynik prawdziwie negatywny (ang. *True Negative*)
TP	wynik prawdziwie pozytywny (ang. *True Positive*)
TPE	estymator Parzena o strukturze drzewiastej (ang. *Tree-structured Parzen Estimator*)
XAI	sztuczna inteligencja możliwa do wyjaśnienia (ang. *Explainable Artificial Intelligence*)
VFD	przemiennik częstotliwości (ang. *Variable Frequency Drive*)
LINEX	liniowo-wykładnicza funkcja straty (ang. *Linear-Exponential*)
MPC	sterowanie predykcyjne oparte na modelu (ang. *Model Predictive Control*)
PRC	krzywa precyzja–czułość (ang. *Precision-Recall Curve*)
ROC	charakterystyka operacyjna odbiornika (ang. *Receiver Operating Characteristic*)
RUL	pozostały okres eksploatacji (ang. *Remaining Useful Life*)

Spis symboli

$p(t)$	ciśnienie w sieci sprężonego powietrza w chwili t	[bar]
$p_{\min}$	próg obciążenia maszyny	[bar]
$p_{\max}$	próg odciążenia maszyny	[bar]
$p_{f,\min}$	dolna granica filtracji odczytów ciśnienia	[bar]
$p_{f,\max}$	górna granica filtracji odczytów ciśnienia	[bar]
$Q_{\text{out}}(t)$	zapotrzebowanie na sprężone powietrze (zakłócenie)	[Nm^3/h]
$d(k)$	komenda sterownika nadrzędnego w kroku k	[–]
$\Delta p(t)$	przyrost ciśnienia między kolejnymi próbkami	[bar]
$c_i(t)$	stan obciążenia i-tej sprężarki w chwili t	[–]
$\hat{c}_i(t)$	przewidywany stan obciążenia i-tej sprężarki	[–]
$\hat{b}$	decyzja modelu po przekształceniu progowym	[–]
T_s	okres próbkowania danych	[s]
L	długość okna obserwacji (ang. *lookback*)	[próbki]
h	horyzont predykcji (długość prognozy sieci)	[próbki]
W	okno decyzyjne, na którym wyznaczana jest ocena decyzji ($W \leq h$)	[próbki]
N	liczba sprężarek w układzie	[–]
N_{VFD}	liczba sprężarek z przemiennikiem częstotliwości	[–]
θ	parametry modelu	[–]
u_1, u_2	liczba jednostek w kolejnych warstwach rekurencyjnych	[–]
p_{drop}	współczynnik porzucania (ang. *dropout*)	[–]
η	współczynnik uczenia (ang. *learning rate*)	[–]
M	liczebność wsadu uczącego (ang. *batch size*)	[próbki]
$\mathcal{L}$	funkcja straty	[–]
$\mathcal{L}_{MSE}$	składowa regresyjna funkcji straty	[–]
$\mathcal{L}_{BCE}$	składowa klasyfikacyjna funkcji straty	[–]
$\mathcal{L}_{reload}$	składowa wykrywania powrotu obciążenia	[–]
$\mathcal{L}_{trend}$	składowa trendu ciśnienia	[–]
w_p	waga składowej regresyjnej funkcji straty	[–]
w_b	waga składowej klasyfikacyjnej funkcji straty	[–]
w_r	waga składowej wykrywania powrotu obciążenia	[–]
w_t	waga składowej trendu ciśnienia	[–]
α	mnożnik kary za błąd fałszywie negatywny	[–]

β	mnożnik kary za błąd fałszywie pozytywny	[–]
$R_{t,\tau,i}$	wskaźnik rzeczywistego powrotu obciążenia	[–]
$D_{t,\tau}$	wskaźnik opadania ciśnienia w horyzoncie predykcji	[–]
C_{FN}	koszt procesowy błędu fałszywie negatywnego	[–]
C_{FP}	koszt procesowy błędu fałszywie pozytywnego	[–]
C_{TP}	koszt procesowy trafnej predykcji obciążenia	[–]
C_{TN}	koszt procesowy trafnej rekomendacji wyłączenia	[–]
θ_q	waga kwantyla w asymetrycznej funkcji straty	[–]
W	okno decyzyjne wokół tranzycji	[próbki]
g	wskaźnik rzeczywistego powrotu obciążenia w oknie	[–]
$\hat{g}$	wskaźnik przewidywanego powrotu obciążenia w oknie	[–]
N_T	liczba wykrytych tranzycji	[–]
N_W	liczba okien decyzyjnych	[–]
N_R	liczba zdarzeń ryzykownych (klasa $TP + FN$)	[–]
N_O	liczba okien umożliwiających wyłączenie (klasa $TN + FP$)	[–]
N_L	liczba tranzycji obciążenia	[–]
μ	średnia długość cykli	[s]
$\tilde{x}$	mediana długości cykli	[s]
σ	odchylenie standardowe długości cykli	[s]
CV	współczynnik zmienności (σ/μ)	[–]
CV_{obc}	współczynnik zmienności okresów obciążenia	[–]
CV_{luz}	współczynnik zmienności okresów biegu jałowego	[–]
γ	współczynnik skośności rozkładu długości okresów	[–]
γ_{luz}	skośność rozkładu okresów biegu jałowego	[–]
$\mu/\tilde{x}$	stosunek średniej do mediany długości cykli	[–]
$\tau_{<h}$	odsetek okresów biegu jałowego krótszych od horyzontu	[%]
S	skuteczność decyzji w oknie decyzyjnym	[%]
S_{FN}	odsetek niewykrytych powrotów obciążenia	[%]
S_{FP}	odsetek decyzji nadmiarowo ostrożnych	[%]
M_U	dopasowanie dokładne dla tranzycji odciążenia	[%]
M_L	dopasowanie dokładne dla tranzycji obciążenia	[%]
R_{FN}	odsetek błędnych decyzji w eksploatacji	[%]
N_D	liczba rekomendacji wyłączenia, po których nie nastąpił powrót obciążenia	[–]
FN_{op}	liczba błędnych decyzji operacyjnych	[–]
N_B	liczba rekomendacji wstrzymanych przez warstwę zabezpieczeń	[–]
M_0	miesiąc uruchomienia systemu	[–]
M_{-k}	k-ty miesiąc przed uruchomieniem systemu	[–]
M_{+k}	k-ty miesiąc eksploatacji	[–]

1 Wprowadzenie

W rozdziale opisano kontekst badań w perspektywie coraz większych wymagań dotyczących optymalizacji energetycznej oraz Przemysłu 4.0, gdzie już dostępne i archiwizowane dane zostaną wykorzystane w nowy sposób. Przedstawiono problematykę układów sprężonego powietrza, problem badawczy wraz z modelem rozpatrywanego systemu sterowania. Rozdział zawiera też cele pracy, tezy badawcze, odnosi się do zakresu oraz ograniczeń rozprawy. Na końcu rozdziału znajduje się podsumowanie struktury rozprawy.

1.1 Kontekst badań: Przemysł 4.0 i optymalizacja energetyczna

Sektor przemysłu jest jednym z największych odbiorców energii elektrycznej w Unii Europejskiej odpowiadając za 24,6% całkowitego zużycia energii w 2023 roku. Za prawie dwie trzecie końcowego zużycia energii w tym sektorze odpowiada energia elektryczna (32,6%) oraz gaz ziemny (31,3%), podczas gdy stałe paliwa kopalne mają swój udział na poziomie 6% [1].

Dyrektywa Parlamentu Europejskiego i Rady (UE) 2023/1791 [2] definiuje m.in. cele w zakresie efektywności energetycznej, zobowiązując państwa członkowskie do zmniejszenia zużycia energii do roku 2030 o co najmniej 11,7% w porównaniu z prognozami z 2020 roku (zużycie energii końcowej w Unii nie może przekroczyć 763 Mtoe). Dyrektywa wprowadza m.in. zasadę „efektywność energetyczna przede wszystkim", która ma przyczynić się do rozwoju Unii m.in. dzięki nowoczesnej, zasobooszczędnej oraz konkurencyjnej gospodarce. Zasada ta powinna być traktowana jako pierwszy wariant przy podejmowaniu decyzji politycznych i inwestycyjnych we wszystkich sektorach.

W świetle dyrektywy zgodnie z zaleceniami komisji [3], małe i średnie przedsiębiorstwa powinny być zachęcane przez państwa członkowskie do poddania się audytom energetycznym i wdrażania zaleceń z nich wynikających. Dla dużych przedsiębiorstw, czyli takich dla których średnie roczne zużycie energii przekracza 85 TJ (ok. 23 611 MWh) w ciągu ostatnich trzech lat, wymagane jest wdrożenie systemu zarządzania energią (do 11 października 2027 roku). W tym celu zalecane jest zastosowanie nowoczesnych technologii takich jak liczniki odzwierciedlające rzeczywiste zużycie danego medium, umożliwiając zdalny odczyt. Regularne audyty energetyczne, a co

za tym idzie, świadomość tego, jak zużywana jest energia, pozwala na potencjalne znaczne oszczędności energii. Najważniejszą normą, jaka pojawia się w tym kontekście, to ISO 50001 [4] systemy zarządzania energią, które przedstawiają przeglądy energetyczne jako fundament cyklu „zaplanuj-wykonaj-sprawdź-działaj". Jednocześnie norma dotycząca audytów energetycznych ISO 50002 [5] oraz norma EN 16247-1 [6] nie są wymagane do certyfikacji systemu zarządzania energią, stanowią jednak praktyczne narzędzie metodologiczne. Od strony automatyki logicznym jest zatem, że budując system przemysłowy, należy zapewnić odpowiedni monitoring parametrów procesu a w szczególności wyposażyć system energetyczny w odpowiednie liczniki od strony przedsiębiorstwa, na których najłatwiej będzie można wykazać oszczędności energii.

Kontekst Przemysłu 4.0 [7], [8] w niniejszej rozprawie odzwierciedlony jest przez wykorzystanie danych zbieranych w procesie produkcji sprężonego powietrza w nowy sposób służący do predykcyjnego sterowania zespołem sprężarek, a co za tym idzie uzyskania efektu optymalizacji pracy sprężarkowni z zapewnieniem większej efektywności energetycznej. Dane, których akwizycja i przetwarzanie są realizowane przez systemy SCADA, zostają wykorzystane do wsparcia pracy lokalnych sterowników nadrzędnych sprężarkowni bez rozbudowy istniejącej infrastruktury. Takie rozwiązanie pozwala poniekąd na uzyskanie elementów inteligencji w zakresie realizowanego zadania. Znaczenie proponowanego rozwiązania w kontekście Przemysłu 4.0 omówiono w sekcji 6.6, a genezę tej koncepcji w sekcji 2.2.1.

1.2 Problematyka układów sprężonego powietrza

W literaturze wskazuje się, że energia elektryczna zużywana na wytwarzanie oraz wykorzystanie sprężonego powietrza w zakładach produkcyjnych może stanowić około 10% zużycia energii elektrycznej całego sektora przemysłowego [9]. Udział ten jest jednak silnie zależny od profilu zastosowań sprężonego powietrza w danym zakładzie produkcyjnym. Opracowania branżowe podkreślają, że sprężone powietrze zwykle odpowiada za znaczącą część rachunku za energię elektryczną, wspominając, że w niektórych sektorach udział ten może być wyraźnie wyższy [10]. Źródła eksperckie z kolei zwracają uwagę, że określenie udziału sprężonego powietrza w poszczególnych gałęziach przemysłu jest trudne ze względu na brak wiarygodnych badań przekrojowych [11]. Z perspektywy ekonomicznej znaczenie tej pozycji jest ugruntowane w strukturze kosztów podczas cyklu życia instalacji: energia elektryczna stanowi dominującą część kosztów operacyjnych sprężarkowni (rzędu 70–80%), podczas gdy udział samego kosztu inwestycji jest relatywnie niski [11]. Podchodząc całościowo, można stwierdzić, że działania poprawiające efektywność energetyczną mają potencjał przynosić wymierne korzyści kosztowe w różnych gałęziach wykorzystania sprężonego powietrza, wpisując się w szersze cele ograniczania zużycia energii i emisji [12].

W zakładach przemysłowych można spotkać się z określeniem sprężonego powietrza jako *czwartego medium* [13] będącego fundamentem działania obok energii elektrycznej, wody i gazu ziemnego. Z doświadczeń własnych autora wynika, że w przemyśle wydobywczym dostępność sprężonego powietrza traktowana jest na równi z energią elektryczną czy wodą — jako medium, którego brak natychmiast wstrzymuje produkcję, a w skrajnych przypadkach może stanowić zagrożenie życia. Sprężone powietrze znajduje szerokie zastosowanie w prawie każdej gałęzi przemysłu: w przemyśle samochodowym może służyć do napędzania narzędzi pneumatycznych czy chwytaków trzymających części, w przemyśle spożywczym i farmaceutycznym do transportu, pakowania i procesów wymagających powietrza o wysokiej czystości, a także w przemyśle tekstylnym, chemicznym, papierniczym i elektronicznym [9]. Za wykorzystaniem sprężonego powietrza przemawia wysoka niezawodność przy zachowaniu prostoty przesyłu, łatwości magazynowania i akumulowania powietrza w zbiornikach buforowych, możliwość dopasowania do wielu zastosowań od prostych elementów pneumatycznych po potężne młoty kuźnicze [14].

1.3 Motywacja i uzasadnienie badań

Autor niniejszej rozprawy rozpoczął pracę w sektorze sprężonego powietrza w 2010 roku, posiadając wcześniejsze szerokie doświadczenie w zakresie automatyki przemysłowej. Wiedza zdobyta na uczelni została rozbudowana przez praktyczne doświadczenie nabyte przy projektowaniu kompletnych systemów sterowania, przechodząc od doboru poszczególnych komponentów, tworzenia dokumentacji elektrycznej, oprogramowania sterowników PLC, zaprojektowania i uruchomienia systemów bazodanowych, implementacji systemów SCADA jak i późniejszego ich utrzymania i analizy zebranych danych. Rozpoczynając pracę w sektorze sprężonego powietrza, autor zauważył, że występuje znacząca potrzeba rozbudowy istniejących systemów sterowania o inteligentniejsze niż wówczas dostępne algorytmy systemów sterowania, które mogłyby lepiej radzić sobie z wyzwaniami, jakie spotykane są na co dzień w przemyśle. O ile sterowniki PLC mogą być zaprogramowane do realizacji mniej lub bardziej skomplikowanych algorytmów sterowania to ich możliwości obliczeniowe są ograniczone, systemy SCADA również często używane są jako rodzaj wsparcia systemów sterowania jak choćby przechowywanie oraz zapisywanie receptur związanych z różnymi produktami, jednak nie oferowano wówczas rozwiązania, które potrafiłoby w sposób inteligentny przewidywać stan systemu sprężonego powietrza i na tej podstawie podejmować decyzje sterujące. Z czasem pojawił się Przemysł 4.0, który miał przenieść instalacje przemysłowe na wyższy poziom rozwoju. Jednym z popularnych zastosowań byłoby wprowadzenie algorytmów umożliwiających symulację pracy linii technologicznej, pozwalając na przygotowanie

oprogramowania sterującego, które będzie adekwatnie odzwierciedlać rzeczywiste warunki pracy i minimalizując czas rozruchu. Takie rozwiązania jednak bazują na procesie, który jest opisany sekwencyjnie i jego decyzje bazują często na zdarzeniach binarnych. W przypadku sprężonego powietrza, mamy do czynienia z systemami różnorodnymi, nieliniowymi, o różnej specyfice, znajdującej się w różnych warunkach pracy, które mogą znacząco różnić się pomiędzy zakładami. Obecnie wspomniane wcześniej systemy wizualizacji czy sterowniki PLC zostały rozbudowane o funkcje umożliwiające implementację algorytmów sztucznej inteligencji, dając użytkownikowi możliwość łatwiejszej implementacji takich rozwiązań lub łatwe sprawdzenie, czy dane rozwiązanie sprawdzi się w danym środowisku. Powszechna dostępność oraz rozwój technologii uczenia głębokiego zachęciły autora do podjęcia próby opracowania inteligentnego systemu sterowania sprężonego powietrza, który mógłby wspomagać istniejące systemy sterowania, minimalizując zużycie energii elektrycznej przy utrzymaniu stabilności ciśnienia. Autor jest przekonany, że istnieje duży potencjał w wykorzystaniu algorytmów uczenia głębokiego do optymalizacji pracy systemów sprężonego powietrza, przynosząc korzyści ekonomiczne, jak i środowiskowe. Efektywne sterowanie sprężonym powietrzem może prowadzić do redukcji emisji gazów cieplarnianych, obniżenia kosztów operacyjnych oraz zwiększenia konkurencyjności przedsiębiorstw.

W ramach niniejszej rozprawy podjęto się opracowania i wdrożenia inteligentnego systemu sterowania, wykorzystującego dwukierunkową architekturę LSTM (BiLSTM) oraz autorskie asymetryczne funkcje straty, dostosowane do specyfiki problemu minimalizacji energii elektrycznej przy zachowaniu ciągłości dostaw sprężonego powietrza. Do weryfikacji opracowanego rozwiązania wybrano 12 obiektów przemysłowych reprezentujących różnorodne warunki pracy, z czego wdrożenie produkcyjne zrealizowano na ośmiu obiektach (czterech pochodzących z badanego zbioru i czterech nowych wybranych na potrzeby wdrożenia). Największe oszczędności energii elektrycznej można osiągnąć przez jak najlepiej dobrane urządzenia, dokonanie zakupów jak najbardziej energooszczędnych sprężarek, jednak w praktyce często dane zakłady przemysłowe mogą dysponować już istniejącymi systemami sprężonego powietrza, które nie są zoptymalizowane pod kątem zużycia energii elektrycznej a ich modernizacja, może być zbyt kosztowna. W takim przypadku inteligentny system sterowania, który może zostać zintegrowany z już istniejącymi systemami sterowania, może przynieść znaczące oszczędności kosztowe przy bardzo niewielkiej inwestycji w system sterowania nadrzędnego wspierany przez algorytmy uczenia głębokiego. Zakres niniejszej rozprawy obejmuje zagadnienie predykcji stanów pracy sprężarek w sprężarkowni, stanowiące najistotniejszy element efektywnego sterowania układem sprężonego powietrza.

Przy pewnych ograniczeniach technologicznych zaproponowano rozwiązanie, które może być implementowane w szerszym gronie zakładów przemysłowych, o ile charakterystyka obiektu jest zbliżona do przypadków omówionych w niniejszej rozprawie.

System został przetestowany na rzeczywistych danych przemysłowych, a jego skuteczność przeanalizowano pod kątem różnych horyzontów predykcji oraz asymetrii kosztów błędów klasyfikacji.

1.4 Problem badawczy

Predykcja w sektorze sprężonego powietrza często bazuje na licznikach energii elektrycznej wykorzystując szereg wskaźników statystycznych takich jak np. CUSUM [14] gdzie przy stworzeniu odpowiednio długich okien czasowych kWh łatwiej obserwować odchyłki od wartości, które świadczą o prawidłowej pracy układu. Przy niewielkim parku maszynowym, opomiarowaniu istotnych fragmentów linii oraz powtarzalności produkcji taki sposób pozwala doskonale na znalezienie anomalii. W tym miejscu można zauważyć pewną analogię, że aby mieć punkt wyjścia dla uczenia sieci neuronowych, należy najpierw nauczyć się prawidłowo działającego systemu. Wypada w tym momencie wspomnieć, że przeprowadzone badania świadczą o tym, że nieprawidłowe działanie układu będzie utrudniać nauczenie się zachowania sprężarek, lub spowoduje, że będzie to niemożliwe [15].

Modelowanie innych urządzeń w sektorze powietrza bez sprężarnia np. dmuchaw, może zostać zrealizowane przy pomocy modeli znanych z teorii sterowania lub bazując na regresji liniowej przy identyfikacji poboru w zadanym okresie [16].

Przytoczone przykłady, o ile sprawdzają się w horyzontach statystycznych i okresach godzinowych, dobowych, miesięcznych przy przewidywaniu w czasie rzeczywistym mogą nie być adekwatne. W sektorze sprężonego powietrza można zetknąć się z regresją liniową oraz gradientem w roli popularnych wskaźników do przewidywania krótkotrwałego zapotrzebowania, jednak takie rozwiązania nie zawsze jest wystarczające, chociaż w jakimś stopniu może sprawdzić się jako dodatkowe zabezpieczenie działania sieci neuronowej.

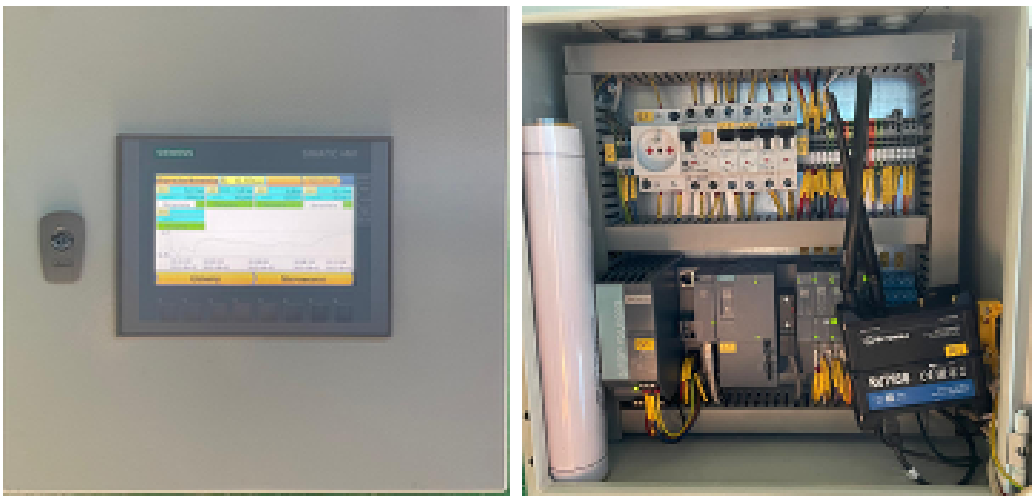


Rysunek 1.1: Sprężarkownia firmy Marani Sp. z o.o. składająca się z pięciu sprężarek śrubowych.

Tytułem wprowadzenia, należy wspomnieć, że sprężarka śrubowa to urządzenie, które jest przeznaczone do wytwarzania sprężonego powietrza. Jej sercem jest stopień śrubowy. Stopień zawiera dwie zestawione ze sobą śruby, które napędzane silnikiem elektrycznym, w wyniku obrotów doprowadzają do kompresji powietrza. Mnogość rodzajów maszyn mogłaby być porównywalna do różnorodności modeli samochodów, gdzie każdy producent będzie miał pewne autorskie rozwiązania, ale zasada działania pozostaje ta sama. Ponieważ zapotrzebowanie na sprężone powietrze może się znacznie zmieniać, zamiast kupować jedną potężną sprężarkę, zestawia się ze sobą kilka maszyn dostarczających powietrze do wspólnego kolektora, dzięki czemu można osiągnąć lepsze dopasowanie do zapotrzebowania. Zespół sprężarek zasilający dany zakład przemysłowy to sprężarkownia. Na rysunku 1.1 przedstawiono przykładową sprężarkownię zbudowaną z pięciu sprężarek śrubowych zrealizowaną przez Marani Sp. z o.o. [17].

Zalecane wytyczne [10], [13] rekomendują synchronizację pracy poszczególnych sprężarek, aby zapewnić jak najbardziej efektywny profil energetyczny. Taki cel zostaje osiągnięty przez wprowadzenie nadrzędnego systemu sterowania, który nadzoruje pracę wszystkich maszyn sprężarkowni realizując wspomniane zalecenia [15]. Na rysunku 1.2 przedstawiono front oraz wnętrze zaprojektowanej szafki realizującej obiektowe sterowanie bazujące na sterowniku PLC. Samo wprowadzenie pomiaru ciśnienia, w miejscu, które odzwierciedla rzeczywiste zapotrzebowanie na sprężone powietrze (zarówno pod względem pomiaru ciśnienia, jak i przepływu powietrza), pozwala na stabilizację, minimalizację liczby pracujących maszyn a co za tym idzie oszczędności energii [11]. Wystarczy wyobrazić sobie układ, który zawiera sześć maszyn, gdzie każda maszyna pracuje na podstawie swojego lokalnego pomiaru ciśnienia oraz lokalnych nastaw: w takim przypadku trudno uniknąć sytuacji niepotrzebnych załączeń czy zakłóceń pracy

pomiędzy maszynami wyposażonymi w przetwornice częstotliwości (gdzie mogą wzajemnie reagować na swoje zmiany prędkości obrotowej silnika głównego i destabilizować układ reagując zbyt agresywnie lub zbyt wolno na zapotrzebowanie). Z wielu typów sprężarek, jakie są dostępne w przemyśle [13], [14] niniejsza rozprawa będzie skupiać się na układach sprężarek śrubowych w związku z rodzajem instalacji, z jakimi autor ma do czynienia na co dzień, na jakich przeprowadzono badania.



Rysunek 1.2: Szafka sterownika nadrzędnego bazująca na sterowniku PLC.

Redukcja ciśnienia roboczego o 1 bar pozwala na osiągnięcie oszczędności energii rzędu 7–8% [13], co czyni precyzyjne utrzymanie minimalnego wymaganego ciśnienia bardzo istotnym elementem zapewnienia efektywności energetycznej sprężarkowni. Zapewnienie efektywności umożliwia zaprojektowanie układu w odpowiedni sposób [13]. Dany zakład może charakteryzować mniejsze lub większe zapotrzebowanie na sprężone powietrze, a przebieg pomiaru ciśnienia będzie świadczył o jego charakterze pracy [18], [19]. W przypadku zakładu, gdzie występuje kilka dużych maszyn wymagających dużych ilości powietrza, charakterystyka będzie inna niż dla odbiorcy realizującego regularną produkcję detali. Produkcję może charakteryzować pewna regularność, mogą zdarzać się okresy zmniejszonego poboru (np. przerwa w pracy), w weekendy maszyny mogą pracować tylko na utrzymanie ciśnienia w instalacji lub być wyłączane. W związku z różnym zapotrzebowaniem, przed hipotetyczną fabryką będą stały zróżnicowane wyzwania zarówno od strony instalacyjnej, kosztowej, jak i sterowniczej.

Przy projektowaniu sprężarkowni poza doborem odpowiedniego rodzaju maszyn należy zapewnić również rezerwę ruchową, która umożliwi bezuderzeniowe przeprowadzanie serwisu maszyn lub będzie stanowić zabezpieczenie ciągłości produkcji w przypadku awarii. Innym istotnym aspektem jest dobór dodatkowego osprzętu związanego z klasą czystości powietrza przeznaczoną w wybranej gałęzi przemysłu (inna klasa jest przewidziana dla przemysłu farmaceutycznego a inna dla przemysłu ciężkiego) [20].

W przypadku rozpatrywanych obiektów powietrze będzie spełniało wymogi dla zastosowania ogólnego powietrza w produkcji zgodnie z ISO 8573-1:2010 1.4.2 z punktem rosy na poziomie +3°C. Bardzo istotną kwestią jest również użycie zbiorników buforowych w których można akumulować zmagazynowane powietrze, ich odpowiedni dobór oraz liczba w przypadku większych wahań zapotrzebowania pozwala na złagodzenie charakterystyki ciśnienia dając systemowi nadrzędnemu czas na reakcję. W układzie sprężonego powietrza poza filtrami pojawiają się również osuszacze. Samo działanie osuszacza może również wpływać na pracę całego systemu. W przypadku awarii i np. zamarzania osuszacza można doprowadzić do ograniczenia przepływu powietrza przez osuszacz.

Przez ostatnie dziesięciolecie nastąpiły również znaczne zmiany w budowie samych sprężarek śrubowych [21]. Cofając się o dwadzieścia–trzydzieści lat na rynku dominowały sprężarki stałoobrotowe lub wyposażone w przetwornice częstotliwości bazujące na silnikach asynchronicznych. Mogły być również jedno lub dwustopniowe. Takie maszyny miały jednak swoje ograniczenia jak np. liczbę rozruchów na godzinę, w skrajnych przypadkach mając ich tylko sześć. W przypadku maszyn wyposażonych w przetwornicę częstotliwości często nie przekraczano 50 Hz jednak było to wystarczające dla ówczesnych rozwiązań. Obecnie na rynku coraz łatwiej dostępne są maszyny bazujące na silnikach synchronicznych z magnesami trwałymi, które nie mają ograniczeń (poza czasem zrzutu ciśnienia po zatrzymaniu) liczby rozruchów na godzinę. Ponieważ te typy sprężarek często wymagają do pracy oleju (chociaż należy wspomnieć, że jak najbardziej w użyciu są sprężarki bezolejowe) starając się zapewnić jak najmniejsze wymagania serwisowe na rynku pojawiły się sprężarki odśrodkowe z łożyskami magnetycznymi, w których wał lewituje w polu magnetycznym (takie rozwiązanie wg deklaracji producenta zapewnia redukcję zużycia energii elektrycznej o 15–20% w porównaniu do sprężarek śrubowych bezolejowych) [22].

Powszechne zastosowania predykcyjne modeli uczenia głębokiego bazują na licznikach lub modelowaniu przepływu sprężonego powietrza [23]. Popularnym rozwiązaniem predykcyjnym jest wykorzystanie sztucznej inteligencji do krótkoterminowego prognozowania zapotrzebowania na energię elektryczną u odbiorców indywidualnych [24]. Model wytrenowany na historii poboru przewiduje obciążenie sieci elektrycznej w horyzoncie najbliższych godzin.

W przypadku rozpatrywanych układów sprężonego powietrza w rozprawie zbadano czy na bazie zebranych danych można przewidywać przyszłe zapotrzebowanie w ograniczonym horyzoncie czasowym i czy decyzje podejmowane na jego bazie będą stabilizować czy destabilizować pomiar ciśnienia na sprężarkowni (co sprawdzono przy wdrożeniu).

Rola sterownika nadrzędnego sprężarkowni jest reaktywna, sterownik musi reagować na spadek lub wzrost ciśnienia sieci przez zapewnienie odpowiedniej liczby maszyn

dostarczających sprężone powietrze. O ile w przypadku rozpatrywania całego układu w horyzontach czasowych kilkunastominutowych, godzinowych, dobowych bez przełożenia na decyzje sterujące można się nie spotkać z karą za nieprawidłową predykcję, to w przypadku błędnej decyzji sterującej można doprowadzić do spadku ciśnienia. Zbyt długotrwały spadek ciśnienia przy dużym zapotrzebowaniu prowadzi do oczekiwania na czas rozprężenia poszczególnych maszyn, podczas gdy ciśnienie operacyjne zakładu spada. Przy niskim poziomie ciśnienia, kolejne urządzenia zakładowe będą zatrzymywane a w ostateczności może dojść do zatrzymania całej produkcji. W tym miejscu należy również podkreślić, że sterownik nadrzędny jest niezbędny dla efektywnej energetycznie pracy sprężarkowni, a jego działanie bazuje często na zespole reguł bez wsparcia uczenia maszynowego, które jest następnym naturalnym etapem poruszonym w niniejszej rozprawie [25].

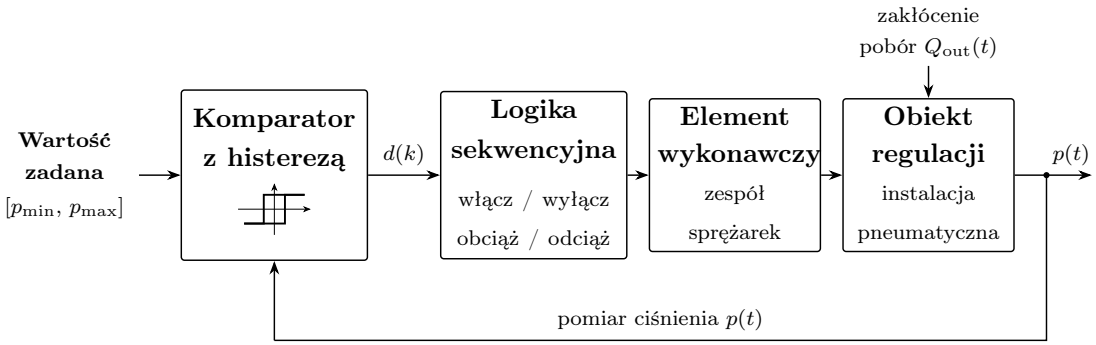
Na samą produkcję zakładową będzie również wpływać nie tylko rytm wyznaczany przez zmianowość czy daną gałąź przemysłu, ale również np. ilość zamówień w danym okresie, koszty energii elektrycznej (np. przeniesienie produkcji do takich okresów, gdy dostępna jest najtańsza energia elektryczna [26]) czy np. rodzaj produkowanych detali. Podobnie załączenia określonych typów urządzeń, które doskonale można obserwować przy użyciu wskaźników statystycznych, widoczne są na przebiegach ciśnienia (spadki ciśnienia w granicach pracy sterownika nadrzędnego bez naruszenia roboczych granic pracy).

Pojawia się zatem pytanie: czy w przypadku różnych zakładów, różnych typów produkcji można ten zmienny, a jednak charakterystyczny dla danego zakładu typ poboru wyodrębnić i opisać przy ograniczonych danych, jakimi dysponuje system nadrzędny? Czy można bazować na chwilowym przebiegu ciśnienia, a następnie na tej bazie przewidywać przyszłość? Czy wykorzystanie sieci neuronowych w tym aspekcie może pomóc, czy jej zastosowanie nie będzie odbiegać od regresji liniowej lub gradientu? Czy przy takim rozwiązaniu można zapewnić stabilność dostaw sprężonego powietrza dla strony klienckiej? Czy brak dostępu do liczników energii elektrycznej oraz pomiarów przepływu będzie uniemożliwiał predykcję?

Ostatnim aspektem, na jaki wypada zwrócić uwagę, jest fakt, że nawet doskonałe wyniki z badań muszą zostać zweryfikowane na rzeczywistej instalacji, ponieważ adekwatność przewidywania można potwierdzić dopiero po tym, gdy ta przewidywana chwila upłynie. Wyznacznikiem będzie w tym wypadku nie tylko brak istotnego spadku ciśnienia, ale przede wszystkim zadowolenie klienta.

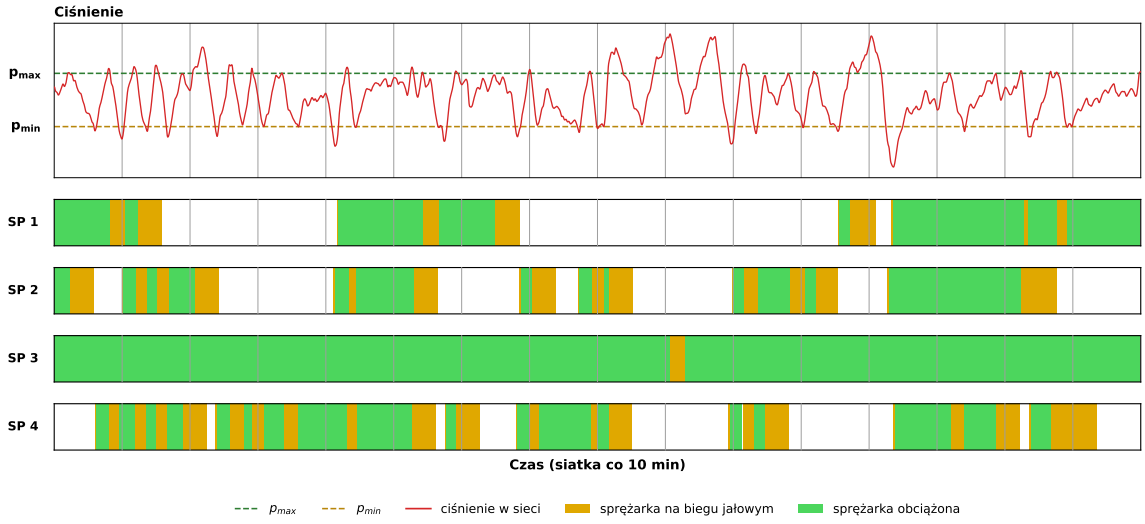
Zadanie sterowania sprężarkami wchodzącymi w skład sprężarkowni zrealizowane jest przez sterownik nadrzędny, który zajmuje się sekwencjonowaniem ich pracy. Schemat sterowania przedstawiono na rysunku 1.3. Pomiar aktualnego ciśnienia $p(t)$ odbywa się w punkcie odzwierciedlającym rzeczywiste zapotrzebowanie. Następnie zmierzone ciśnienie jest porównywane przez komparator względem zakresu utworzonego przez

granice ciśnienia $[p_{\min}, p_{\max}]$, który wydaje decyzje $d(k) \in \{-1, 0, +1\}$ odpowiadające zmniejszeniu, brakowi reakcji lub zwiększeniu wydajności układu. Sterownik nadrzędny na bazie decyzji oraz własnej logiki realizuje włączenie, wyłączenie, obciążenie lub odciążenie danej sprężarki. Zapotrzebowanie na sprężone powietrze $Q_{\text{out}}(t)$ jest nieprzewidywalnym zakłóceniem wpływającym na prawidłowość pracy instalacji pneumatycznej zakładu. Sieć sprężonego powietrza zawiera w sobie rurociągi, jak i zbiorniki buforowe. Pomiar aktualnego ciśnienia $p(t)$ jest zatem sygnałem sprzężenia zwrotnego dla komparatora.



RYSUNEK 1.3: Struktura nadrzędnego sekwencyjnego sterowania zespołem sprężarek. Wartością zadaną jest zakres ciśnienia $[p_{\min}, p_{\max}]$, a komparator z histerezą generuje komendy dwupołożeniowe $d(k) \in \{-1, 0, +1\}$, bez aktywnej regulacji w zakresie granic.

Na rysunku 1.4 przedstawiono rzeczywisty przebieg ciśnienia wraz ze stanami pracy sprężarek śrubowych. Stan tłoczenia (obciążenia) oznaczony jest kolorem zielonym. Stan bez sprężania (bieg jałowy/bieg luzem) oznaczono kolorem ciemnożółtym. Kolor biały reprezentuje stan wyłączenia maszyny.



RYSUNEK 1.4: Przebieg ciśnienia i stanów pracy sprężarek rzeczywistej instalacji. Pominięto wartości ciśnienia oraz oznaczenia czasu.

Sterownik nadrzędny reaguje sekwencyjnie wyłącznie na przekroczenie progów $p_{\min}$ i $p_{\max}$. Pomiędzy wspomnianymi granicami ciśnienia ze strony sterownika nadrzędnego nie odbywa się regulacja. Sprężarka SP3 jest maszyną wiodącą.

1.5 Cele pracy

Celem rozprawy jest opracowanie i weryfikacja warstwy predykcyjnej wspomagającej istniejące systemy sterowania sprężarkowni. Proponowane rozwiązanie nie zastępuje sterownika nadrzędnego realizującego funkcje regulacyjne, lecz dostarcza mu inteligentnego wsparcia umożliwiającego podejmowanie wyprzedzających decyzji sterujących w celu minimalizacji zużycia energii elektrycznej przy zachowaniu ciągłości dostaw sprężonego powietrza.

Cele szczegółowe to:

1. opracowanie wytycznych dla doboru okresu próbkowania danych treningowych sieci neuronowej, wraz z metodyką ich przygotowania w sposób weryfikujący wierność odwzorowania procesu produkcyjnego,
2. opracowanie architektury sieci rekurencyjnej dedykowanej dla predykcji stanów obciążenia sprężarek śrubowych,
3. definicja indywidualnej funkcji straty odzwierciedlającej specyfikę problemu wraz ze sposobem doboru jej parametrów,
4. opracowanie wskaźników ewaluacyjnych informujących o bezpieczeństwie proponowanego rozwiązania,
5. analiza wpływu horyzontu predykcji na jakość decyzji sterujących,
6. opracowanie regułowej warstwy bezpieczeństwa gwarantującej stałość dostaw sprężonego powietrza wraz z mechanizmem nadzoru operacyjnego po wdrożeniu,
7. walidacja proponowanej metodyki na rzeczywistych danych przemysłowych oraz w warunkach praktycznego wdrożenia na obiektach przemysłowych.

1.6 Tezy badawcze

1. Predykcyjne wsparcie nadrzędnego procesu sterowania zespołem sprężarek z wykorzystaniem uczenia głębokiego obliguje do zindywidualizowanego podejścia, które wymaga doboru następujących elementów: okresu próbkowania zbioru uczącego (odzwierciedlającego dynamikę obiektu), funkcji straty (wraz z ustaleniem

jej parametrów oraz kosztów procesowych), horyzontu predykcji, poziomów reguł bezpieczeństwa (zabezpieczenie stabilności dostaw). Niedotrzymanie tych założeń prowadzi do wskaźników oceny nieodpowiadających rzeczywistemu procesowi, co przy wdrożeniu produkcyjnym skutkuje niską jakością decyzji wspierających.

2. Przydatność opracowanego modelu predykcyjnego wymaga wskaźników oceny bazujących na rzeczywistych, pojedynczych zdarzeniach decyzyjnych z ukierunkowaniem na koszt błędu procesowego. Klasyczne podejścia bazujące na ocenie wyznaczanej dla pojedynczych próbek, odchyłce charakterystyki przepływu ciśnienia lub przepływu sprężonego powietrza bazującej na błędzie średniokwadratowym nie oddają wiernie ryzyka wdrożeniowego (gdzie błąd zachwiania ciągłości dostaw jest większym kosztem niż ominięte możliwości oszczędności energii).

3. Zastosowanie dwukierunkowej sieci rekurencyjnej BiLSTM zapewnia lepszą zdolność modelowania procesu niż wykorzystanie rekurencyjnych sieci jednokierunkowych LSTM oraz GRU, jednak nie wyklucza możliwości ich zastosowania. Sieć dwukierunkowa pozwala na adekwatniejszą identyfikację sytuacji, gdy możliwe jest bezpieczne wyłączenie maszyny, która nie będzie obciążana w rozpatrywanym horyzoncie czasowym.

4. Zaproponowana metodyka zakłada wykorzystanie już istniejącej infrastruktury systemu sterowania nadrzędnego bazującej na sterowniku PLC, a zakres jej stosowalności jest możliwy do wyznaczenia przed wdrożeniem. Po wdrożeniu wymagany jest ciągły nadzór operacyjny dla utrzymania efektów (ciągła eksploatacja wiąże się z mechanizmami degradacji, które można wykryć z wyprzedzeniem).

1.7 Zakres i ograniczenia pracy

Praca obejmuje następujące zagadnienia:

- analityczny przegląd metod uczenia maszynowego i głębokiego w sterowaniu przemysłowym,
- teoretyczne podstawy rekurencyjnych sieci neuronowych LSTM i BiLSTM,
- opracowanie autorskiej asymetrycznej funkcji straty dostosowanej do specyfiki problemu,
- eksperymenty na danych rzeczywistych pochodzących z 12 obiektów przemysłowych,
- analizę wpływu horyzontu predykcji oraz architektury sieci na jakość predykcji,

- walidację systemu w rzeczywistych warunkach procesowych na obiektach wdrożeniowych.

Praca nie obejmuje następujących zagadnień:

- projektowania ani implementacji sprzętowej warstwy sterowania — system działa jako warstwa nadrzędna współpracująca z istniejącymi sterownikami PLC,
- analizy kosztów utrzymania i awarii sprężarek,
- bezpośredniego sterowania napędami sprężarek — system ogranicza się do warstwy predykcji i decyzji o wyłączeniu.

1.8 Struktura rozprawy

Rozprawa składa się z siedmiu rozdziałów, których kolejność odzwierciedla przebieg przeprowadzonych prac. W pierwszej kolejności zidentyfikowano problem, opracowano metodykę, sprawdzono ją na danych historycznych, a następnie przeprowadzono pełne wdrożenie produkcyjne na reprezentatywnych obiektach spośród wyznaczonych grup. Na końcu wyciągnięto wnioski z wielomiesięcznej (806 dni) eksploatacji.

Rozdział 1 wprowadza w problematykę układów sprężonego powietrza bazującego na sprężarkach śrubowych, przedstawia kontekst problemu w świetle narastających wymagań efektywności energetycznej w Unii Europejskiej oraz ograniczenia kaskadowych sterowników nadrzędnych. Rozdział zawiera również motywację stojącą za rozpoczęciem badań, sformułowanie problemu badawczego, cele oraz tezy rozprawy.

Rozdział 2 dokonuje przeglądu literatury w zakresie ewolucji przemysłowych systemów sterowania prowadzącej do Przemysłu 4.0 z uwzględnieniem technologii, jakich znajomość jest niezbędna dla zrozumienia architektury systemów sterowania i miejsca w nich projektowanego modułu wsparcia nadrzędnych systemów sterowania sprężonym powietrzem. Rozdział zawiera również opis architektury rekurencyjnych sieci neuronowych używanych przy modelowaniu szeregów czasowych w sektorach sprężonego powietrza i energetyce oraz metod uczenia biorących pod uwagę asymetryczne koszty. Na końcu rozdziału opisano zauważoną lukę badawczą oraz przedstawiono tabelaryczne porównanie cech wybranych prac z niniejszą rozprawą.

Rozdział 3 zawiera opis zaproponowanej metodyki, definicję zadania predykcji, sposobu przygotowania danych, proponowane architektury sieci, budowę funkcji straty. Opisuje również system wskaźników oceny wyznaczany dla zdarzeń decyzyjnych biorąc pod uwagę koszt błędnych decyzji w odniesieniu do procesu.

Rozdział 4 zawiera wyniki badań przeprowadzonych na dwunastu wybranych obiektach przemysłowych podzielonych na trzy grupy w zależności od budowy sprężarkowni. Zawiera również studium ablacji porównawczych architektur sieci neuronowych, analizę horyzontu predykcji i zbiorcze zestawienie rezultatów.

Rozdział 5 opisuje wdrożenie przeprowadzone na ośmiu obiektach, z których cztery należą do omówionych już układów, a cztery nowe dają możliwość sprawdzenia proponowanej metodyki. Opisano zasady działania modułu SI wspierającego decyzje, warstwę reguł bezpieczeństwa, nadzór operacyjny oraz wyniki uzyskane w trakcie wdrożenia. Poświęcono również fragment rozdziału na studium przypadku obiektu wymagającego zindywidualizowanego podejścia.

Rozdział 6 zawiera interpretację uzyskanych wyników w odniesieniu do poszczególnych składowych systemu, wnioski płynące ze wdrożenia, porównanie z pracami dostępnymi w literaturze z zaznaczeniem fragmentów autorskich.

Rozdział 7 zawiera podsumowanie pracy i weryfikację tez badawczych. Opisuje praktyczne zalecenia dla wdrożeń, jak i podejmuje dyskusję dotyczącą kierunków dalszych badań.

Do rozprawy dołączono załączniki A i B zawierające pseudokod podstawowych modułów systemu oraz proponowaną strukturę plików konfiguracyjnych.

2 Przegląd literatury

Rozdział zawiera przegląd literatury, dotyczący układów sprężonego powietrza w przemyśle uwzględniając stronę projektową, której zrozumienie jest kluczowe dla zrozumienia problemu badawczego. Opisuje kontekst historyczny, który jest ściśle powiązany z proponowanymi rozwiązaniami sterowania wykorzystującymi współpracę SCADA-/PLC/HMI oraz przechodzi do wykorzystania uczenia głębokiego w sterowaniu przemysłowym. Kolejnym poruszonym aspektem jest uczenie nastawione na koszt oraz opis funkcji strat w uczeniu maszynowym. Na końcu rozdziału opisano identyfikację luki badawczej.

2.1 Układy sprężonego powietrza w przemyśle

2.1.1 System sprężonego powietrza

W normie ISO 11011:2013 [27] zdefiniowano system sprężonego powietrza jako składający się z trzech podsystemów:

- zasilania, przekształcający energię źródłową w energię sprężonego powietrza,
- transmisji, zajmujący się rozprowadzeniem i dostarczeniem energii sprężonego powietrza do miejsca wykorzystania,
- zapotrzebowania, który zawiera w sobie wszystkich odbiorców sprężonego powietrza (w tym również uwzględnia bezproduktywne straty).

Sprężarkownia stanowi istotny element infrastruktury energetycznej zakładu, której zadaniem jest wytworzenie takiej ilości sprężonego powietrza, która umożliwia działalność produkcyjną. Zadanie to zrealizowane jest przez utrzymywanie w całej instalacji odbiorczej ciśnienia na wystarczająco wysokim poziomie. W skład sprężarkowni mogą wchodzić maszyny różnych rodzajów pochodzące od różnych producentów w różnych konfiguracjach (np. podstawowe zapotrzebowanie będzie zapewniała turbosprężarka a sprężarki śrubowe będą odpowiedzią na zmienne zapotrzebowanie), a ten dobór będzie się przekładał na efektywność wykorzystania energii elektrycznej oraz możliwości optymalizacji.

Norma ISO 11011:2013 podaje również metodykę oceny efektywności energetycznej, przykładowe metryki, szablony dotyczące wszystkich istotnych aspektów systemu

sprężonego powietrza gdzie znaczącą rolę odgrywa pomiar ciśnienia, przepływu sprężonego powietrza oraz mocy (zużycie energii elektrycznej). Te trzy składniki pozwalają przedsiębiorstwu na efektywną ocenę efektywności i wskazanie kierunków poprawy. W przypadku niniejszej rozprawy dane dotyczące przepływu oraz mocy pozostały po stronie odbiorczej, przez co w badaniach bazowano na pomiarze ciśnienia odzwierciedlającego rzeczywiste zapotrzebowanie. Nie rozpatrywano pomiarów ciśnienia poszczególnych maszyn jednak w praktyce przemysłowej całościowe podejście do sprężarkowni oraz narzucenie wymogu np. nieprzekraczania określonej liczby kWh na wyprodukowany 1 Nm^3 sprężonego powietrza [25] jest wystarczające.

W poprzednim rozdziale wspomniano o tym, że sprężone powietrze jest czwartym medium, jednak charakteryzuje je niższa sprawność przetwarzania energii. Szacuje się, że od 80% do 93% energii elektrycznej jest tracone w postaci ciepła [13]. W niniejszej rozprawie w związku z praktyką inżynierską oraz rozpatrywaną generacją maszyn przyjmuje się wartość 70%. Potencjał odzyskania tej energii szacowany jest na poziomie od 50% do 93% [13], a odzysk możliwy jest przez wymienniki ciepła, dogrzewanie pomieszczeń lub inne rozwiązania wykorzystujące ciepło odpadowe [28].

2.1.2 Dobór właściwych maszyn i ich ograniczenia

Pobór sprężonego powietrza w przemyśle może charakteryzować się znaczną zmiennością a jego charakter często powiązany jest z gałęzią przemysłu jakiej dotyczy. W przypadku odbiorców o stabilnym, jednostajnym zapotrzebowaniu najwyższą efektywność energetyczną (wyrażoną wspomnianym wcześniej współczynnikiem kWh/Nm^3) osiągają sprężarki odśrodkowe (turbosprężarki). Jedno z rozwiązań proponowanych przez Marani Sp. z o.o. zawiera sprężarki odśrodkowe wraz z chłodnią wentylatorową, gwarantując rzeczywisty współczynnik na poziomie ok. 85 kWh/Nm^3, które można zestawić ze współczynnikiem na poziomie ok. 120 kWh/Nm^3 [25] dla sprężarek śrubowych po optymalizacji sprężarkowni. Wadą turbosprężarek jest jednak brak możliwości redukcji poboru energii elektrycznej przy zmniejszonym zapotrzebowaniu na powietrze — nadmiar produkcji jest odprowadzany przez tłumik na zewnątrz przy niezmienionym poborze energii. Takie rozwiązanie w warunkach zmiennego zapotrzebowania prowadzi do znacznego pogorszenia wskaźnika efektywności energetycznej. Dodatkowym ograniczeniem jest dopuszczalna liczba rozruchów, wynosząca np. jedynie dwa na dobę, co wyklucza możliwość elastycznego sterowania tą maszyną w odpowiedzi na zmieniające się zapotrzebowanie. Niektórzy producenci zalecają wizję lokalną po każdym zatrzymaniu maszyny, aby upewnić się, że elementy wirujące zostały faktycznie zatrzymane. Również w tym wypadku trzeba się liczyć ze zjawiskiem pompażu, polegającym na gwałtownych i cyklicznych wahaniach przepływu i ciśnienia w sprężarce odśrodkowej

przy zbyt małym zapotrzebowaniu, co może prowadzić do uszkodzenia maszyny i jest jednym z kluczowych ograniczeń eksploatacyjnych tego typu urządzeń [29].

Na podstawie doświadczeń własnych autora, zarówno w zakresie projektowania systemów sterowania nadrzędnego, jak i ewaluacji istniejących instalacji zgodnie z normami ISO, zaobserwowano, że w polskim przemyśle często pojawia się zmienne zapotrzebowanie na sprężone powietrze mogące drastycznie różnić się w zależności od aktualnego rytmu produkcji. W okresach przestojów lub zmniejszonej produkcji praca jednej maszyny może okazać się wystarczająca, podczas gdy w szczytowych momentach konieczne będzie uruchomienie kilku sprężarek (w sposób dostosowany do charakteru danego zakładu oraz przewidzianych zasad przyłączenia do sieci elektrycznej). W takich warunkach powszechnie stosowane są sprężarki śrubowe, które dzięki modułowej budowie, sprawdzonej niezawodności, wieloletniemu doświadczeniu producentów oraz możliwości przeprowadzania remontów generalnych i pracy w układach wielomaszynowych stanowią dojrzałą i elastyczną odpowiedź na wymagania przemysłowych układów sprężonego powietrza.

Sprężarki śrubowe przy zapewnieniu odpowiednich warunków pracy, czyli wentylacji, temperatury otoczenia oraz przy przestrzeganiu planowanych przeglądów, oferują niezawodną pracę w trybie ciągłym. Aktywne chłodzenie (w przypadku sprężarek chłodzonych olejem, oczywiście są też rozwiązania bezolejowe determinowane wymaganiami dotyczącymi klasy czystości powietrza zgodnie z ISO 8573 [20]) pozwala na uniknięcie przestojów. Sprężarki chłodzone olejem są również najbardziej przystępne cenowo. W ostatnich latach w związku z poszukiwaniem energooszczędnych rozwiązań na rynku coraz częściej pojawiają się sprężarki z przetwornicą częstotliwości, która umożliwia najlepszą adaptację do zmiennego zapotrzebowania. W poprzednim dziesięcioleciu takie maszyny oferowane były z silnikami asynchronicznymi a obecnie coraz częściej wykorzystywane są konstrukcje bazujące na silnikach synchronicznych z magnesami trwałymi [30], [31]. W przypadku silników asynchronicznych liczba startów zależy od mocy silnika oraz czasu rozprężania, przy silnikach synchronicznych dopuszczalna liczba startów jest ograniczona czasem rozprężania (chociaż praktyka inżynierska tutaj zwraca uwagę, że należy zachować rozsądek i unikać sytuacji mogących ograniczyć życie sprężarki). Należy zwrócić uwagę, że przez czas rozprężania sprężarka ma zablokowaną możliwość startu. Maszyny z przetwornicą częstotliwości oferują bezuderzeniowy, płynny rozruch z niewielkimi prądami rozruchowymi co jest dużo korzystniejsze dla maszyny jak i dla rozdzielni ją zasilającej. Te usprawnienia jednak wiążą się z wyższą ceną i nie zawsze w przypadku klienta będzie to najbardziej opłacalne rozwiązanie (chociaż w niniejszej rozprawie wykazano, że dynamika takich układów przy dobrym poborze zostawia niewiele miejsca na poprawę). Adaptacja do zmiennego zapotrzebowania jest szczególnie istotna dla elementów wykonawczych wykorzystujących sprężone powietrze oraz podnosi jakość produkcji.

Użycie zaawansowanej konstrukcji sprężarki wymaga jej odpowiedniego zabezpieczenia od strony sterującej. Potrzebny jest stały monitoring kluczowych parametrów obiegu chłodzącego (temperatura oleju), parametrów ciśnienia (ciśnienie zewnętrzne i wewnętrzne), bardziej zaawansowane maszyny również oferują zabezpieczenia w postaci pomiarów prądów czy napięcia. Aby zaproponować tańsze rozwiązania dla klienta końcowego można się spotkać z maszynami w których przetwornica częstotliwości została wyposażona w sterownik sprężarkowy co pozwala na minimalizację kosztów budowy w związku z rezygnacją z dodatkowego sterownika. Różni producenci reprezentują również różne podejście co do zasad oraz algorytmów sterowania mimo, że podstawowa zasada działania pozostaje analogiczna. Aspekty te są często pomijane w praktyce, tymczasem ustalone przez producenta zasady sterowania sprężarkami bezpośrednio przekładają się zarówno na możliwe do uzyskania oszczędności energii, jak i na poziom bezpieczeństwa systemu produkcyjnego.

Z jakimi ograniczeniami, skutkującymi potencjalnymi oszczędnościami energii elektrycznej, należy się liczyć na poziomie pojedynczej sprężarki:

- rodzaj użytej konstrukcji będzie determinował dopuszczalną ilość startów na godzinę,
- konstrukcja będzie też wpływać na czas rozprężania po zatrzymaniu, który ma zapewnić łagodny kolejny rozruch,
- elektryczny układ rozruchowy — czy będzie bazował na rozruchu bezpośrednim tj. gwiazda-trójkąt czy będzie wyposażona w układ łagodnego rozruchu tzw. soft–start,
- czy sprężarka będzie wyposażona w przetwornicę częstotliwości,
- jak dobrane są nastawy regulatorów PID sprężarki wyposażonej w przetwornicę, jak szybko następuje adaptacja do zmienności,
- jak szybko po uruchomieniu sprężarki można dostarczać powietrze,
- filozofia współpracy z systemem nadrzędnym — czy sprężarka będzie bazowała na swoich lokalnych nastawach w trybie sterowania zdalnego, jakie komendy sterujące są przewidziane, jakie dostępne są możliwości,
- jakie specyficzne rozwiązania sterowania zdalnego są oferowane przez danego producenta,
- w jaki sposób będzie prowadzona wymiana danych z systemem nadrzędnym czy będzie zrealizowana przez protokoły komunikacyjne (Profinet, Profibus, Modbus RTU, Modbus TCP itd.) czy przez sygnały elektryczne binarne (jeśli tak to jaka

ich pula jest dostępna, czy istnieje tylko włącz/wyłącz czy dostępne jest również sterowanie obciążeniem),

- czy występują jakieś ograniczenia od strony programowej, które mogą powodować niedeterministyczną odpowiedź na komendy sterujące (np. brak statusu gotowości oraz potrzeba wprowadzenia dodatkowej logiki weryfikującej stan),

- wydajność maszyny podawana jest wg normy ISO 11011 [27] przez co często pojawia się niezrozumienie co do faktycznej wydajności sprężarki.

Wymienione ograniczenia, mimo że często nie ma o nich mowy w dokumentacji producenta, muszą zostać wzięte pod uwagę przy tworzeniu wydajnego i niezawodnego całościowego systemu sterowania. Sterownik nadrzędny musi dostarczyć funkcjonalność, która mogła zostać pominięta na poziomie sprężarki.

2.1.3 Struktura sprężarkowni

Sam mikroprocesorowy sterownik sprężarki nie ma informacji o tym, jakie ciśnienie występuje w punkcie największego zapotrzebowania w związku z czym, co potwierdza również norma ISO 11011:2013, konieczny jest sterownik nadrzędny, aby zsynchronizować pracę wielu maszyn. Różni producenci oferują różne rozwiązania w tym zakresie, jednak w przypadku niniejszej rozprawy zdecydowano się na rozwiązanie bazujące na sterowniku PLC. Takie rozwiązanie przy umiejętności programowania umożliwia nie tylko współpracę z maszynami należących do różnorakich producentów zgodnie z wcześniejszymi założeniami, ale również daleko idącą integrację z dodatkowym wyposażeniem, jakie może znaleźć się na sprężarkowni. W przypadku liczników energii jednak powszechną praktyką jest, że informacja o ich stanie pozostaje po stronie odbiorcy.

W skład sprężarkowni poza wybranym rodzajem sprężarek wchodzą również inne elementy, których zadaniem jest zapewnienie odpowiednich własności sprężonego powietrza na wyjściu. Niektóre z nich również mogą mieć wpływ na efektywność energetyczną całego systemu, a ich dobór powinien być dostosowany do charakterystyki zapotrzebowania oraz wymagań produkcyjnych. Do podstawowych elementów sprężarkowni należą:

- określona liczba sprężarek, która powinna zostać dobrana w sposób pozwalający na unikanie przestojów oraz zapewnienie warunków serwisowych przy elastycznym dostosowaniu do zmiennego zapotrzebowania, najlepiej w układach wielosprężarkowych jest posiadać przynajmniej jedną maszynę redundantną,

- układ uzdatniania powietrza - system separacji zanieczyszczeń (odolejanie, filtracja) oraz osuszania,

- instalacja dystrybucyjna, łącząca ze sobą poszczególne komponenty dostarczając sprężone powietrze na wyjście,
- zbiorniki buforowe, których niewłaściwy dobór, złe rozmieszczenie lub brak może prowadzić do niestabilności ciśnienia oraz pogorszenia efektywności energetycznej,
- w niektórych przypadkach również urządzenia pomiarowe, takie jak przepływomierze, analizatory parametrów sieci elektrycznej czy liczniki energii, które mogą być wykorzystywane do monitorowania i optymalizacji pracy systemu zgodnie z wymaganiami komisji europejskiej dotyczącymi efektywności energetycznej,
- wyposażenie służące do odzysku ciepła odpadowego, które może być wykorzystywane do ogrzewania mediów (np. wody) lub dogrzewania pomieszczeń (przez systemy wentylacji) co pozwala na zwiększenie całkowitej efektywności energetycznej systemu sprężonego powietrza,
- system sterowania nadrzędnego integrujący wszystkie elementy sprężarkowni, które mają taką możliwość umożliwiający optymalizację pracy całego systemu, a nie tylko poszczególnych maszyn.

Najlepszy system nadrzędnego sterowania nie będzie miał możliwości optymalizacji, jeżeli na etapie budowy układu popełniono błędy projektowe. Największe negatywne przełożenie będzie miał dobór nieodpowiednich maszyn, które może cechować brak wystarczającej elastyczności oraz brak odpowiedniej liczby lub wielkości zbiorników buforowych, które są kluczowe dla zapewnienia stabilności ciśnienia. Również odpowiednia klasa pozostałego osprzętu ma wpływ na potencjalne oszczędności, jednak bardziej zaawansowane rozwiązania będą wyraźnie droższe w zakupie od podstawowych rozwiązań, które w wielu przypadkach mogą być wystarczające.

Bardzo istotnym aspektem jest również wykorzystanie sprężarek śrubowych z przetwornicą częstotliwości, które adaptują się do zmiennego zapotrzebowania dobierając aktualne obroty silnika na bazie wewnętrznego regulatora PI lub PID opartego o pomiar ciśnienia wyjściowego. Tutaj jednak należy zapewnić odpowiednią organizację pracy uwzględniającą pozostałe maszyny.

Mając dobrany prawidłowo układ wielosprężarkowy, sprężarki bazowe w nim występujące będą pracować w celu pokrycia podstawowego zapotrzebowania (najlepiej w okolicach punktu pracy, ponieważ tam mają największą wydajność) a sprężarki wyposażone w przetwornicę częstotliwości będą reagować na zmiany zapotrzebowania.

W świetle niniejszej rozprawy wprowadzono rozróżnienie na trzy rodzaje układów, jakie zostały przebadane, a z jakimi można się spotkać w przemyśle dla układów wielosprężarkowych:

- *typ A* układ bez sprężarek z przetwornicą częstotliwości, gdzie obciążenie/odciążenie lub włączenie/wyłączenie następuje na bazie aktualnego poboru,
- *typ B* układ z jedną sprężarką wyposażoną w przetwornicę częstotliwości, która odpowiada za dopasowanie się do zmiennego zapotrzebowania, a pozostałe sprężarki są obciążane lub odciążane sekwencyjnie,
- *typ C* układ z kilkoma sprężarkami, z których każda jest wyposażona w przetwornicę częstotliwości.

Każdy typ układu wymaga innego podejścia do sterowania, a jego adaptacja do zmiennego zapotrzebowania będzie się różnić. Autor w tym miejscu chciał zwrócić uwagę, że aby przetestować metody uczenia zaproponowane w niniejszej rozprawie, należy najpierw doprowadzić układ sterowania do stanu nominalnej pracy. Jeżeli w układzie pojawiają się elementy, które destabilizują działanie układu, czyli stanowią zakłócenie, to należy najpierw w jak największym stopniu je wyeliminować. Zaobserwowano w praktyce, że brak możliwości nauczenia się zachowania układu może wynikać z nieprawidłowej konfiguracji danej maszyny, niewłaściwej częstotliwości zebranych pomiarów lub np. osprzętu tj. zamarzanie osuszacza. Spotkano się również z dławieniem całego układu w wyniku nieprawidłowo ustawionej klapy wyjściowej w instalacji klienckiej, którą możemy interpretować jako ingerencję czynników nieprzewidywalnych.

2.1.4 Problemy eksploatacyjne i kierunki optymalizacji

Mimo tego, że sprężarkownie budowane są często na bazie nowoczesnych technologii, to w praktyce często działają w sposób daleki od optymalnego. Charakter pracy sprężarkowni podczas jej cyklu życia potrafi być zmienny w zależności od charakteru produkcji realizowanego przez danych zakład. W okresach dużego zapotrzebowania na produkty oferowane przez dany zakład sprężarkownia może być wykorzystywana w sposób intensywny, podczas gdy w okresach przestojów lub zmniejszonej produkcji może być wykorzystywana w sposób nieefektywny. Na cykliczność i powtarzalność produkcji mogą również wpływać przepisy środowiskowe, które mogą wymuszać ograniczenia co do poziomu hałasu w ciągu doby lub konieczność wyłączenia w weekendy. Wysokie ceny energii elektrycznej mogą zachęcać zakłady do pracy w określonych godzinach przy najkorzystniejszych taryfach.

W związku z tym, że często sprężarkownia może być projektowana na bazie największego zapotrzebowania, w okresach o mniejszym poborze może być wykorzystywana w sposób nieefektywny. Należy się liczyć w takich warunkach z następującymi ograniczeniami:

- przewymiarowanie układu, zbyt duże maszyny nie będą w stanie pracować w optymalnym punkcie pracy, co przy zmniejszonym zapotrzebowaniu będzie prowadzić do pogorszenia efektywności energetycznej (większa maszyna, będzie zużywać więcej energii elektrycznej na biegu jałowym)
- zbyt wysokie ciśnienie robocze, które może kompensować za niedobory projektowe instalacji dystrybucyjnej,
- nieszczelności w sieci dystrybucyjnej, które mogą wymuszać ciągłą pracę maszyn w okresach przestojów,
- nieoptymalny układ sekwencyjny pracy sprężarek, który może prowadzić do pracy w nieefektywnych punktach pracy,
- niewłaściwe wykorzystanie sprężarek z przetwornicą częstotliwości lub brak możliwości dynamicznego wpływu na ich pracę.

Brak zaadresowania tych problemów będzie miał negatywny wpływ związany zarówno z kosztami traconej energii elektrycznej, jak i gorszymi warunkami eksploatacji maszyn, które mogą pracować w warunkach skracających ich żywotność. Sama reorganizacja pracy sprężarkowni, realizowana przez sterownik nadrzędny, pozwala znacznie ograniczyć wspomniane straty, a w wielu przypadkach uchronić przed ponoszeniem dużych nakładów inwestycyjnych. Ograniczenie poziomu ciśnienia roboczego do minimalnego poziomu jest również prostym sposobem oszczędności.

2.1.5 Tradycyjne metody optymalizacji

W celu optymalizacji istniejących układów można przyjąć różne strategie, w których naukowcy oraz inżynierowie będą mogli wykorzystać swoją dziedzinową wiedzę i umiejętności w obszarach, które na pozór mogą wydawać się od siebie niezależne, jednak będą prowadzić do wymiernych oszczędności. W przypadku doskonałego opomiarowania można np. wprowadzić licznik statystyczny CUSUM, którego poziom będzie informował o typowym zużyciu energii a jego przekroczenie będzie sygnałem do podjęcia działań i jak we wspomnianym artykule [14] pozwala wykryć nieszczelność w samej maszynie produkcyjnej. Można również skupić się na samej optymalizacji pracy sprężarek przez ustalenie odpowiedniej sekwencji [25] i uzyskać akceptowalny współczynnik efektywności energetycznej bez wymiany maszyn. Nadal najprostszym sposobem optymalizacji będzie wprowadzenie sprężarek wyposażonych w przetwornicę częstotliwości, co znacząco ogranicza czas pracy sprężarkowni na biegu jałowym [15] jednak wiąże się z kosztem zakupu maszyny.

W związku ze zwiększeniem zainteresowania możliwościami wykorzystania sztucznej inteligencji w przemyśle, coraz częściej pojawiają się również propozycje rozwiązań

bazujących na uczeniu maszynowym do optymalizacji pracy sprężarkowni. Pojawiają się modele predykcyjne prognozujące zapotrzebowanie na sprężone powietrze, co pozwala na bardziej efektywne planowanie pracy maszyn [23]. Popularne zastosowanie to wykorzystanie uczenia głębokiego w celu przewidywania awarii [32], jednak takie podejście wymaga często dodatkowego, kosztownego opomiarowania układu.

2.2 Uczenie głębokie w sterowaniu przemysłowym

2.2.1 Kontekst historyczny — od klasycznej automatyki do Przemysłu 4.0

Przed pojawieniem się programowalnych sterowników logicznych PLC układy sterowania realizowane były przez wykorzystanie przekaźników elektromagnetycznych oraz regulatorów czasowych realizujących mniej lub bardziej złożone zadania funkcji sterowania. Biorąc pod uwagę, że w przypadku bardziej skomplikowanych procesów technologicznych, jak np. w rozwijającym się w latach 60. XX wieku przemyśle samochodowym liczba przekaźników mogła sięgać nawet kilku tysięcy, to w praktyce utrzymanie takiego układu sterowania było bardzo trudne i wymagało doskonale przeszkolonego personelu biegłego w tworzonych rozwiązaniach. Awaria jednego przekaźnika mogła zachwiać cały proces produkcyjny a identyfikacja wadliwego komponentu była czasochłonna i wymagająca odpowiedniego doświadczenia. Warto przytoczyć anegdotę zasłyszaną od automatyka pracujących w przemyśle na Śląsku w latach 70., że takie układy przekaźnikowe pozwoliły im nabyć odporności psychicznej, gdy podczas zatrzymania zakładu w związku z awarią przekaźnika mieli za plecami cały personel kierowniczy czy nawet dyrekcję. Współcześnie w przypadku sprężarek śrubowych nadal można się spotkać z maszynami z lat 90. XX wieku gdzie cała logika sterowania została zrealizowana bez sterownika programowalnego na bazie ówczesnych rozwiązań. Należy wspomnieć, że nowoczesnych układach rola logiki sterującej zbudowanej na przekaźnikach lub stycznikach może pojawić się w celu zabezpieczenia komponentów maszyny np. popularny układ rozruchowy gwiazda-trójkąt, mający nie dopuścić do jednoczesnego zasilenia wszystkich styczników i uszkodzenia silnika (przy równoczesnym zmniejszeniu prądu rozruchowego), jednak użycie dedykowanego sterownika zabezpieczającego maszynę, jak i pamiętającego o niezbędnych przeglądach, jest wskazane do zapewnienia jej jak najdłuższej żywotności. Wracając do końca lat 60., należy podkreślić, że rosnąca złożoność procesów technologicznych wymagała bardziej elastycznego oraz skalowalnego podejścia do sterowania.

W odpowiedzi na rosnące wymagania przemysłu w zakresie automatyzacji procesów

technologicznych oraz wspomnianych problemów, na początku lat 70. XX wieku pojawiły się pierwsze programowalne sterowniki logiczne PLC, które wykroczyły poza przemysł samochodowy i zostały wykorzystane w takich gałęziach przemysłu jak przemysł spożywczy, napojów, metalowy, wytwórczy oraz celulozowo-papierniczy [33]. Sterowniki PLC pozwoliły na zastąpienie zawodnych rozwiązań bazujących na przekaźnikach uniwersalną, modułową, łatwo modyfikowalną oraz skalowalną strukturą nadającą się do wymagających zastosowań oferując dodatkowe funkcje logiki programowej. Sterowniki PLC niosą ze sobą również pewne ograniczenia, takie jak stosunkowo niewielka ilość pamięci (nawet przy użyciu większych kart pamięci celem rozbudowy, nie zaoferują więcej niż komputer), oraz tego, że nie w każdej dziedzinie ich zastosowanie będzie niosło ze sobą korzyści (jak choćby w przypadkach, gdy dedykowane rozwiązania mikroporocesorowe mogą dostarczyć tańsze rozwiązanie sterowania) [34]. Mając ciąg technologiczny zbudowany z kilkunastu maszyn w przypadku wystąpienia awarii na jednej z nich, operator procesu również musiałby znaleźć się w jej okolicy, aby na bazie ograniczonych komunikatów (świetlna sygnalizacja awarii) dokonać diagnostyki oraz naprawy, podczas której nie odbywa się produkcja, a zamówione towary będą miały opóźnienie z dotarciem do klienta końcowego. Równolegle do rozwoju sterowników programowalnych, rozpoczęły rozwijać się kolejne dwie gałęzie badawcze, które dały podstawy pod dalszą rewolucję przemysłową a korzystanie z jej osiągnięć jest dzisiaj na porządku i znajduje odzwierciedlenie w sektorze sprężonego powietrza. Sama logika sterowania nie rozwiązywała problemu komunikacji na linii operator - maszyna, a potężne układy wyposażone w dziesiątki lampek czy przycisków były trudne w obsłudze.

W 1965 roku, zauważono potrzebę wprowadzenia interfejsu na linii człowiek-maszyna przez użycie dotykowego (wówczas pojemnościowego) ekranu dla kontrolera lotów w przemyśle samolotowym [35],co było naturalnym rozwiązaniem dla podejmowania decyzji, a zaproponowany panel dotykowy dał podwaliny pod nowoczesne panele operatorskie. W 1967 roku pojawiło się już określenie *ang. Touch Displays: A Programmed Man-Machine Interface*, które zwraca uwagę na potrzebę stworzenia przejrzystego interfejsu pomiędzy człowiekiem a maszyną. W przemyśle pierwsze wdrożenie panelu operatorskiego, którego możemy określić już nowoczesnym mianem HMI *ang. Human-Machine Interface*, pojawiło się w Europejskiej Organizacja Badań Jądrowych CERN [36] w celu sterowania akceleratorem cząstek SPS (*ang. Super Proton Synchrotron*). Takie rozwiązanie umożliwia pominięcie wspomnianych wcześniej sygnałów optycznych za pomocą graficznego interfejsu oraz wprowadzenie dotykowego interfejsu z wykorzystaniem ekranu jako przycisku. Tutaj również pojawia się ciekawy aspekt, ponieważ mamy do czynienia z umiejscowieniem panelu w sterowni *ang. Control Room*. Jednak szerokie zastosowanie w przemyśle, praktycznie do teraz, umożliwiło wykorzystanie ekranów rezystancyjnych, umożliwiających pracę z wykorzystaniem rysika przez wprowadzenie sensora współrzędnych [37] oraz powiązanie przezroczystego sensora z

monitorem CRT [38]. Obecnie ekran dotykowy służący do obsługi sprężarki śrubowej jest standardowym wyposażeniem maszyny. Tymczasem pod koniec lat 70. XX wieku przemysł został wyposażony w koncepcyjne założenia umożliwiające budowę skomplikowanych maszyn i układów oraz ich szybką diagnostykę, jednak potrzebny był kolejny aspekt wspomnianej rewolucji. Pojawiła się potrzeba nadzoru nie tylko lokalnego, lecz także oddalonego, co doprowadziło do powstania systemów wizualizacji.

Mając opracowane sterowanie przy użyciu sterowników PLC, diagnostykę i obsługę z wykorzystaniem paneli operatorskich HMI pojawiła się potrzeba zaawansowanego przetwarzania danych pomiarowych w postaci wykraczającej poza sterowanie maszynowe mniej polegające na fizycznej obecności operatora procesu. Tak naturalne obecnie systemy, jak zbieranie danych ze stacji pogody, analiza danych z systemu kolejowego, paliwowego, systemów energetycznych czy wodociągowych wymagała rozwiązań, które pozwalałyby na zbieranie i przetwarzanie danych z lokacji znajdujących się w dużych odległościach [39]. Jeszcze przed pojawieniem się sterowników PLC konstruowano rozwiązania telemetryczne umożliwiające zbieranie danych z oddalonych czujników przy użyciu dostępnych wówczas sygnałów elektrycznych, linii telefonicznych (również prywatnych) czy przy wykorzystaniu radia (do momentu aż w związku z rozwojem sieci radiowych ograniczył się znacząco zakres dostępnych częstotliwości). System użyty w rozprawie do przetwarzania i zbierania danych pomiarowych wykorzystał również koncepcje wtedy opracowane i tam, gdzie to było możliwe, użyto sieci GSM oraz modułów telemetrycznych dla potrzeb zbierania i przetwarzania danych pomiarowych. W związku z dynamicznym rozwojem komputerów możliwe było wprowadzenie scentralizowanego punktu sterowania, gdzie dane zbierane z rozproszonych układów zostały zebrane i przetworzone dostarczając, w czasie rzeczywistym, informacji o stanie monitorowanego systemu. Od końca lat 70. pojawiły się systemy, które możemy określić mianem SCADA (*ang. Supervisory Control and Data Acquistion*). Wspomniane Nadrzędne Systemy Sterowania oraz Zbierania Danych Pomiarowych stały się fundamentem cyfryzacji przemysłu, integracji danych, co stanowi ważny element Przemysłu 4.0.

Systemy SCADA poza wspomnianymi funkcjami nadzorczymi nad instalacjami przemysłowymi łączą w sobie zadania wizualizacji, alarmowania, akwizycji oraz archiwizacji danych procesowych jak i dostarczają możliwości zdalnego sterowania. Systemy SCADA definiuje się jako nadrzędne systemy nadzoru i akwizycji danych, umożliwiające monitorowanie, sterowanie oraz archiwizację parametrów procesowych w czasie rzeczywistym. Dane z rozproszonych modułów takich jak sterowniki PLC czy Zdalne Jednostki Terminalowe (*ang. RTU – Remote Terminal Unit*) realizujące monitoring oraz sterowanie danego procesu, przy pomocy protokółów komunikacyjnych udostępniane są lokalnym stacjom operatorskim HMI jak i serwerom SCADA warstwy nadrzędnej systemu sterowania. Komunikacja pomiędzy warstwami odbywa się przy pomocy protokołów komunikacyjnych jak np. nadal popularny Modbus RTU, DNP3 czy

IEC 60870-5, które zapewniają niezawodną wymianę danych w czasie rzeczywistym, odporność na zakłócenia oraz mechanizmy kontroli integralności danych [40].

Równolegle do rozwoju SCADA, w zakładach przemysłowych pojawiła się potrzeba integracji sterowania ciągłego w ramach jednej platformy, co doprowadziło do powstania systemów DCS (*ang. Distributed Control System*). Ponieważ wraz z rozwojem technologii zaciera się różnica między zaawansowanymi sterownikami PLC, których instalacje mogą zawierać dziesiątki modułów rozproszonych a dedykowanymi danej gałęzi przemysłu rozwiązaniami systemów rozproszonych DCS, w aplikacjach o charakterze procesowym i krytycznym (np. sektor energetyczny) dominującym rozwiązaniem nadal pozostaje klasyczna architektura rozproszona (MacDowall, 2004). Systemy te rozwijały się równolegle do wspomnianych dotychczas systemów PLC, HMI oraz SCADA. Systemy DCS koncentrują się na sterowaniu procesami ciągłymi i oferują spójną, hierarchiczną architekturę systemu sterowania budowaną w obrębie danego zakładu z wykorzystaniem redundancji czy określonego typu sieci komunikacyjnych. Czołowi producenci jak Siemens czy ABB oferują dedykowane rozwiązania, które pozwalają na zapewnienie zintegrowanej, scentralizowanej bazy procesowej całego zakładu przy zapewnieniu natywnej redundancji na każdym poziomie sterowania ciągłego (co w systemach bazujących na PLC wymaga dodatkowej konfiguracji lub jest niemożliwe lub nieopłacalne). Należy zwrócić uwagę, że SCADA jest systemem nadrzędnym integrującym pracę wielu lokalnych sterowników, natomiast DCS jest systemem sterowania, w którym sterowanie i nadzór są integralną częścią jednej architektury.

W kontekście niniejszej rozprawy SCADA stanowi warstwę nadrzędną odpowiedzialną za zbieranie danych pomiarowych z zespołu sprężarek oraz ich udostępnianie do dalszej analizy predykcyjnej, podczas gdy sterowniki PLC realizują lokalne sterowanie zespołem maszyn wchodzących w skład sprężarkowni. Istotny jest wybór sterownika PLC ze względu na architekturę, która umożliwia zaprojektowanie i wdrożenie inteligentnego wsparcia decyzji procesowych. Taki podział ról jest zgodny z architekturą współczesnych systemów automatyki i stanowi fundament dla rozwiązań Przemysłu 4.0, w których kluczową rolę odgrywa integracja danych, ich przetwarzanie oraz wykorzystanie w inteligentnych algorytmach sterowania. Rozwój SCADA i DCS zapoczątkował konwergencję warstwy operacyjnej (OT) z warstwą informatyczną (IT), co stanowi fundament współczesnych architektur Przemysłu 4.0.

Biorąc pod uwagę architekturę systemów produkcyjnych (jak w przypadku systemów sprężonego powietrza) wypada wspomnieć o modelu architektury referencyjnej opracowanej przez Uniwersytet Purdue [41], który podkreśla korzyści wynikające ze strukturyzacji systemów produkcyjnych. Proponowane kaskadowe podejście wymiany danych umożliwia śledzenie przepływu informacji i produkcji na wszystkich szczeblach

przedsiębiorstwa. Klasyczny model zakładał separację warstwy technologii operacyjnych OT (*ang. Operational Technology*), obejmującej urządzenia sterujące procesami fizycznymi, oraz warstwy technologii informatycznych IT (*ang. Information Technology*), odpowiedzialnej za przetwarzanie, analizę i dystrybucję danych w systemach biznesowych. Potężne ilości zbieranych danych, nie powinny pozostawać tylko na poziomie sterowania a zostać zintegrowane z systemami analitycznymi, bazami danych, serwerami aplikacyjnymi oraz usługami sieciowymi. Zacieranie granic pomiędzy warstwami IT i OT jest niezbędne w kontekście rosnącej roli systemów predykcyjnych, które wymagają dostępu do danych procesowych w czasie rzeczywistym [8]. Wspomniana konwergencja IT/OT stworzyła fundament pod rozwój Internetu Rzeczy (IoT), w którym moduły RTU, sterowniki PLC, sprężarki, liczniki energii itp. stały się elementami sieci komunikującej się nie tylko w obrębie zakładu, lecz także z systemami nadrzędnymi, usługami chmurowymi oraz platformami analitycznymi (w zależności o polityki cyberbezpieczeństwa danego zakładu). IoT rozszerza klasyczne podejście SCADA o możliwość gromadzenia danych z tysięcy urządzeń, ich przetwarzanie brzegowe (*ang. edge computing*), a następnie wykorzystanie w algorytmach predykcyjnych i systemach wspomagania decyzji. W przeciwieństwie do tradycyjnych architektur, w których dane przepływały kaskadowo przez kolejne warstwy modelu Purdue, IoT umożliwia ich równoległy przepływ do wielu odbiorców, co mogłoby zwiększyć elastyczność systemów sterowania. W kontekście systemów sprężonego powietrza koncepcja IoT mogłaby dostarczyć wielu dodatkowych możliwości analitycznych oraz rozszerzyć zakres dostępnych danych, jednak urządzenia badane w niniejszej rozprawie nie zostały przystosowane do integracji w ramach architektury IoT, co naturalnie ogranicza zakres możliwości.

Rozwój sterowników PLC, systemów SCADA, jak i rozproszonych architektur sterowania DCS stworzył fundament pod trzecią rewolucję przemysłową określaną mianem Przemysłu 3.0. Odejście od zawodnego, przekaźnikowego sterowania i przejście do cyfrowej automatyzacji procesów technologicznych dostarczyło potężne możliwości monitorowania maszyn, archiwizacji parametrów pracy oraz danych procesowych, jak i otworzyło szerokie możliwości analityczne (a co za tym idzie m.in. predykcyjne). Konstrukcja robotów przemysłowych, maszyn CNC, a współcześnie nawet robotów humanoidalnych [42] umożliwiło znaczący wzrost wydajności i powtarzalności produkcji.

Skala automatyzacji wymusiła integrację danych z wielu warstw modelu hierarchicznego, co doprowadziło do konwergencji technologii operacyjnych (OT) i informatycznych (IT). Dane produkcyjne zaczęto traktować jako zasób wspierający procesy decyzyjne i optymalizację pracy maszyn. Warto wspomnieć o koncepcji, pojawiającej się w kontekście Przemysłu 3.0, inteligentnych sieci energetycznych określonych jako „Internet energetyczny", w którym rozproszone źródła energii komunikują się i wymieniają nadwyżkami pomiędzy sobą [43]. Koncepcja ta wykracza poza automatykę przemysłową, jednak wskazuje na rosnącą rolę cyfrowej komunikacji i integracji danych

w skali całych gospodarek, która stała się fundamentem kolejnej rewolucji. Idea ta wpisuje się w trend efektywności energetycznej związany również z układami sprężonego powietrza. Sprężarki śrubowe mogą być źródłem ciepła odpadowego pochodzącego z obiegu olejowego oraz stanowić element lokalnych systemów odzysku energii, będący elementem rozproszonej gospodarki energetycznej.

Przemysł 4.0 stanowi kolejny etap ewolucji systemów produkcyjnych. Klaus Schwab w 2016 roku zauważył, że ludzkość stoi u progu rewolucji, która zmienia sposób życia, pracy oraz wzajemnej komunikacji a jej zakres jest bez precedensu w historii ludzkości [7]. Dla automatyki przemysłowej kluczową rolę odgrywa w niej głębsza integracja technologii operacyjnych (OT) z informatycznymi (IT) oraz wykorzystanie zaawansowanych metod analitycznych z tym związanych. W stosunku do Przemysłu 3.0, którego podstawą była automatyzacja i cyfrowe sterowanie, Przemysł 4.0 koncentruje się na tworzeniu systemów cyber-fizycznych CPS (ang. *Cyber-Physical Systems*), posiadających zdolność do autonomicznej wymiany danych i podejmowania decyzji w czasie rzeczywistym [44]. Współczesne zakłady produkcyjne generują ogromne ilości danych procesowych, które dzięki przenikaniu się warstw IT/OT mogą być przetwarzane przez algorytmy sztucznej inteligencji bezpośrednio na poziomie produkcji.

Przemysł 4.0 obejmuje szereg technologii: Internet Rzeczy (IoT), przetwarzanie brzegowe, systemy chmurowe, komunikację sieciową w czasie rzeczywistym oraz zaawansowane metody analityczne takie jak uczenie maszynowe i uczenie głębokie. W opozycji do modelu Purdue Przemysł 4.0 umożliwia równoległy przepływ informacji pomiędzy warstwami sterowania, nadzoru oraz analityki, pozwalając na wdrażanie inteligentnych algorytmów sterowania [45].

W kontekście niniejszej rozprawy szczególne znaczenie ma predykcja stanów pracy sprężarek śrubowych na podstawie danych procesowych. Uczenie głębokie, jako jeden z kluczowych elementów Przemysłu 4.0, pozwala na uchwycenie złożonych wzorców wewnątrz zależności czasowych danych pomiarowych, co jest podstawą projektowanego systemu sterowania. Zaproponowany system nie wymaga rozbudowywania sieci pomiarowej czy modelowania elementów fizycznego układu.

2.2.2 Sieci rekurencyjne (LSTM, BiLSTM, GRU)

Głównym zadaniem modelowania procesu sprężania powietrza było znalezienie mechanizmów, które będą dobrze odzwierciedlać specyfikę i dynamikę procesu. Obserwując proces w 20-sekundowych odstępach (z racji wykorzystania sieci komórkowej do zbierana danych oraz ograniczeń użytych rozwiązań monitoringu) otrzymano informację o aktualnym przebiegu pomiaru ciśnienia oraz aktualnie dostarczających powietrze maszynach. W związku z tym dane w chwili t będą skorelowane z danymi z poprzednich próbek. Charakter tej zależności być może okaże się cykliczny, może będzie silnie

związany z wartościami poprzedzającymi, a może będzie sięgał dziesiątki tysięcy próbek wstecz. Odpowiedzią na taki charakter procesu, są rekurencyjne sieci neuronowe RNN(ang. *Recurrent Neural Networks*) [46], które są wyposażone w rekurencyjne zależności. Tradycyjna sieć z propagacją do przodu (ang. *feedforward*) traktuje każdą próbkę niezależnie, przez co traci istotny w tym wypadku, kontekst czasowy, przez co nie nadaje się wypadku szeregów czasowych, które mają złożoną dynamikę [47].

Jednym z głównych problemów z jakim można się spotkać przy sieciach rekurencyjnych jest zjawisko zanikającego gradientu (ang. *vanishing gradient*), który uniemożliwia uczenie modelu na danych, które pojawiają się w długim horyzoncie czasowym. Problem ten został zauważony w 1991 roku w pracy dyplomowej Seppa Hochreitera i opisany w publikacji z 1998 roku [48] podczas gdy popularyzacja i niezależna analiza nastąpiła w 1994 roku [49]. W algorytmie sieci rekurencyjnej propagacji wstecznej przez czas BPTT (ang. *Backpropagation Through Time*) gradient błędu jest mnożony przez macierze wag w każdym kroku czasowym. Gdy wartości własne tych macierzy są mniejsze od jedności, gradient wykładniczo zanika wraz z odległością czasową, co uniemożliwia sieciom uczenie się zależności między odległymi w czasie zdarzeniami. Zatem nasuwa się wniosek, że klasyczna sieć RNN jest w stanie uwzględniać kontekst jedynie kilku ostatnich kroków czasowych, co czyni ją nieadekwatną dla procesów przemysłowych gdzie istotne zdarzenia, takie jak zmiana profilu produkcji, może poprzedzać ich skutki o wiele godzin. Gdy wartości własne macierzy będą większe od jedności, można się spotkać z problemem eksplodującego gradientu, który również destabilizuje proces uczenia.

W odpowiedzi na opisane problemy pojawiła się architektura sieci wyposażonej w komórki pamięci długo–krótkotrwałej LSTM (ang. *Long Short-Term Memory*) wprowadzona przez Hochreitera [47]. Różnica w stosunku do klasycznej sieci w jej wypadku to wprowadzenie dwóch równoległych tras informacji tj. stanu ukrytego oraz stanu komórki. Stan komórki (ang. *cell state*) przepływa przez kolejne kroki czasowe z minimalnymi zmianami wynikającymi z bramek i niesie w sobie pamięć o ciągu czasowym. Bramka zapominająca (ang. *forget gate*) odpowiedzialna jest za decyzje o tym, które informacje z poprzedniego stanu komórki należy odrzucić. Bramka wejściowa (ang. *input gate*) decyduje o tym jakie nowe informacje zostaną zapisane. Bramka wyjściowa (ang. *output gate*) odpowiada za to, które elementy stanu komórki przejdą do stanu ukrytego. Takie rozwiązanie pozwala na „zamrożenie" ważnej informacji na wiele kroków czasowych. Sieć może zapamiętać, że sprężarka przeszła w stan obciążenia wiele godzin temu i uwzględnić ten stan przy predykcji nadchodzących stanów, nawet gdy w międzyczasie wystąpiły inne zdarzenia. Podsumowując, zaproponowany w sieci LSTM mechanizm bramek wejściowej, zapominającej i wyjściowej pozwala sieci selektywnie zapamiętywać i zapominać informacje, co jest niezbędne przy modelowaniu procesów przemysłowych z różną rozpiętością zależności czasowych.

Mechanizm LSTM został uproszczony w publikacji z 2014 roku [50], podczas gdy nazwa nowej architektury i jej szczegółowa analiza pojawiły się w późniejszej publikacji [51], gdzie nową architekturę, określono mianem bramkowej jednostki rekurencyjnej GRU (*Gated Recurrent Unit*). Zaproponowane rozwiązanie cechuje połączenie bramki zapominającej i wejściowej w jedną bramkę aktualizacji oraz eliminację członu stanu komórki, dzięki czemu zmniejszono liczbę parametrów sieci o jedną czwartą przy utrzymaniu zbliżonej jakości modelowania zależności długookresowych. Należy zwrócić uwagę, że takie uproszczenie sieci jest szczególnie istotne, gdy proces wymaga ciągłego douczania, co będzie się przekładało na mniejszą ilość obliczeń, jakie należy cyklicznie wykonywać. Odpowiedź na to, który wariant sieci będzie lepszy jest niejednoznaczna, ponieważ wyniki będą silnie zależały od charakterystyki modelowanego procesu [52]. W przypadku prowadzonych badań, empirycznie zaobserwowano, że przy stosunkowo krótkich oknach obserwacji (kilka próbek) różnice między LSTM a GRU są marginalne, natomiast przy wykorzystaniu w uczeniu danych pochodzących wielogodzinnych trendów produkcyjnych dodatkowe człony komórki LSTM dawały lepsze wyniki. Dla rozwiązania produkcyjnego wybór konkretnej architektury powinien zostać zweryfikowany eksperymentalnie dla ustalonego zbioru danych określonego obiektu przemysłowego.

Kolejny istotny aspekt, to kierunek przetwarzania sekwencji, wspomniane architektury LSTM oraz GRU przetwarzają sekwencję wyłącznie w jednym kierunku, który jest zgodny z intuicyjnym fizycznym przebiegiem procesu, czyli od przeszłości do chwili bieżącej (od lewej do prawej). Jednak trzeba się również liczyć z tym, że pojawiają się pewne zależności, które przy rozpatrywaniu sekwencji w kierunku przeciwnym mogą być łatwiej zauważalne a przy jednokierunkowej zostaną utracone (w publikacji własnej zauważono przypadek, gdy pierwsza przewidywana próbka przez LSTM jest predykowana mniej adekwatnie niż kolejne [53]). Na danych zebranych w równych odstępach czasu, rozważając chwilę w danym punkcie, dostępne są dane zarówno z przeszłości, jak i przyszłości co pozwala na dobranie okna obserwacji niosącego w sobie te informacje. W 1997 roku pojawiła się architektura, która nadaje się do wykorzystania tych informacji i naturalnie odzwierciedla podane zależności, jest to architektura dwukierunkowej sieci LSTM czyli BiLSTM [54]. Sieć ta z założenia przetwarza sekwencję z obu stron tzn. od lewej do prawej jak i od prawej do lewej a wyniki obu przebiegów są łączone, co daje lepszą reprezentację każdego punktu w sekwencji. Odnosząc się do systemu sterowania sprężonym powietrzem, tranzycja obciążenia lub odciążenia będzie poprzedzona określonym kształtem przebiegu ciśnienia przed i po wystąpieniu zdarzenia. BiLSTM w tym wypadku może efektywniej zidentyfikować wzorzec tranzycji niż sieć LSTM lub GRU biorąca pod uwagę tylko jeden kierunek przetwarzania.

Warto przytoczyć wzór, który doskonale pokazuje charakter architektury dwukierunkowej BiLSTM [54]:

$$\overrightarrow{h}_t = \text{LSTM}_{\text{forward}}(x_t, \overrightarrow{h}_{t-1}) \tag{2.1}$$

$$\overleftarrow{h}_t = \text{LSTM}_{\text{backward}}(x_t, \overleftarrow{h}_{t+1}) \tag{2.2}$$

$$h_t = [\overrightarrow{h}_t; \overleftarrow{h}_t] \tag{2.3}$$

gdzie $\overrightarrow{h}_t$ i $\overleftarrow{h}_t$ oznaczają odpowiednio ukryte stany sieci przetwarzającej sekwencję do przodu (z przeszłości w przyszłość) i do tyłu (z przyszłości w przeszłość). Wspomnianą teorię o tym, że sieć biorąca pod uwagę sekwencję w obu kierunkach może efektywniej rozpoznawać wzorzec obciążenia w oknie obserwacji, potwierdzono empirycznie w badaniach własnych, gdzie BiLSTM wykazało przewagę nad jednokierunkowym LSTM przy predykcji binarnych stanów obciążenia. Spostrzeżenie co do rozpatrywania układów sprężonego powietrza przy ewaluacji danych w kierunku przeciwnym zaproponowano w 2025 roku [53] nie używając wtedy jeszcze architektury dwukierunkowej.

Po dobraniu architektury sieci oraz wykonaniu pierwszych treningów na zebranych danych, następuje zwykle zetknięcie ze zjawiskiem dryftu koncepcyjnego (ang. *concept drift*) [55]), który jest wyraźnie obecny przy procesie modelowania sprężonego powietrza. Ukryta charakterystyka wiążąca dane wejściowe oraz przewidywanie zdarzeń w przyszłości będzie zmienna w czasie. W praktyce oznacza to, że dla dwóch takich samych zestawów danych wejściowych, następujące zdarzenia w przyszłości mogą być różne. Dane treningowe, zebrane dla potrzeby uczenia głębokiego w ramach niniejszej rozprawy, pochodzące z procesu sterowania sprężarkami są związane z kilkoma ograniczeniami. Przede wszystkim istotne jest, aby dane zostały zebrane podczas prawidłowej pracy układu. Wystąpienie okresu awarii, planowanych przestojów, krótkofalowych zmian konfiguracji systemu sterowania (np. w związku z ustawieniami wentylacji, która może różnić się w zależności od pory roku czy wprowadzania rozwiązań odzysku ciepła odpadowego) będzie prowadziło do zebrania nieprawidłowych wzorców, które nie będą reprezentatywne dla normalnych wzorców pracy, przez co będą zaburzać proces uczenia. Następnie aspekt wykrywania zdarzeń obciążenia, czyli tych, które sieć ma wykrywać, będzie prowadził do braku równowagi klas tj. w stosunku do całej liczby próbek, zdarzenia mogą być niezwykle rzadkie, również może zabraknąć danych opisujących zapisanie konkretnej maszyny, ponieważ takie zdarzenie nie wystąpiło podczas całego procesu uczenia (w związku z tym sieć nie będzie wiedziała, jaka jest waga określonej odstawionej maszyny i w wypadku jej użycia, będzie zawsze przewidywała nieprawidłowy obecnie stan). Ostatnim aspektem będzie rozpiętość danych w czasie, które będą dobrze odzwierciedlać dynamikę obiektu. Może się okazać, że potrzebne będą tygodnie lub miesiące (w zależności od częstotliwości zmian układu oraz okresu próbkowania), co wpływa na opóźnienie wdrożenia systemu. Zaobserwowano w przypadku badań, że

dla obiektów o średniej dynamice, optymalny okres wynosi około trzech tygodni, zbiór zbyt krótki nie zawiera dostatecznie zróżnicowanych wzorców, natomiast zbiór zbyt długi może zawierać wzorce nieaktualne ze względu na dryft koncepcyjny [53].

Podsumowując zjawisku dryftu koncepcyjnego w przypadku sprężonego powietrza [56], zauważono, że statystyczne własności danych w czasie mogą wynikać z wielu czynnników takich jak sezonowa zmienność produkcji, zmiana profilu produkcji (np. rodzaju i ilości produkowanych elementów), modyfikacje konfiguracji sprężarek (np. optymalizacja nastaw regulatorów PI sprężarek z przetwornicą częstotliwości), zużycia samych sprężarek (gorsza wydajność, niewykonanie przeglądów), innych czynników (wprowadzenie układów odzysku ciepła), które mogą wpływać na eksploatację całego układu. Zauważalne będzie wówczas pogorszenie adekwatności predykcji mimo tego, że teoretycznie fizyczna konfiguracja układa pozostała niezmienna. W tym momencie gdyby bazować tylko na predykcji sieci neuronowej, pojawiające się błędy mogłyby stanownić niebezpieczeństwo zachwiania dostaw sprężonego powietrza. Poza koniecznością douczania (ang. *retreningu*) modelu zalecane jest wprowadzenie dodatkowych ograniczeń, które będą stanowiły „zawór bezpieczeństwa" pracy dla całego układu. Szczegółowo rodzaje dryftu koncepcyjnego omówiono w sekcji 2.7.

Na wczesnym etapie badań, pojawiło się pytanie, czy sieć, która byłaby nauczona na jednym rodzaju obiektu, byłaby również efektywna w przypadku innego, co sprawdzono przez bezpośrednie zastosowanie modelu do predykcji ciśnienia na innych obiektach, bez jego douczania [18]. Możliwość transferu wiedzy pomiędzy różnymi instalacjami przemysłowymi, o zbliżonej charakterystyce jest znana w literaturze i określona jest pojęciem uczenia transferowego (ang. *transfer learning*) [57]. W przypadku gdy uczenie wymaga dużych mocy obliczeniowych, posiadanie punktu startowego skracałoby czas i koszt wdrożenia. Dla rozpatrywanych układów sprężonego powietrza, transfer wiedzy jest jednak utrudniony ze względu na indywidualny charakter specyfiki każdego obiektu — różnice będą dotyczyły innej konfiguracji instalacji odbiorczej, odmiennego układu sprężarek, charakterystyce odbiorców oraz profilu produkcji, przez co wzorce opracowane na jednym obiekcie, nie będą mogły zostać zgeneralizowane i przeniesione na inny obiekt. Zaobserwowano empirycznie, że próby bezpośredniego zastosowania modelu referencyjnego na nowym obiekcie, bez douczania skutkują dużo gorszymi wynikami niż wytrenowanie modelu od postaw. Najlepiej w związku z tym byłoby wykorzystać techniki „uczenia się, jak się uczyć" (ang. *meta-learning*) [58] i/lub dostrajania (ang. *fine-tuning*) [59], które powinny zapewnić szybszą adaptację modelu do nowego obiektu przy zachowaniu istniejącej architektury.

Przy proponowanej metodyce (która została opisana w rozdziale 3) pojawia się jeszcze jeden aspekt, związany z interpretacją działania modelu, który może stanowić niedogodność w procesie wdrożenia. Przed wdrożeniem przeprowadzono analizę przewidywanych wyników wspólnie z potencjalnym operatorem systemu, przechodząc

przez wykresy okien kolejnych zdarzeń decyzyjnych. Z perspektywy człowieka, będąc przyzwyczajonym do deterministycznych algorytmów sterowania stosowanych w sterownikach PLC, można odczuwać trudność w zrozumieniu decyzji podejmowanych przez model głębokiego uczenia, który z perspektywy użytkownika funkcjonuje jako „czarna skrzynka". Interpretowalność modeli uczenia maszynowego (ang. Explainable AI, XAI) jest ciągle rozwijanym obszarem badań [60], a metody takie jak SHAP [61] czy LIME [62] umożliwiają wyjaśnianie predykcji poprzez przypisywanie wag cechom wejściowym. Jednak ich praktyczne zastosowanie w sieciach rekurencyjnych działających w czasie rzeczywistym jest obciążające obliczeniowo, przez co ich użyteczność w systemach sterowania jest ograniczona. Z tego względu decyzja o wyłączeniu sprężarki jest podejmowana autonomicznie przez model, lecz dla sterownika PLC stanowi jedynie rekomendację, która może być obwarowana dodatkowymi warunkami bezpieczeństwa. Takie rozwiązanie zmniejsza opór wdrożeniowy, ponieważ sterownik PLC zachowuje pełną odpowiedzialność za bezpieczeństwo procesu chroniąc przez potencjalnym przestojem.

2.2.3 Zastosowania predykcji w energetyce

W sektorze energetycznym popularnym zastosowaniem jest predykcja zapotrzebowania na energię elektryczną (ang. *load forecasting*) i prawdopodobnie jest to jedno z najlepiej zbadanych zastosowań sieci rekurencyjnych w energetyce. Przewidywanie obciążenia sieci w różnych horyzontach czasowych jest niezbędne dla efektywnego zarządzania siecią, planowania rezerw mocy oraz optymalizacji kosztów zakupu energii na rynkach [63]. Sieci LSTM [24] w tym przypadku są wyraźnie lepsze niż modele klasyczne (jak np. ARIMA — *AutoRegressive Integrated Moving Average*) czy regresja liniowa. Modelowanie nieliniowych wzorców zużycia energii wynika ze zmienności zachowań konsumentów, zmienności produkcji czy różnych warunków pogodowych. Krótkoterminowe przewidywania (zakres minut lub godzin) doskonale sprawdziły się przy analizie australijskich sieci elektroenergetycznych, gdzie zauważono potrzebę odchodzenia od klasycznego zarządzania sieciami dystrybucyjnymi w kierunku sieci inteligentnych (ang. *Smart Grid*), gdzie architektura LSTM efektywnie modeluje nieliniowe, sekwencyjne zależności w historycznych szeregach czasowych oddając rozdział mocy. Taki układ sieci jest w stanie adekwatnie aproksymować nagłe piki i spadki obciążeń, co stanowi krytyczny element stabilizacji nowoczesnych systemów przemysłowych opartych na zmiennych parametrach zasilania, jednak zauważono również, że w celu bardziej adekwatnej predykcji należy dokonać profilowania odbiorców [24].

Predykcja krótkoterminowa sprawdza się przy predykcji bieżącego zapotrzebowania na energię, natomiast planowanie rozbudowy infrastruktury elektroenergetycznej czy

zarządzanie rynkiem energii wymaga modelowania w horyzontach średnio– i długoterminowych. Biorąc pod uwagę wspomniane wcześniej ograniczenia i np. sezonowość, literatura wskazuje znaczną przewagę sieci dwukierunkowych nad jednokierunkowymi (zauważono to również w badaniach własnych wokół sprężonego powietrza). BiLSTM lub rozwiązania hybrydowe lepiej sprawdzą się przy inteligentnych sieciach energetycznych, znajdując ukryty kontekst, przy rozpatrywaniu przeszłych i przyszłych stanów [64]. Drugim rozwiązaniem w przypadku predykcji długoterminowych, jest integracja sieci z mechanizmem uwagi (ang. *attention mechanism*) czy strukturami konwolucyjnymi, co umożliwa analizę nieliniowych zależności trendów obciążeń węzłów energetycznych pomimo wysokiego poziomu zakłóceń danych wejściowych [65].

W odniesieniu do układów sprężonego powietrza, już wcześniej zauważono użyteczność wykorzystania wskaźników statystycznych bazujących na licznikach energii elektrycznej [14], gdzie samo znalezienie odchyłek od stanu normalnego, pozwala na wykrycie anomalii w działaniu maszyny. W przypadku sprężonego powietrza, licznik sprężonego powietrza, dostarczający informacji o ilości zużytego medium, wymaga ciągłego dotrenowywania by nadążyć za zmieniającymi się trendami zapotrzebowania [23]. Taka informacja pozwala na identyfikację wzorców zużycia w różnych horyzontach czasowych co pozwala na optymalizację kosztów energii. W przypadku niniejszej rozprawy tylko strona odbiorcza dysponowała pełnymi danymi odnośnie zużycia energii i wyprodukowanego powietrza, jednak ścisła współpraca z klientem w tym zakresie, przyniosła wymierne korzyści energetyczne. Samo wprowadzenie harmonogramu produkcji i np. zmiana priorytetów maszyn w momencie mniejszego zapotrzebowania na sprężone powietrze (np. w okresie przerw) pozwala na wymierne oszczędności. Opomiarowanie maszyn pod względem ilości produkowanego powietrza jak i energii elektrycznej w połączeniu z obniżeniem ciśnienia do najmniejszego poziomu pozwalającego na regularną produkcję daje istotne korzyści. Jednak predykcja w dłuższych horyzontach czasowych, bardzo cenna z perspektywy planowania oraz doboru maszyn do konfiguracji sprężarkowni, nie dostarcza informacji, niezbędnych do podejmowania decyzji w czasie rzeczywistym. Na potrzeby systemu sterowania potrzebne są decyzje w horyzoncie sekund, być może maksymalnie kilku minut a decyzje np. godzinne nie są użyteczne (np. nagłe załączenie nowej maszyny produkcyjnej po stronie odbiorcy, może błyskawicznie spowodować pojawienie się dryftu koncepcyjnego). Zjawisko to jest szczególnie widoczne na wykresach RMSE, gdzie odchyłka pomiędzy statystyczną predykowaną wartością licznika wyprodukowanego powietrza wydaje się śledzić trend, ale w praktyce dla systemu sterowania, rekomendacja na jej bazie mogłaby oznaczać, nie włączaj maszyny przy spadku ciśnienia i doprowadzić do zatrzymania produkcji.

Kolejnym zastosowaniem, które wydaje się być dominujące, w przypadku maszyn

oraz sieci rekurencyjnych jest wykorzystanie możliwości predykcyjnego utrzymania ruchu PdM (ang. *Predictive Maintenance*), które pozwala na przewidywanie awarii maszyn na podstawie danych z czujników przed ich faktycznym wystąpieniem [52]. Minimum pomiarów, w jakie powinna być wyposażona sprężarka śrubowa w celu zabezpieczenia jej funkcjonowania, to dwa pomiary analogowe odpowiedzialne za pomiar ciśnienia wyjściowego oraz pomiar temperatury oleju na stopniu śrubowym. W ich wypadku, gdy nastąpi przekroczenie określonego poziomu progu alarmowego, następuje ostrzeżenie, a powyżej kolejnego poziomu awaryjne wyłączenie. Niektórzy producenci korzystają również ze śledzenia gradientu charakterystyki temperatury oleju, ponieważ zbyt szybkie nagrzewanie będzie świadczyć o nieprawidłowym działaniu technologicznego obiegu olejowego. W przypadku literatury jednak, badacze zalecają wprowadzenie dodatkowego (często kosztownego) opomiarowania (np. czujniki dgrań), na bazie którego sieci takie jak LSTM czy BiLSTM mogą być użyte do poszukiwania anomalii w przebiegach drgań, ciśnienia oleju, temperatury oleju czy prądu silnika. Takie działanie pozwala teoretycznie na działania prewencyjne i wykonanie niezbędnych przeglądów przed awarią komponentów, w praktyce z własnego doświadczenia w zakresie pomiarów wibracji autor uważa, że działania prewencyjne są bardzo trudne i wymagają szerokiego, kosztownego badania, które może przewyższać koszty pojedynczej maszyny [66]. Mimo tych trudności, większość zastosowań uczenia głębokiego w przemyśle, koncentruje się właśnie na PdM, z założeniem poniesienia dodatkowych kosztów infrastrukturalnych [67]. W przypadku niniejszej rozprawy, tak aby zbadać możliwości wdrożenia na jak największej ilości obiektów, ograniczono się wyłącznie do pomiaru ciśnienia wyjściowego w punkcie poboru oraz stanach binarnych sprężarek. Parametry są powszechnie dostępne w każdym układzie sprężonego powietrza ze sterownikiem nadrzędnym, bez konieczności instalacji dodatkowych czujników. Wypada natomiast wspomnieć, że np. prowadząc analizę obiektu, który, nie miałby możliwości odczytu danych przez protokoły komunikacyjne, instalacja pomiaru ciśnienia, oraz pojedynczych przekładników prądowych dla każdej z maszyny systemu również byłoby adekwatne dla stworzenia analogicznego modelu układu.

Ponieważ producenci, dostarczają wraz z maszynami własne sterowniki nadrzędne, ingerencja w ich algorytmy sterowania może być trudna lub niemożliwa. Autor zalecałby wykorzystanie do potrzeb sterowania nadrzędnego własnego sterownika PLC, który gwarantuje błyskawiczną integrację w zakresie zarówno sterowania maszynami, jak i wyposażenia obiektu w funkcje charakterystyczne dla Przemysłu 4.0. W związku z tym w literaturze, pojawia się luka jak chodzi o obszar predykcji binarnych stanów operacyjnych maszyn przemysłowych w krótkim horyzoncie czasowym oraz użycie ich jako podstawy decyzji sterujących. Przytoczone dotychczas prace skupiają się na predykcji wartości ciągłych (zużycie energii, wyprodukowane powietrze, charakterystyka przepływu) lub na detekcji anomalii wskazujących na zbliżającą się awarię, nie zaś

na przewidywaniu stanów operacyjnych (obciążenie/bieg jałowy) w horyzoncie gwarantującym bezpieczeństwo podjęcia decyzji o wyłączeniu maszyny. Kolejną nowością jak chodzi proponowane podejście jest jawne uwzględnienie asymetrii kosztów błędów predykcji w funkcji straty (większość prac traktuje problem jako symetryczny błąd klasyfikacji). Należy podkreślić, że w rzeczywistości koszt fałszywie negatywnej predykcji (wyłączenie sprężarki, która była potrzebna może prowadzić do spadków ciśnienia) jest fundamentalnie inny od kosztu fałszywie pozytywnej (utrzymanie włączonej, a w rzeczywistości niepotrzebnie pracującej sprężarki) [68], [69].

Układy bazujące na uczeniu głębokim i opisanych wcześniej architekturach w zastosowaniach energetycznych nie pozwalają na jednoznaczne wskazanie, która architektura sprawdza się najlepiej. Wyniki będą przede wszystkim zależeć od dostępnych danych, dynamiki obiektu, horyzontu predykcji oraz przyjętej metryki oceny [70]. Dla zadania predykcji obciążenia sieci elektroenergetycznej zauważono, że BiLSTM wykazuje przewagę przy krótszych horyzontach, gdy dane dla obu kierunków są osiągalne. Dla potrzeb sterowania sprężarkami krótki horyzont predykcji (od 2 do 5 min) znajduje się w obszarze, gdzie BiLSTM wykazuje empiryczne zalety potwierdzone w badaniach własnych [71].

2.3 Uczenie wrażliwe na koszt i asymetryczne koszty

W poprzedniej sekcji zwrócono uwagę na asymetryczność kosztów błędów klasyfikacji, przy czym popularne podejście w uczeniu maszynowym zakłada ich symetryczność (tzn. fałszywie pozytywna i fałszywie negatywna klasyfikacja jest traktowana równorzędnie). W przypadku przemysłu, koszt nieprzewidzenia awarii maszyny będzie dużo większy niż koszt jej prewencyjnego zatrzymania. Koszt utrzymania sprężarki na biegu jałowym niepotrzebnie przez kilka minut, będzie mniejszy niż wyłączenie sprężarki, które doprowadzi do spadku ciśnienia w związku z nieprawidłową decyzją wyłączenia. Opisanymi przypadkami zajmuje się uczenie wrażliwe na koszt (ang. *cost-sensitive learning*), które stanowi rodzinę metod jawnie uwzględniających asymetrię kosztów błędów w procesie uczenia modelu [72].

2.4 Funkcje straty w zadaniach z asymetrycznym kosztem

2.4.1 Funkcje straty w uczeniu maszynowym

Funkcja straty (ang. *loss function*) w uczeniu maszynowym jest funkcją, której minimum model będzie poszukiwał w trakcie procesu uczenia. Funkcja może nieść w sobie informacje o karze w przypadku błędnej klasyfikacji lub wielkości odchyłki

predykcji stanu analogowej wartości procesowej od predykowanej. Na jej bazie, model otrzymuje wytyczne co do dynamiki czy nawet charakteru obiektu, które dla człowieka, mogą być trudne w zaobserwowaniu lub pozostawać niedostrzegalne.

W sektorze sprężonego powietrza, w związku z bazowaniem procesów uczenia na wartościach analogowych takich jak ciśnienie względne lub przepływ powietrza [23] powszechną praktyką jest budowa funkcji straty wokół błędu średniokwadratowego MSE (ang. *Mean Squared Error*) (2.4). Taka ocena wymaga dodatkowej weryfikacji ze względu na to, że pozornie niewielka odchyłka dla całego zbioru może odzwierciedlać sytuacje, gdy model nie reaguje prawidłowo na dane wejściowe a np. znajduje odpowiednią średnią statystyczną. Empirycznie zauważono również, że gdy następuje śledzenie trendów charakterystyki to oczywistym jest, że model będzie lepiej reagował i wykrywał zmiany, jednak w praktyce wskazane jest sprawdzenie danych predykcji przez operatora przed ostatecznym wdrożeniem modelu.

$$\mathcal{L}_{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{2.4}$$

gdzie N oznacza liczbę próbek, y_i rzeczywistą wartość zmiennej wyjściowej, a $\hat{y}_i$ wartość wyjściową przewidywaną przez model.

W przypadku zadania przydzielenia danych do określonej klasy, czyli klasyfikacji binarnej, powszechnym jest wykorzystanie entropii krzyżowej BCE (ang. *Binary Cross-Entropy*) (2.5):

$$\mathcal{L}_{BCE} = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)\right] \tag{2.5}$$

gdzie $y_i \in \{0, 1\}$ oznacza rzeczywistą etykietę klasy, $\hat{y}_i \in (0, 1)$ prawdopodobieństwo przynależności do klasy pozytywnej przewidywane przez model, a N liczbę próbek w zbiorze uczącym.

W kontekście przemysłowym należy zwrócić uwagę na problem braku równowagi klas (ang. *class imbalance*), gdzie zdarzenia krytyczne, gdy są rzadkie, powodują zachowanie w którym model może osiągnąć wysoką dokładność ogólną przez przewidywanie zawsze klasy dominującej, ignorując całkowicie klasę rzadką [73]. Najprostszą metodą kompensacji jest wprowadzenie mechanizmu ważenia klas (ang. *class weighting*), gdzie próbki klasy rzadkiej otrzymują wyższy współczynnik wagi w funkcji straty, jednak metoda ta nie uwzględnia różnic w kosztach poszczególnych typów błędów. Takie rozwiązanie można by np. zastosować w przypadku próbek obciążenia sprężarek przy dysponowaniu wiedzą o tym, kiedy obciążenie nieaktywnych maszyn może się pojawić, jednak nie było potrzeby podjęcia takiej próby w niniejszej rozprawie.

2.4.2 Funkcja straty ogniskowej i warianty asymetryczne

Inną popularną architekturą funkcji straty jest funkcja straty ogniskowej FL (ang. *Focal Loss*) [74]. Jej popularne zastosowanie pozwala na wykrywanie obiektów znajdujących się na obrazach. Funkcja oferuje dynamiczne ważenie próbek w zależności od pewności predykcji modelu, przez co pozwala na rozwiązanie problemu braku równowagi klas:

$$\mathcal{L}_{FL} = -\alpha_t(1 - p_t)^\gamma \log(p_t) \tag{2.6}$$

gdzie p_t oznacza prawdopodobieństwo przypisane klasie rzeczywistej, α_t jest współczynnikiem ważenia klas, $\gamma > 0$ jest parametrem skupienia (ang. *focusing parameter*), który redukuje wagę łatwych próbek (dobrze klasyfikowanych, $p_t \to 1$) i koncentruje uczenie na trudnych przypadkach. Gdy $\gamma = 0$ funkcja straty ogniskowej redukuje się do ważonej entropii krzyżowej. Przytoczone oznaczenia α_t i γ odnoszą się do oryginalnej notacji pracy i nie są tożsame z oznaczeniami przyjętymi w rozprawie.

Bezpośrednią metodą oceny asymetrii kosztów jest tablica pomyłek. W przypadku klasyfikacji binarnej można zdefiniować cztery koszty: C_{TP}, C_{TN}, C_{FP}, C_{FN}, gdzie zazwyczaj przyjmuje się $C_{TP} = C_{TN} = 0$ (poprawna predykcja nie generuje kosztu) oraz $C_{FN} \gg C_{FP}$ dla problemów, gdzie pominięcie zdarzenia jest bardziej kosztowne niż fałszywy alarm. W takim wypadku funkcja straty minimalizowana podczas uczenia przyjmuje postać:

$$\mathcal{L}_{cost} = C_{FP} \cdot FP + C_{FN} \cdot FN \tag{2.7}$$

W kontekście sterowania sprężarkami stosunek C_{FN}/C_{FP} może być wyższy o rzędy wielkości. Zatrzymanie produkcji z powodu braku sprężonego powietrza generuje straty nieporównywalnie wyższe niż kilkuminutowe utrzymanie zbędnej sprężarki na biegu jałowym. Proponowana w niniejszej rozprawie funkcja straty `industrial` stanowi rozwinięcie tego podejścia, łącząc składową regresyjną (MSE dla ciśnienia względnego), klasyfikacyjną (BCE dla stanów binarnych) oraz asymetryczną karę za FN, dostosowaną do specyfiki układów sprężonego powietrza.

Wykorzystanie tablicy pomyłek ma jednak ograniczenie, ponieważ traktuje każdy błąd FN jednakowo, niezależnie od stopnia pewności błędnej predykcji. W literaturze spotyka się w związku z tym kary rosnące wraz z pewnością błędnej predykcji, budowane na bazie odległości wyjścia modelu od progu decyzyjnego. W niniejszej rozprawie przyjęto rozwiązanie odmienne: kara nakładana jest na decyzję po przekształceniu progowym, ponieważ sterownik nadrzędny otrzymuje rekomendację wyłączenia (lub nie), a negatywny koszt procesowy powstaje wyłącznie wtedy, gdy decyzja okaże się błędna. Uzasadnienie tego wyboru oraz jego konsekwencje dla procesu uczenia przedstawiono w rozdziale 3.

2.4.3 Asymetria w prognozowaniu szeregów czasowych

Zapotrzebowanie na sprężone powietrze w niniejszej rozprawie odzwierciedlone jest przez charakterystykę przebiegu ciśnienia oraz stan obciążenia sprężarek śrubowych w czasie. Ten zbiór zmiennych stanowi zespół nieprzewidywalnych lub przewidywalnych wzorców charakterystyk procesowych, których trafna predykcja wiąże się z asymetrycznym ryzykiem procesowym. W literaturze [75] zauważono, że jeżeli zmienność układu pozwala na adekwatną predykcję, to taka predykcja będzie wiązała się odpowiednią oceną np. ryzyka finansowego. Kolejnym bardzo istotnym opisanym aspektem jest dowiedzenie, że przewidywalność zmienności jest związana z horyzontem czasu oraz okresem rozpatrywanych dynamik. W przypadku długich horyzontów czasowych powyżej ustalonego progu predykcja może nie być możliwa. Posiadanie funkcjonalnego modelu predykcyjnego przekłada się na adekwatną ocenę ryzyka procesowego. Długość horyzontu predykcji została zbadana w sekcji analizy długości horyzontu predykcji 4.8, zjawisko nieadekwatnego początkowego modelu omówiono w sekcji studium przypadku 5.9.

Różny koszt asymetrycznego błędu predykcji obecny jest w rzeczywistości na co dzień. Często przytaczanym przykładem jest zagadnienie gazeciarza (ang. *newsvendor/newsboy problem*), który ma oszacować decyzję zakupu odpowiedniej liczby gazet na bazie relacji pomiędzy kosztem przeszacowania (niesprzedania towaru) a niedoszacowania (brak towaru do sprzedaży dla chętnych do zakupu klientów). Już w latach 50. XX wieku podjęto się zagadnienia, mającego pomóc przy optymalnym zorganizowaniu wyrobów gotowych, biorąc pod uwagę potrzebę zmaksymalizowania wartości netto produktów ze świadomością kosztu przez dostarczenie modeli matematycznych biorących pod uwagę różne koszty procesowe [76]. Zauważono również, że estymacja przy użyciu metody MSE jest bardzo wrażliwa na wartości odstające (ang. *outlier*) a założenie, że błędy w modelach mają rozkład normalny jest często życzeniowe [77]. Zaproponowana przez wspomnianych badaczy metoda uzyskuje odporność na anomalię przez minimalizację straty bezwzględnej ważonej asymetrycznie dzięki wprowadzeniu kwantyli regresyjnych. Zaproponowany kwantyl regresyjny ma asymetryczną funkcję kosztu (współczesny odpowiednik to ang. *pinball loss*) w której błędy dodatnie (niedoszacowanie predykcji) są mnożone przez wagę θ_q a błędy ujemne (przeszacowanie predykcji) przez wagę $(1 - \theta_q)$. To rozwiązanie umożliwia na wskazanie modelowi czy istotniejszy jest koszt zaniżenia, czy zawyżenia wyniku co przy samym symetrycznym MSE nie jest możliwe. W rozprawie nie zastosowano straty kwantylowej, a asymetria została wprowadzona przez kryterium walidacyjne omówione w sekcji ewaluacji 3.5.

Przykład użyteczności podejścia asymetrycznego omawia publikacja [78], która podaje praktyczne zastosowanie jako nawracający problem szacowania szczytowego poziomu wody przy projektowaniu i budowie tamy. Jeżeli poziom wody w tamie zostanie

przekroczony w związku z niedoszacowaniem to tama zostanie przerwana (kara wykładnicza) co wiąże się z katastrofą. Z drugiej strony nadmiarowy zapas prowadzi do niepotrzebnego zużycia zasobów ale nie doprowadzi do katastrofy (kara liniowa). Zaproponowanym rozwiązaniem jest użycie istniejącej wówczas funkcji straty LINEX (ang. *Linear-Exponential*) umożliwiającej definicję błędów krytycznych (człon rosnący wykładniczo) oraz niekrytycznych (człon narastający liniowo).

Kolejną perspektywą z jaką można się zetknąć w przemyśle jest podejście stochastyczne. W przypadku sterowania predykcyjnego w warunkach niepewności literatura zachęca do wykorzystania modeli MPC (ang. *Model Predictive Control*), które stwarzają pole do optymalizacji [79]. W przypadku sprężonego powietrza długie utrzymywanie maszyn na biegu jałowym w celu bycia gotowym do obciążenia jest założeniem najgorszego możliwego wariantu, którego kosztem będą znaczne straty energii. W przeglądzie omówiono podejście bycia świadomym ograniczeń modelu i zamiast bezwzględnych zasad wprowadzenie ograniczenia probabilistycznego (ang. *chance constraint*), które dopuszcza np. spadek ciśnienia przez niewielki określony czas pracy. W związku ze zmienną dynamiką układu sprężonego powietrza zalecane byłoby ciągłe uczenie modelu na bazie nowych danych pomiarowych, co pozwala na adaptację do nowych warunków. W przypadku predykcji stanu obciążenia sprężarek śrubowych w niniejszej rozprawie zrezygnowano z podejścia stochastycznego. Oczekiwana jest decyzja czy maszynę można bezpiecznie wyłączyć zamiast kary rosnącej wraz z odchyleniem od rzeczyiwstego przebiegu, czy prawdopodobieństwo wystąpienia spadku ciśnienia. Położono nacisk na otrzymanie binarnej decyzji, rezygnując z kary proporcjonalnej lub probabilistycznej na rzecz kary progowej dającej jednoznaczną decyzję. Metody stochastyczne wymagają znajomości rozkładu zapotrzebowania na sprężone powietrze, którego nie da się wiarygodnie estymować bez pomiaru przepływu. Proponowany system z założenia działa wyłącznie na pomiarze ciśnienia i stanach obciążenia maszyn.

2.5 Ocena modeli przy asymetrycznych kosztach

W zależności od specyfiki danej sprężarkowni, jej budowy oraz zmiennego profilu produkcyjnego decyzje podejmowane o obciążaniu maszyn będą się różnić. Zauważalny może być w związku z tym brak równowagi klas, a popularna krzywa ROC (ang. *Receiver Operating Characteristic*) biorąca pod uwagę zdarzenia TN będzie dawać zbyt optymistyczne wyniki. Wystarczy podać przykład maszyny, która jest maszyną rezerwową i prawie nigdy nie jest obciążana. Literatura zauważa, że w wypadku braku balansu danych najlepszym wyjściem będzie użycie oceny bazującej na PRC (precyzja-czułość ang. *Precision-Recall*) [80]. Taka ocena w związku z uwzględnieniem TP wśród wszystkich predykcji pozytywnych, pozwala na trafne przewidywanie predykowanej skuteczności klasyfikacji przy braku balansu [81]. W rozprawie zamiast krzywej ROC

wykorzystano wskaźniki bazujące na decyzjach, dokładność S, dopełnienie do czułości $S_{FN} = (1 - FN)$ jako wskaźnik bezpieczeństwa oraz odsetek decyzji nadmiarowo ostrożnych (S_{FP}), które omówiono w sekcji ewaluacji 3.5.

W przypadku operowania na danych rozłożonych w czasie (tak jak ma to miejsce w przypadku obserwacji procesu zapotrzebowania na sprężone powietrze) wystąpienie stanu anomalii może być poprzedzone określoną sekwencją. W związku z tym zamiast rozpatrywać wskaźniki oceny biorąc pod uwagę wszystkie rozpatrywane punkty, proponuje się wziąć pod uwagę przedział występowania anomalii i metryki na nich bazujące [82]. Publikacja [83] zauważa, że szeregi czasowe pojawiające się m.in. w energetyce, w celu wykrywania w nich anomalii wymagają podejścia łączącego ciągły proces uczenia oraz dokonywania predykcji. Zaproponowano w nim podział zbioru na „okna anomalii" oraz metodykę postępowania umożliwiającą obiektywną ocenę różnych modeli. Wczesna detekcja w obrębie okna jest nagradzana, a karane jest niewykrywanie anomalii (zaproponowana kara jest asymetryczna). W rozprawie zaproponowano metryki oceny bazujące na oknach decyzyjnych, czyli takich, które powiązane są ze wcześniejszym wystąpieniem zdarzenia odciążenia.

Przegląd dotyczący różnych podejść do zagadnienia wyznaczenia pozostałego okresu eksploatacji RUL (ang. *Remaining Useful Life*) różnych urządzeń technicznych podaje szereg wytycznych, które pozwalają na adekwatne zbudowanie modelu predykcyjnego [32], [52]. Podobnie jak w niniejszej rozprawie zwrócono uwagę na podejście kompleksowe wdrażane etapami. Najpierw wymagane jest utworzenie modelu, który będzie odpowiadał procesowi pod względem diagnostycznym i prognostycznym. Istotne jest m.in. określenie celu, wybranie danych wejściowych, wzięcie pod uwagę zakłóceń procesowych oraz różnych kosztów. Zauważono również, że sieci neuronowe pozwalają na modelowanie przy braku wiedzy eksperckiej, która wymagana byłaby przy budowie systemu ekspertowego. Bardzo istotne jest również posiadanie danych, które faktycznie odzwierciedlają proces. Wybór modelu również wymaga dobrania odpowiednich metryk. W rozprawie zaproponowano kompleksowe podejście dotyczącego zebrania, przetwarzania, modelowania, oceny oraz nadzoru operacyjnego z wybiórczym dotrenowywaniem po wdrożeniu jednak nie dotyczy ono zagadnienia predykcyjnego utrzymania ruchu.

2.6 Wybór modelu i strojenie hiperparametrów

Jednym ze sposobów optymalizacji hiperparametrów jest obok wyszukiwania ręcznego wykorzystanie przeszukiwania za pomocą siatki kombinacji (ang. *grid search*). W przypadku niewielkiej liczby parametrów jak np. użycie kilku parametrów w zaproponowanej funkcji straty taka metoda może okazać się wystarczająca gwarantując w przestrzeni dyskretnej optymalność [84]. Przy większej przestrzeni przeszukiwania

autorzy [85] udowadniają empirycznie, że przeszukiwanie losowe może owocować doborem modeli o takiej samej lub lepszej jakości w zestawieniu z siatką kombinacji. Ta ostatnia również nie sprawdzi się w przypadku wielowymiarowości.

Lepszym sposobem wyszukiwania optymalnych hiperparametrów, w związku ze zbyt dużą siatką przeszukiwania, jest algorytm optymalizacji parametrów metodą TPE (ang. *Tree-Structured Parzen Estimator*) [86]. Algorytm skupia się na przeszukiwaniu w obszarach, w których prawdopodobieństwo poprawy wyniku jest największe. Autorzy zauważają również, że często niektóre hiperparametry nie mają istotnego znaczenia dla wyniku końcowego. Metoda ta jest również efektywna obliczeniowo oszczędzając przez to zasoby komputerowe.

Przy prowadzeniu badań nad rozprawą problemem na początku było znalezienie odpowiedniej konfiguracji startowej sieci neuronowej do dalszej optymalizacji. W literaturze zauważono, że dobór odpowiednich hiperparametrów często wymagał wiedzy eksperckiej w związku z czym zaproponowano automatyczną metodę dostrajania w ramach optymalizacji bayesowskiej [86]. Autorzy dowodzą, że proponowane algorytmy przewyższają dostępne wówczas metody i pozwalają osiągnąć lub przekroczyć wyniki osiągnięte dzięki ekspertom. W kontekście niniejszej rozprawy, optymalizacja bayesowska jest drugim, opcjonalnym etapem proponowanej metodyki.

Ostatnim aspektem, jaki należy wziąć pod uwagę jest kryterium walidacyjne uczonej sieci. Zgodnie z opisanymi wcześniej metodami można otrzymać odpowiadający problemowi model. W związku z zakłóceniami występującymi w procesie należy wybrać model z takiej epoki [87], który będzie spełniał zdefiniowane wymagania. W tym celu można wykorzystać mechanizm wcześniejszego zatrzymania (ang. *early stopping*) co pozwoli na przerwanie procesu uczenia po znalezieniu minimum spełniającego zdefiniowane założenia. Ponieważ w rzeczywistości podczas uczenia minimów może być wiele, do tego również może dochodzić do przeuczenia modelu, kryterium wcześniejszego zatrzymania odgrywa bardzo ważną rolę. W proponowanej w rozprawie metodyce wcześniejsze zatrzymanie zostało wykorzystane z asymetrycznym kryterium walidacyjnym, pozwalając na znalezienie optymalnej sieci.

2.7 Dryft koncepcyjny i uczenie w eksploatacji

W przypadku sprężarkowni na jej działanie będą miały wpływ różne czynniki a jej charakterystyka może zmieniać się w ciągu roku zarówno w związku ze zmianą pory roku jak i zmianą profilu produkcji zakładu czy awariami maszyn. W takim wypadku efektywność wdrożonego modelu może spadać z czasem a zaobserwowane zjawisko można opisać pojęciem dryftu koncepcyjnego. Literatura proponuje przegląd rozwiązań w zakresie wykrywania dryftu koncepcyjnego oraz adaptacji do nowych warunków [88]. Ważnym aspektem jest zdawanie sobie sprawy z tego zjawiska i uwzględnienie go przy

projektowaniu (co zostało uwzględnione w niniejszej rozprawie). Publikacja [89] podejmuje się taksonomii zjawiska dryftu jak i przedstawienia strategii postępowania. Zgodnie z nią sezonowość oznaczałaby dryft powtarzalny (ang. *recurring concepts*), zmiana profilu produkcji lub zużycie maszyn to dryft inkrementalny (ang. *incremental drift*), stopniowy (ang. *gradual drift*) występowałby gdy nowy i stary stan występują naprzemiennie w związku np. z awarią i powrotem z awarii określonej maszyny. Ostatnim rodzajem to nagła skokowa zmiana (ang. *sudden drift*) gdy dotychczasowe wzorce przestają się pojawiać, w przypadku sprężonego powietrza, oznaczałoby to np. wymianę jednej lub kilku maszyn sprężarkowni. Zalecane strategie postępowania to podejście aktywne, takie jak w niniejszej rozprawie, gdy wskaźnik nadzoru operacyjnego ciągle monitoruje efektywność podejmowanych decyzji a jego praktyczna realizacja została zrealizowana w rozdziale 5. Drugie podejście jest pasywne, bez używania detektora ale z ciągłym dotrenowywaniem. W przytoczonej publikacji zaproponowane jest zarządzanie modelami i przełączanie np. w związku z sezonowością bieżącego modelu na model bardziej efektywny.

2.8 Identyfikacja luki badawczej

Na wstępie należy zaznaczyć, że producenci sprężarek śrubowych, którzy dostarczają dedykowane sterowniki nadrzędne, rzadko publikują szczegóły dotyczące algorytmów sterowania. Potwierdzają oczywiście spełnienie wspomnianych na wstępie norm jak i dostarczanie, ich zdaniem, najlepszych możliwych rozwiązań sterowania. Otrzymanie charakterystyk energetycznych proponowanych maszyn w rzeczywistych warunkach, również może być trudne, a przede wszystkim należy zwrócić uwagę, że dane podawane są dla normy ISO 1217:2009 [90] podczas gdy rzeczywiste warunki pracy wymagają uwzględnienia zmian sezonowych, w szczególności temperatury powietrza wlotowego. W związku z tym przegląd literatury w tym obszarze opiera się w znacznej mierze na normach ISO, wytycznych branżowych oraz doświadczeniach własnych autora zdobytych podczas wdrożeń przemysłowych.

Przeprowadzony przegląd literatury wskazuje, że istniejące podejścia do optymalizacji układów sprężonego powietrza koncentrują się na dwóch kierunkach. Pierwszy z nich obejmuje klasyczne systemy sterowania nadrzędnego oparte na deterministycznych regułach sekwencjonowania maszyn [25], które nie posiadają zdolności predykcyjnych i reagują wyłącznie na bieżący stan układu. Drugi kierunek obejmuje metody uczenia maszynowego stosowane głównie w predykcyjnym utrzymaniu ruchu oraz predykcji zużycia energii na podstawie danych z liczników i przepływomierzy [23], [52].

TABELA 2.1: Porównanie wybranych prac z niniejszą rozprawą. ✓ — cecha występuje, - — nie występuje.

Praca	Predykcja stanów bin.	Sieć dwukier.	Asymetria kosztów	Predykcja ciśnienie	Ocena zdarzeniowa	Walidacja przemysł.
Sterowanie układami sprężonego powietrza						
Salomone-González i in. (2025) [25]	–	–	–	✓	–	✓
Cabello Eras i in. (2020) [14]	–	–	–	–	–	✓
U.S. Department of Energy (2016) [13]	–	–	–	–	–	✓
Mesbah (2016) [79]	–	–	✓	–	–	–
Uczenie maszynowe w układach sprężonego powietrza						
Bezold i in. (2025) [23]	–	–	–	–	–	✓
Serradilla i in. (2020) [32]	–	–	–	–	–	✓
Khan i in. (2026) [52]	–	–	–	–	–	–
Predykcja szeregów czasowych w energetyce						
Kong i in. (2019) [24]	–	–	–	–	–	–
Pavlatos i in. (2023) [64]	–	✓	–	–	–	–
Alizadegan i in. (2025) [70]	–	✓	–	–	–	–
Wang i in. (2023) [65]	–	–	–	–	–	–
Uczenie wrażliwe na koszt						
Elkan (2001) [72]	–	–	✓	–	–	–
He i Garcia (2009) [73]	–	–	✓	–	–	–
Lin i in. (2017) [74]	–	–	✓	–	–	–
Asymetria kosztu w prognozowaniu						
Arrow i in. (1951) [76]	–	–	✓	–	–	–
Koenker i Bassett (1978) [77]	–	–	✓	–	–	–
Zellner (1986) [78]	–	–	✓	–	–	–
Ocena i detekcja zdarzeń w szeregach czasowych						
Lavin i Ahmad (2015) [83]	–	–	✓	–	✓	–
Tatbul i in. (2019) [82]	–	–	–	–	✓	–
Prace własne						
Kasprzyk i Gałuszka (2023) [18]	–	–	–	✓	–	✓

Praca	Predykcja stanów bin.	Sieć dwukier.	Asymetria kosztów	Predykcja ciśnienie	Ocena zdarzeniowa	Walidacja przemysł.
Kasprzyk i Gałuszka (2024) [68]	✓	–	–	✓	–	✓
Kasprzyk (2026) [71]	✓	✓	✓	✓	✓	✓
Niniejsza rozprawa	✓	✓	✓	✓	✓	✓

Na podstawie przeglądu literatury zidentyfikowano następujące luki badawcze:

1. brak prac stosujących jawną asymetrię kosztów do predykcji stanów binarnych jako podstawy decyzji sterujących w krótkim horyzoncie wraz z walidacją produkcyjną,

2. istniejące metody predykcji bazują na rozbudowanym opomiarowaniu (liczniki energii, przepływomierze), niedostępnym często od strony sterownika nadrzędnego w większości instalacji przemysłowych (dodatkowe opomiarowanie łączy się z dodatkowym kosztem) — brak podejść operujących wyłącznie na pomiarze ciśnienia i stanach binarnych sprężarek,

3. niedostateczna analiza wpływu horyzontu predykcji na jakość decyzji sterujących w układach sprężarkowych w kontekście zabezpieczenia przy wykorzystaniu czasu rozprężania maszyn,

4. brak uwzględnienia dryftu koncepcyjnego przy doborze optymalnej długości okna treningowego dla danych przemysłowych o zmiennym profilu produkcji,

5. niewystarczająca dokumentacja praktycznych wdrożeń systemów predykcyjnych w skali wieloobiektowej w rzeczywistych warunkach przemysłowych.

Niniejsza praca stanowi odpowiedź na te luki przez opracowanie dedykowanych funkcji straty: łączącej w sobie śledzenie wartości analogowej i predykację stanów binarnych, oraz asymetryczną funkcję straty uwzględniającą nierównomierny koszt. Praca daje również wytyczne co do budowy własnych funkcji straty, lepiej dostosowanych do specyfiki danego obiektu. Zaproponowano również systematyczną metodykę ewaluacji, która wprowadza metryki (S_{FN}, S_{FP}) pozwalające na ocenę bezpieczeństwa wdrożenia, wyznaczane dla rzeczywistych zdarzeń tranzycji odciążenia. Walidację wykonano na 12 obiektach przemysłowych, z wdrożeniem produkcyjnym na 8 z nich — bez konieczności rozbudowywania opomiarowania obiektu.

Wyniki prac badawczych prowadzonych w ramach niniejszej rozprawy były publikowane w latach 2022–2026. Wczesne prace powiązane były z wykorzystaniem środowiska Matlab i próbowały znaleźć wzorce w rozpatrywanych charakterystykach zapotrzebowania na sprężone powietrze, jak i uniwersalną architekturę [18], [91]. Badane były

także potencjalne możliwości oszczędności energii elektrycznej dla różnych grup układów popularnych w przemyśle [15]. Następnie w związku z planowanym wdrożeniem zdecydowano się na przejście do środowiska Python. W wyniku zmiany środowiska programistycznego oraz odkrycia nowych zależności pojawiły się kolejne publikacje [53], [68], [69]. Sprawdzono również klasyczne podejście do modelowania przy użyciu danych bazujących na licznikach energii w sektorze produkcji powietrza przez dmuchawy [16]. W 2026 roku doświadczenia z wdrożeń zostały opisane oraz przedstawione na konferencji krajowej [71].

Zestawienie publikacji związanych bezpośrednio z realizacją rozprawy przedstawiono w tabeli 2.2.

Tabela 2.2: Publikacje własne związane z realizacją rozprawy.

Rok	Tytuł publikacji	Miejsce publikacji
2022	*LSTM networks in prediction of the demand for compressed air depending on the type of production*	ICONST EST 2022, Budva, Montenegro
2023	*Prediction of compressed air demand depending on the type of production with the use of neural networks* DOI: 10.12913/22998624/165989	Advances in Science and Technology Research Journal, vol. 17, nr 3, s. 154–159
2023	*Energy savings possibilities gained from neural network application in compressed air supervisory control systems* DOI: 10.1109/MMAR58394.2023.10242517	27th International Conference on Methods and Models in Automation and Robotics (MMAR), s. 279–285
2024	*Frequency driven screw compressors state prediction in modern air control system*	European Simulation and Modelling Conference 2024, Ostend, s. 252–258
2024	*Predykcja stanów sprężarek śrubowych z użyciem sieci LSTM w celu wsparcia nadrzędnych systemów sterowania nowoczesnych przemysłowych układów sprężonego powietrza*	*Automatyzacja procesów dyskretnych. Teoria i zastosowania*, s. 133–144
2025	*Frequency driven screw compressors unload state prediction in modern air control system*	23rd Industrial Simulation Conference 2025, Skövde, Sweden, s. 61–65

Rok	Tytuł publikacji	Miejsce publikacji
2026	*Redukcja zużycia energii elektrycznej w przemysłowych układach sprężonego powietrza z wykorzystaniem uczenia głębokiego — doświadczenia z wdrożeń*	*Automatyzacja procesów dyskretnych. Teoria i zastosowania*, s. 159–169

3 Metodyka badań

Rozdział zawiera opis definicji rozpatrywanego problemu wraz ze schematem proponowanej metodyki. Przedstawia następnie szczegółowo przyjęte założenia dotyczące autorskiej asymetrycznej funkcji straty, architektury sieci dwukierunkowej BiLSTM oraz metodykę optymalizacji hiperparametrów. W rozdziale opisana jest również zaproponowana metoda ewaluacji mająca zminimalizować ryzyko wdrożeniowe. Na końcu rozdziału umieszczono informacje o wykorzystywanym środowisku procesowym badań.

3.1 Sformułowanie problemu

3.1.1 Definicja i założenia

Przedmiotem niniejszej rozprawy jest opracowanie i weryfikacja systemu predykcji stanów obciążenia sprężarek śrubowych w przemysłowych układach sprężonego powietrza, w oparciu o ograniczone dane procesowe, jak również przekazanie wytycznych co do integracji rozwiązania w ramach rozwiązań Przemysłu 4.0. Problem sformułowany jest jako zadanie predykcji wielokrokowej w przód (ang. *multi-step ahead prediction*) [92], w którym na podstawie historii obserwacji wybranych wartości w oknie czasowym L system przewiduje stany obciążenia sprężarek oraz przebieg ciśnienia w horyzoncie h kroków w przód.

Dla układu złożonego z N sprężarek śrubowych, o być może różnych charakterystykach pracy, w chwili t dostępny jest wektor obserwacji
$\mathbf{x}_t = [p(t), c_1(t), c_2(t), \ldots, c_N(t)]$, gdzie $p(t)$ oznacza ciśnienie względne w sieci sprężonego powietrza w punkcie odbioru, a $c_i(t) \in \{0, 1\}$ to stan obciążenia i-tej sprężarki. Bity są już wstępnie przefiltrowane przez warunki logiczne, które dopuszczają dane tylko w momencie, gdy układ jest włączony, wszystkie maszyny są widoczne przez sieć komunikacyjną (statn COM = 1) i nie występują na maszynach aktywne alarmy (ALARM = 0). Sekwencja wejściowa wstecz (ang. *lookback*) obejmuje L kolejnych wektorów obserwacji:

$$\mathbf{X}_t = [\mathbf{x}_{t-L+1}, \mathbf{x}_{t-L+2}, \ldots, \mathbf{x}_t] \in \mathbb{R}^{L\times(N+1)} \tag{3.1}$$

Sekwencja docelowa (horyzont predykcji) obejmuje h kolejnych wektorów bezpośrednio następujących po sekwencji wejściowej:

$$\mathbf{Y}_t = [\mathbf{x}_{t+1}, \mathbf{x}_{t+2}, \ldots, \mathbf{x}_{t+h}] \in \mathbb{R}^{h\times(N+1)} \tag{3.2}$$

Następnie tworzone są pary $(\mathbf{X}_t, \mathbf{Y}_t)$ metodą przesuwającego się okna (ang. *sliding window*) z krokiem 1 próbki. Przesunięcie wynika z założenia, że przyszłe zdarzenia, będą wynikały z sekwencji wejściowej w przeszłości. Dla szeregu o długości T powstaje $T - L - h + 1$ par treningowych. Model uczy się zatem odwzorowania:

$$\hat{\mathbf{Y}}_t = f_\theta(\mathbf{X}_t) \tag{3.3}$$

gdzie parametry θ optymalizowane są tak, aby minimalizować zaproponowaną asymetryczną funkcję straty. Należy zwrócić uwagę, że podejście to jest autoregresyjne — te same dane służą jako wejście ($\mathbf{X}$) oraz wyjście ($\mathbf{Y}$), będąc przesunięte względem siebie o L próbek.

Założeniem metodyki jest ograniczenie danych wejściowych do parametrów, które są niezbędne do prowadzenia prawidłowej regulacji przez sterownik nadrzędny, a zatem muszą być dostępne w jego standardowej konfiguracji. Takie rozwiązanie powoduje brak konieczności instalacji dodatkowych czujników takich jak przepływomierze sprężonego powietrza lub liczniki energii elektrycznej (czy analizatory parametrów sieci dostarczające informacji o chwilowym i długofalowym zużyciu energii w przypadku gdy uzyskanie danych z licznika jest niemożliwe). Przyjęte założenie wynika z obserwacji podczas wdrożeń, gdzie dostęp do danych jak np. z liczników energii pozostaje często po stronie klienckiej i jest niedostępny dla systemu sterowania.

3.1.2 Dane wejściowe

Na bazie przeprowadzonych eksperymentów dane wejściowe systemu ograniczono do sygnałów dostępnych w standardowej konfiguracji sprężarkowni wyposażonej w sterownik nadrzędny:

- trend ciśnienia względnego w sieci sprężonego powietrza w punkcie odbioru $p(t)$ [bar] — pomiar ciągły odzwierciedlający aktualne zapotrzebowanie na sprężone powietrze,
- stany binarne obciążenia poszczególnych sprężarek $c_i(t) \in \{0, 1\}$ — gdzie wartość 1 oznacza pracę sprężarki pod obciążeniem (tłoczenie), a wartość 0 bieg jałowy lub postój,
- opcjonalnie: gradient ciśnienia $\Delta p(t) = p(t) - p(t-1)$ jako dodatkowa cecha pochodna informująca o kierunku zmian ciśnienia.

Należy zwrócić uwagę, że istotnym warunkiem dotyczącym jakości danych treningowych, jest prawidłowe działanie układu w momencie ich rejestracji. Próbki zebrane

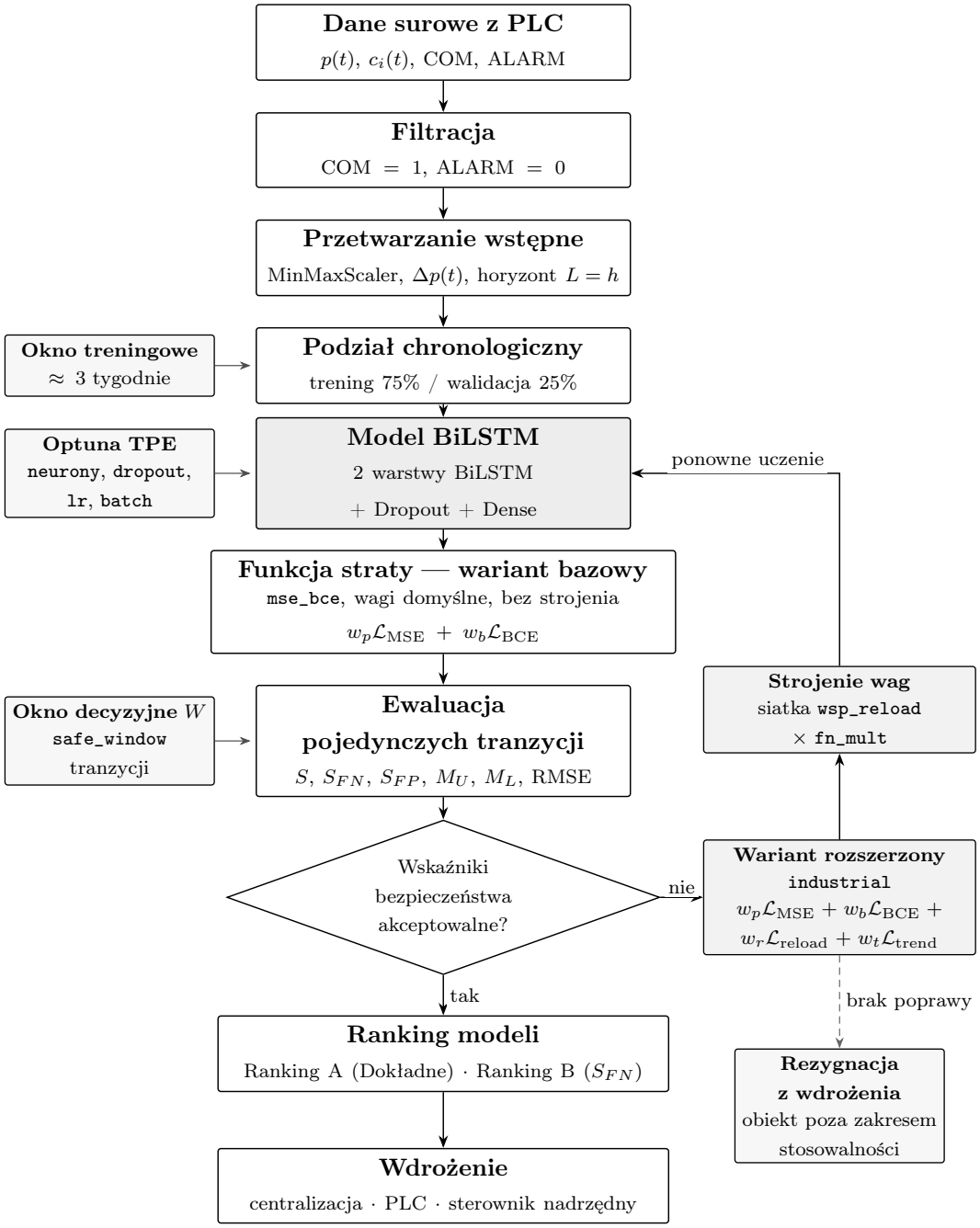
podczas awarii losowych sprężarek, braki komunikacji maszyn ze sterownikiem nadrzędnym, aktywne alarmy, stany przejściowe po włączeniu systemu (dobijanie ciśnienia w instalacji po wyłączeniu) nie reprezentują normalnych wzorców pracy i mogą zaburzać proces uczenia (w zależności od charakterystyki dynamicznej układu). W związku z tym dane treningowe poddawane są filtracji — do zbioru uczącego wprowadzane są wyłącznie próbki spełniające warunki poprawnej komunikacji (COM = 1) oraz braku aktywnych alarmów (ALARM = 0). Standardowo te sygnały są weryfikowane lub wypracowywane przez sterownik nadrzędny, rejestrowane wraz z danymi procesowymi, co umożliwia automatyczną selekcję próbek bez konieczności ingerencji operatora (natomiast przy braku wystarczającej ilości danych weryfikacja operatora systemu jest konieczna co do weryfikacji możliwości wdrożenia).

Wcześniej wspomniano, że system nie wymaga danych z liczników energii elektrycznej ani pomiarów przepływu sprężonego powietrza, co ułatwia proces wdrożenia. Dane te, mile widziane i cenne z perspektywy analizy energetycznej, są w praktyce przemysłowej często niedostępne dla systemu sterowania ze względów prawnych i/lub organizacyjnych (np. zakłady przemysłowe mogą mieć dedykowane systemy monitorujące system energetyczny z zabronieniem komunikacji tego systemu z innymi ze względu na wrażliwość danych).

Bardzo ważnym parametrem konfiguracyjnym jest okres próbkowania danych. W badaniach własnych zastosowano okres próbkowania $T_s = 20$ s, co odpowiada trzem próbkom na minutę, przy oknie lookback dobieranym symetrycznie do horyzontu predykcji, $L = h \in \{6, 9, 12, 15\}$ próbek (2–5 minut historii). Wybór ten jest kompromisem między rozdzielczością czasową niezbędną do wykrywania szybkich tranzycji a ilością danych możliwych do zebrania przez system akwizycji. Należy podkreślić, że zbyt długi okres próbkowania (powyżej 60 s) uniemożliwia adekwatne modelowanie dynamiki tranzycji odciążenia, które typowo trwa 30–120 s w rozpatrywanych obiektach. Okres ten podyktowany jest również ograniczeniami ilości danych przesyłanych z wykorzystaniem sieci komórkowej, gdzie zbyt duża częstotliwość próbkowania może generować nadmierny ruch sieciowy i zwiększać koszty transmisji danych. Bardzo ważną obserwacją tutaj również jest fakt występowania dynamiki samego układu. Wyprzedzając późniejsze wnioski, może zdarzyć się, że obiekt ma tak dużą dynamikę, że w ciągu 20 sekund, będzie dochodziło do więcej niż jednej tranzycji obciążenia/odciążenia, co będzie widoczne przez brak możliwości nauczenia się charakterystyki obiektu. W takim przypadku niezbędne jest zwiększenie częstotliwości próbkowania, nawet do czasu np. 5 s co pozwoli na uchwycenie wszystkich niezbędnych zależności. Tak intensywna praca maszyn, będzie skracała ich żywotność, ale system predykcyjny może pomóc w optymalizacji ich pracy, mimo że zaleceniem byłby audyt sprężonego powietrza oraz modernizacja obiektu.

3.1.3 Schemat metodyki

Rysunek 3.1 przedstawia schemat blokowy proponowanej metodyki, obejmujący etapy od akwizycji danych do wdrożenia produkcyjnego systemu predykcyjnego.



Rysunek 3.1: Schemat blokowy proponowanej metodologii predykcji stanów obciążenia sprężarek śrubowych. Postępowanie rozpoczyna się od wariantu bazowego funkcji straty z wagami domyślnymi; wariant rozszerzony wraz ze strojeniem wag wprowadzany jest dopiero wtedy, gdy ocena wskaźników bezpieczeństwa okaże się niewystarczająca.

3.2 Autorska asymetryczna funkcja straty

3.2.1 Koncepcja i uzasadnienie

Standardowe funkcje straty stosowane przy klasyfikacji binarnej — binarna entropia krzyżowa (BCE) oraz błąd średniokwadratowy (MSE) — traktują wszystkie błędy predykcji jako równoważne. W przypadku problemu predykcji stanów obciążenia sprężarek zaproponowano odejście od tego założenia i wprowadzenie wag, które odzwierciedlają istotność każdego członu. Entropia krzyżowa będzie odpowiadać klasyfikacji zachowania układu sterowania, natomiast błąd średniokwadratowy będzie odpowiadać regresji ciśnienia sprężonego powietrza. Wprowadzenie wag pozwala na ustalenie, który z członów funkcji straty jest istotniejszy w procesie uczenia i lepiej odzwierciedla rzeczywistą dynamikę układu.

Fałszywie negatywna predykcja stanu obciążenia (FN) — czyli sytuacja, gdy model przewiduje możliwość wyłączenia sprężarki, podczas gdy jest ona potrzebna do utrzymania ciśnienia w sieci, może prowadzić do niedoboru sprężonego powietrza i zatrzymania produkcji. Z drugiej strony fałszywie pozytywna predykcja (FP) niesie ze sobą mniejsze ryzyko operacyjne, prowadzi do zbędnego utrzymywania sprężarki na biegu jałowym czego wynikiem jest „tylko" możliwość utraty oszczędności energii. Koszt takiego błędu jest dużo niższy niż ryzyko zachwiania stabilności dostaw sprężonego powietrza. Jednak tutaj należy również zwrócić uwagę na fakt, że nie zawsze tak będzie, że decyzja utrzymania sprężarki na biegu jałowym będzie bezpieczna, w skrajnych przypadkach może prowadzić do wyłączenia maszyny po jakimś odliczonym czasie, wyłączenia jej, a następnie braku możliwości rozruchu przez czas rozprężania co też może prowadzić do spadku ciśnienia.

Gdyby potraktować powyższe uwagi jak koszty [72] to w kontekście sterowania sprężarkami koszt fałszywie negatywnej predykcji C_{FN} — obejmujący straty wynikające z zatrzymania produkcji, ponownego uruchomienia maszyn oraz potencjalnych kar umownych — jest wielokrotnie wyższy niż koszt fałszywie pozytywnej predykcji C_{FP} [74], sprowadzającego się jedynie do zbędnego zużycia energii elektrycznej podczas biegu jałowego sprężarki (ze wspomnianym ograniczeniem).

$$C_{FN} \gg C_{FP} \tag{3.4}$$

3.2.2 Składowe funkcji straty

W odpowiedzi na obserwowany proces dostaw sprężonego powietrza, zaproponowano autorską funkcję straty, w której każdy człon niesie w sobie informację, która

pozwala na lepsze odzwierciedlenie dynamiki układu oraz ma umożliwić bardziej adekwatne uczenie modelu. W części własnych publikacji przedstawiono uproszczoną funkcję straty, która uwzględniała składowe BCE oraz MSE wraz z wagami, natomiast założeniem było, że wprowadzenie dodatkowych członów pozwoli na lepsze odzwierciedlenie dynamiki układu. Zaproponowana, pełna przemysłowa funkcja straty `industrial` jest sumą ważoną dwóch składowych uczących oraz dwóch składowych walidacyjnych, które zaproponowano w trakcie badań:

$$\mathcal{L}_{total} = w_p \cdot \mathcal{L}_{MSE} + w_b \cdot \mathcal{L}_{BCE} + w_r \cdot \mathcal{L}_{reload} + w_t \cdot \mathcal{L}_{trend} \tag{3.5}$$

gdzie domyślnie $w_p = 0{,}15$, $w_b = 0{,}10$, $w_r = 0{,}60$, $w_t = 0{,}15$.

Składowa regresyjna $\mathcal{L}_{MSE}$ minimalizuje błąd predykcji ciśnienia:

$$\mathcal{L}_{MSE} = \frac{1}{T \cdot h} \sum_{t=1}^{T} \sum_{\tau=1}^{h} \big(p(t+\tau) - \hat{p}(t+\tau)\big)^2 \tag{3.6}$$

Obecność tej składowej stabilizuje uczenie przez ciągły sygnał gradientu nawet w okresach bez zmian obciążenia. Składowa odzwierciedla okresy zapotrzebowania i jego braku. W praktyce nadążanie za tą składową pozwala na lepsze dopasowanie modelu do charakterystyki.

Składowa klasyfikacyjna $\mathcal{L}_{BCE}$ minimalizuje błąd predykcji stanów binarnych sprężarek:

$$\mathcal{L}_{BCE} = -\frac{1}{T \cdot h \cdot N} \sum_{t=1}^{T} \sum_{\tau=1}^{h} \sum_{i=1}^{N} \big[c_i \log \hat{c}_i + (1-c_i) \log(1-\hat{c}_i)\big] \tag{3.7}$$

Założeniem było tutaj, że wprowadzenie składowej klasyfikacyjnej pozwala na uchwycenie pewnych trendów zapotrzebowania, które nie są łatwo identyfikowane przez trend ciśnienia. Czyli oddaje pewne wzorce, które mogą stanowić trzon lub niezmienność obiektu.

Składowa wykrywania ponownego obciążenia $\mathcal{L}_{reload}$ wprowadza asymetryczną karę za błędy predykcji tranzycji odciążenia. Kara za FN jest α razy większa niż za FP:

$$R_{t,\tau,i} = \begin{cases} 1, & \text{gdy } c_i(t+\tau) < 0{,}5 \ \wedge \ c_i(t+\tau+1) \geq 0{,}5 \\ 0, & \text{w przeciwnym przypadku} \end{cases} \tag{3.8}$$

$$\hat{b}_{t,\tau,i} = \begin{cases} 1, & \text{gdy } \hat{c}_i(t+\tau+1) \geq 0{,}5 \\ 0, & \text{w przeciwnym przypadku} \end{cases} \tag{3.9}$$

$$\mathcal{L}_{reload} = \frac{1}{T \cdot (h-1) \cdot N} \sum_{t,\tau,i} \Big[\alpha \cdot R_{t,\tau,i} \cdot \big(1 - \hat{b}_{t,\tau,i}\big) + \beta \cdot \big(1 - R_{t,\tau,i}\big) \cdot \hat{b}_{t,\tau,i} \Big] \tag{3.10}$$

gdzie R oznacza rzeczywiste zdarzenie powrotu obciążenia ($0 \to 1$), $\hat{b}$ oznacza decyzję modelu po przekształceniu progowym (zaokrąglenie do wartości binarnej), α jest mnożnikiem kary za FN (domyślnie $\alpha = 12$), a β mnożnikiem kary za FP (domyślnie $\beta = 1$). Wartość progu zaokrąglenia w proponowanej metodyce jest stały i wynika z równoważnego kosztu przydziału (wpływ zaokrąglenia na wyniki badano w publikacji [15]), a kwestia kar proporcjonalnych [77], [78] została omówiona przy przeglądzie literatury w sekcji asymetrycznego kosztu 2.4.3. W składowej $\mathcal{L}_{reload}$ błędy FN i FP wyznaczane są dla pojedynczych kroków horyzontu predykcji, ponieważ funkcja straty operuje na wyjściu modelu, a nie na zdarzeniach, natomiast wskaźniki oceny modelu (podrozdział 3.5) odnoszą się do zdarzeń decyzyjnych w oknach tranzycji.

Działanie składowej $\mathcal{L}_{reload}$ wymaga wyjaśnienia w odniesieniu do rozpatrywanego modelu sterowania oraz mechanizmu uczenia. Sterownik nadrzędny otrzymuje tylko i wyłącznie rekomendację wyłączenia, stany pośrednie jak np. cząstkowe prawdopodobieństwo wyłączenia nie są tutaj istotne, istotne jest to, czy należy wyłączyć sprężarkę (nie występuje stan częściowego wyłączenia sprężarki). Z tego powodu kara w składowej nakładana jest nie na surowe wyjście sieci, a na decyzję otrzymaną po przekształceniu progowym. W związku z tym proponowana kara, skupiona jest na koszcie procesowym, spowodowanym błędną decyzją, a nie na odchyleniu predykcji od żądanego poziomu. Niepewna, ale trafna decyzja jest traktowana tak, jak decyzja podjęta z dużą pewnością. Składowe funkcji straty $\mathcal{L}_{reload}$ oraz $\mathcal{L}_{trend}$ nie uczestniczą w procesie propagacji wstecznej, lecz zostają wyznaczone na zbiorze walidacyjnym, odpowiadając za mechanizm wcześniejszego zatrzymania procesu uczenia i selekcji modelu do wdrożenia. Na proces uczenia największy wpływ ma składowa regresyjna oraz klasyfikacyjna (krok zejścia gradientowego), odwzorowujące przebieg procesu. Podsumowując, zadaniem uczenia jest odwzorowanie dynamiki układu, a wybór modelu produkcyjnego odbywa się na podstawie walidacji, a zatem rzeczywistego kosztu procesowego. Założenie asymetrii kosztów (3.4) (odwzorowane przez wagę w_r oraz mnożnik α) realizowane jest na poziomie wyboru modelu przeznaczonego do wdrożenia. Rozwiązanie zostało sprawdzone po implementacji i było zgodne z oczekiwaniami. Jeżeli za wagi innych składowych przyjąć zero, składowe związane z przekształceniem progowym przyjmują niezerowe wartości straty przy zerowym gradiencie względem wyjścia modelu, natomiast składowe regresyjna i klasyfikacyjna dają gradient niezerowy. Jako kierunek dalszych badań w rozdziale 7.5 wskazano wariant bazujący na różniczce zamiast przekształcenia progowego.

Składowa trendu ciśnienia $\mathcal{L}_{trend}$ ma za zadanie wykrywać momenty w horyzoncie predykcji, w których ciśnienie spada, co wiążę się ze wzrostem zapotrzebowania na sprężone powietrze i karząc za występujące wówczas błędy krytyczne. Składowa oddziałuje przez kryterium wyboru modelu produkcyjnego. Empirycznie zaobserwowano, że jej wpływ zależy od charakterystyki obiektu — dla jednego z obiektów wyposażonych w falownik prowadziła do wyboru modelu zbyt zachowawczego co opisano w rozdziale 4.

$$D_{t,\tau} = \begin{cases} 1, & \text{gdy } p(t+\tau+1) < p(t+\tau) \\ 0, & \text{w przeciwnym przypadku} \end{cases} \tag{3.11}$$

$$\mathcal{L}_{trend} = \frac{1}{T \cdot (h-1) \cdot N} \sum_{t=1}^{T} \sum_{\tau=1}^{h-1} \sum_{i=1}^{N} D_{t,\tau} \cdot R_{t,\tau,i} \cdot \left(1 - \hat{b}_{t,\tau,i}\right) \tag{3.12}$$

Maska D jest wyznaczona z rzeczywistego przebiegu, a nie z predykcji, przez co składowa trendu stanowi podzbiór zdarzeń karanych przez $\mathcal{L}_{reload}$, zawężony do momentów spadku ciśnienia.

3.2.3 Wariant bazowy mse_bce

Autor uważa, że sama składowa regresyjna, jest niewystarczająca w procesie uczenia działania procesu dostarczania sprężonego powietrza stąd bazowy wariant asymetrycznej funkcji straty zawiera dwie składowe wraz z wagami `mse_bce`:

$$\mathcal{L}_{mse_bce} = w_{MSE} \cdot \mathcal{L}_{MSE} + w_{BCE} \cdot \mathcal{L}_{BCE}, \quad w_{MSE} + w_{BCE} = 1 \tag{3.13}$$

Domyślne wartości wag to $w_{MSE} = 0{,}3$ oraz $w_{BCE} = 0{,}7$ [69].

Wariant ten nie zawiera asymetrycznej kary za FN i służy jako podstawa dla oceny wkładu składowych `industrial`.

3.3 Architektura sieci BiLSTM

3.3.1 Uzasadnienie wyboru architektury

Artykuły z własnych publikacji głównie bazowały na sieci LSTM [47], jednak w trakcie badań, zgodnie z wykonanym przeglądem literatury, zauważono to, co inni badacze w sektorze energetycznym, czyli, że sieć dwukierunkowa (BiLSTM [54]) w krótkich horyzontach czasowych może lepiej odzwierciedlić zależności sekwencyjne. Co za tym idzie sieć może dawać lepsze wyniki predykcji, ponieważ długofalowe modelowanie jest trudne [48], [49]. Przesłanki przemawiające za wykorzystaniem BiLSTM można wypunktować w następujący sposób:

- dane procesowe ze sprężarek są szeregami czasowymi o silnych zależnościach sekwencyjnych,
- tranzycje odciążenia posiadają charakterystyczny kształt przebiegu ciśnienia zarówno poprzedzający jak i następujący po zdarzeniu — sieć dwukierunkowa widząca sekwencję w obu kierunkach może efektywniej identyfikować te wzorce,
- wyniki eksperymentów porównawczych potwierdziły przewagę BiLSTM nad LSTM i GRU[50], [51] pod względem kluczowych metryk.

3.3.2 Konfiguracja sieci

W zastosowanej architekturze sieci można wyróżnić pięć następujących warstw:

1. **Warstwa wejściowa** — tensor $[L \times (N+1)]$, gdzie $L \in \{6, \ldots, 15\}$ (lookback), $N+1$ — ciśnienie i stany N sprężarek,
2. **Dwie warstwy BiLSTM** — z u_1 i u_2 jednostkami, z warstwą Dropout między nimi,
3. **Warstwa Dense** — transformacja do wymiaru $h \cdot (N+1)$,
4. **Warstwa Reshape** — tensor $[h \times (N+1)]$,
5. **Wyjście split** — ciśnienie przez aktywację liniową, stany sprężarek przez sigmoidę $\mathrm{sigm}(\cdot)$.

3.3.3 Kierunek przetwarzania danych

Należy zwrócić uwagę na istotność właściwego zebrania oraz przetworzenia posiadanych danych, wszystkie pomiary muszą zostać odpowiednio przygotowane. Kierunek przetwarzania danych przedstawia się następująco:

1. **Filtracja** — COM = 1, ALARM = 0 (bity poprawności — semantyka jak w sekcji 3.1),
2. **Normalizacja** — MinMaxScaler fit tylko na zbiorze treningowym,
3. **Gradient ciśnienia** — opcjonalne $\Delta p(t)$,
4. **Sekwencje lookback** — okno przesuwne $L \in \{6, \ldots, 15\}$ (lookback),
5. **Podział treningowy/walidacyjny** — train 75% / val 25%.

Dla większości rozpatrywanych przypadków optymalny okres treningowy wyniósł około 3 tygodni (przy typowym 4-tygodniowym oknie danych i podziale 75%/25% odpowiada to trzem tygodniom treningu i tygodniowi walidacji) — zbiór krótszy nie zawiera dostatecznie zróżnicowanych wzorców, natomiast zbyt długi może zawierać wzorce nieaktualne ze względu na dryft koncepcyjny [55], [56].

Do zatrzymania procesu uczenia wykorzystano mechanizm wczesnego zatrzymania (ang. *early stopping*) monitorujący wartość funkcji straty wyznaczaną na zbiorze walidacyjnym [87]. W przypadku braku poprawy wartości funkcji straty przez 20 kolejnych epok, proces uczenia zostaje zatrzymany. Do wdrożenia przechodzi model, który osiągnął najlepszą stratę walidacyjną.

3.4 Metodyka optymalizacji hiperparametrów

Zaproponowana metodyka optymalizacji hiperparametrów prowadzona jest w dwóch etapach. Pierwszy etap skupia się na znalezieniu poszczególnych wag składowych funkcji straty najpierw dla wariantu `mse_bce` (składowe w_p, w_b) za pomocą przeszukiwania siatki kombinacji [85]. W przypadku gdy wskaźniki oceny w kolejnych krokach nie będą wystarczające, następuje powrót do kroku przeszukiwania siatki kombinacji dla wariantu funkcji straty `industrial` (w_p, w_b, w_r, w_t wraz z mnożnikiem kary FN α). Dobór tych parametrów związany jest z charakterystyką obiektu, a nie z właściwościami samej sieci. Przeszukiwanie odbywa się zgodnie z przyjętym kryterium selekcji modelu na bazie zbioru walidacyjnego, który zapewnia minimalizację liczby FN. Zagadnienia doboru modelu omówiono w odpowiedniej sekcji przeglądu literatury 2.6.

Drugim etapem jest optymalizacja bayesowska oparta na estymatorach Parzena o strukturze drzewiastej (ang. *TPE - Tree-structured Parzen Estimator*) [86]. W jej wypadku ustalana jest liczba jednostek w kolejnych warstwach rekurencyjnych (u_1, u_2), współczynnik regularyzacji dropout p_{drop}, współczynnik uczenia η, rozmiar batch M, a także waga składowej reload w_r, mnożnik α oraz waga składowej regresyjnej w_p.

3.4.1 Dwuetapowe podejście

Uzasadnieniem podejścia dwuetapowego jest to, że przestrzeń parametrów funkcji straty jest niewielka i możliwa w interpretacji, dzięki czemu nie wymaga dużych mocy obliczeniowych, a wpływ poszczególnych wag jest obserwowalny. Przestrzeń parametrów sieci jest bardzo rozległa, więc zdecydowano o wykorzystaniu metody adaptacyjnej. W rozdziale 4 przy porównywaniu poszczególnych obiektów, zatrzymano się na etapie pierwszym przy stałych parametrach sieci, dzięki czemu otrzymano porównywalność pomiędzy obiektami. Drugi etap znajduje zastosowanie przy strojeniu modeli dla obiektów, dla których konfiguracja bazowa okazała się niewystarczająca.

3.4.2 Siatka kombinacji

Rozważane są dwie asymetryczne funkcje straty: `mse_bce` oraz `industrial`.

Dla wariantu `mse_bce`: siatka 9 kombinacji $w_{MSE} \in \{0{,}1, \ldots, 0{,}9\}$.

Dla wariantu `industrial`: $w_r \in \{0{,}68,\ 0{,}73,\ 0{,}78,\ 0{,}83,\ 0{,}88\}$ oraz $\alpha \in \{8,\ 10,\ 12,\ 14,\ 16\}$, co daje 25 kombinacji. Konfiguracja przewiduje również wymiar $w_p \in \{0{,}10,\ 0{,}15,\ 0{,}20\}$, wykorzystywany wybiórczo dla pojedynczych obiektów.

Wagi nie są normalizowane, zmiana w_r zmienia wagę tej składowej względem pozostałych, których wartości pozostają stałe. Kryterium: minimalizacja liczby `FN`.

We wszystkich przebiegach eksperymentów pierwszy etap wskazał $\alpha = 12$ oraz $w_r = 0{,}78$ z kolei etap bayesowski, przeszukujący α w zakresie ciągłym, podnosił tę wartość dla części obiektów wdrożeniowych.

Waga w_b pozostaje ustalona i stanowi punkt odniesienia dla pozostałych wag, ponieważ składowe nie są normalizowane, a o zachowaniu modelu decyduje ich wzajemny stosunek.

3.4.3 Optymalizacja bayesowska (Optuna)

Przestrzeń poszukiwań [85], [86]:

- trzy warstwy: $u_1 \in [144, 192]$ (krok 16), $u_2 \in [64, 128]$ (krok 16), $u_3 \in [32, 64]$ (krok 8),
- dropout $p_{drop} \in [0{,}30, 0{,}45]$ (krok 0,05),
- learning rate $\eta \in [0{,}003, 0{,}008]$ (skala logarytmiczna),
- batch $M \in \{2048,\ 4096,\ 8192\}$,
- wagi straty: $w_r \in [0{,}78, 0{,}88]$ (krok 0,02), $\alpha \in [14{,}0, 20{,}0]$ (krok 0,5), $w_p \in [0{,}18, 0{,}26]$ (krok 0,02).

Kryterium: minimalizacja `composite_score_with_fn` łączącego S_{FN} z ogólną dokładnością, 30 prób na obiekt.

3.5 Metodyka ewaluacji

3.5.1 Okno decyzyjne

W sekcji 2.1.2 poruszono temat czasu rozprężania maszyny, przez okres którego sprężarka nie może zostać ponownie uruchomiona. Zatrzymanie i rozprężanie stwarza ryzyko zachwiania stabilności dostaw sprężonego powietrza w przypadku nietrafionej decyzji. W związku z tym przyjęto założenie, że model ma zapewnić jak największe

bezpieczeństwo właśnie dla tego okresu. Koncepcja ta skutkuje wprowadzeniem trzech różnych wielkości czasowych. Pierwszą z nich jest długość okna obserwacji L, która określa liczbę próbek historycznych podawanych na wejście sieci. Drugą wartością jest horyzont predykcji h określający liczbę przewidywanych przez sieć próbek (na całej jego długości wyznaczany jest błąd RMSE). Ostatnią wartością jest okno decyzyjne W definiujące na ilu pierwszych przewidywanych próbkach ma opierać się decyzja wyłączenia. Wskaźniki bezpieczeństwa S, S_{FN} oraz S_{FP} wyznaczane są wyłącznie dla okna W. Przyjęta długość okna decyzyjnego, wynikająca zatem z przyjętego z zapasem czasu rozprężania maszyny, wynosi 2 min, co przy okresie próbkowania 20 s odpowiada 6 próbkom, niezależnie od przyjętego horyzontu predykcji.

Tabela 3.1: Parametry czasowe modelu i ich rola w ocenie dla przykładowego okresu próbkowania $T_s = 20$ s.

Oznaczenie	Znaczenie	Rola	Przykład
L	okno obserwacji	wejście sieci	12 próbek (4 min)
h	horyzont predykcji	wyjście sieci, RMSE	12 próbek (4 min)
W	okno decyzyjne	decyzja, S, S_{FN}, S_{FP}	6 próbek (2 min)

Ewaluacja prowadzona jest w kontekście wykrytych tranzycji odciążenia ($1 \rightarrow 0$), ponieważ to właśnie w tych momentach sterownik nadrzędny może podjąć decyzję o wyłączeniu sprężarki. Dla każdej tranzycji definiowane jest okno decyzyjne $W = [t + w_{start},\ t + w_{end}]$ wokół rzeczywistego momentu tranzycji t (domyślnie $w_{start} = 0$, $w_{end} = 6$ próbek). Przyjęcie okna decyzyjnego zamiast oceny na poziomie pojedynczych próbek odpowiada stosowanej w praktyce ocenie detekcji zdarzeń w szeregach czasowych, gdzie trafność odnoszona jest do przedziału czasu obejmującego zdarzenie, a nie do pojedynczej chwili [82], [83]. Przyczyny takiego wyboru omówiono w sekcji 2.5.

Użyte wskaźniki odnoszą się do decyzji w ramach jednego rozpatrywanego zbioru tranzycji, a nie wszystkich próbek w ramach okna sekwencji. Decyzja FN oznacza, że sieć rekomendowała wyłączenie maszyny, podczas gdy rzeczywistość wskazała na konieczność jej uruchomienia. Wskaźniki FP oraz FN niosą w sobie informację o rzeczywistych kosztach błędnych decyzji systemu sterowania. Rzeczywista kara za zdarzenia FN jest wyższa niż w przypadku FP — dla zdarzenia FN następuje rekomendacja wyłączenia sprężarki, która w rzeczywistości jest potrzebna, co może prowadzić do spadku ciśnienia w sieci. W przypadku zdarzenia FP sprężarka zostaje utrzymywana na biegu jałowym, co niesie stratę energetyczną, lecz nie ryzykuje wstrzymania produkcji.

Dla każdej tranzycji oraz rozpatrywanej sprężarki wyznaczane są dwa wskaźniki binarne, agregujące stany w oknie decyzyjnym:

$$g = \begin{cases} 1, & \text{gdy w oknie } W \text{ występuje próbka ze stanem } c(\tau) = 1 \\ 0, & \text{w przeciwnym przypadku} \end{cases} \tag{3.14}$$

$$\hat{g} = \begin{cases} 1, & \text{gdy w oknie } W \text{ model przewiduje } \hat{c}(\tau) \geq 0{,}5 \\ 0, & \text{w przeciwnym przypadku} \end{cases} \tag{3.15}$$

gdzie g oznacza rzeczywiste wystąpienie powrotu obciążenia w oknie (maszyna będzie potrzebna), a $\hat{g}$ — powrót przewidziany przez model. Decyzyjnie $\hat{g} = 0$ oznacza rekomendację możliwości wyłączenia maszyny na czas okna. Klasyfikacja tranzycji wynika wprost z pary $(g, \hat{g})$:

TP — $(1, 1)$: powrót obciążenia wystąpił i został przewidziany — trafne utrzymanie maszyny w gotowości,

TN — $(0, 0)$: powrót nie wystąpił i nie był przewidziany — trafna rekomendacja wyłączenia, zrealizowana oszczędność energii,

FP — $(0, 1)$: przewidziano powrót, który nie nastąpił — nadmierna ostrożność, utracona możliwość oszczędności,

FN — $(1, 0)$: powrót wystąpił, lecz nie został przewidziany — rekomendacja wyłączenia maszyny, która okazała się potrzebna; błąd krytyczny dla bezpieczeństwa dostaw.

Wypada jeszcze wspomnieć o jednym aspekcie, a mianowicie klasie TN. Ta klasa, to właśnie jedyna klasa realizująca oszczędność. Gdy model przewiduje brak powrotu obciążenia, a maszyna faktycznie nie wraca — sterownik wyłączając ją oszczędza energię elektryczną przez pominięcie utrzymywania maszyny na biegu jałowym.

3.5.2 Metryki ewaluacyjne i bezpieczeństwa

W związku z brakiem równowagi klas zrezygnowano ze wskaźników oceny bazujących na krzywej ROC na rzecz miar bazujących na precyzji i czułości [80], [81]. Wprowadzono następujące metryki ewaluacyjne [71]:

$$S = \frac{TP + TN}{TP + TN + FP + FN} \cdot 100\% \tag{3.16}$$

$$S_{FN} = \frac{FN}{TP + FN} \cdot 100\% \tag{3.17}$$

$$S_{FP} = \frac{FP}{TP + TN + FP + FN} \cdot 100\% \tag{3.18}$$

gdzie TP (prawdziwie pozytywne, ang. *True Positive*) oznacza liczbę trafnych predykcji obciążenia, TN (prawdziwie negatywne, ang. *True Negative*) trafnych predykcji biegu jałowego, FP błędnych predykcji obciążenia, a FN — błędnych predykcji biegu jałowego, będących najgroźniejszym rodzajem błędu.

Wskaźnik S wyraża ogólną skuteczność predykcji w oknie tranzycji — wartość 100% oznacza, że dla każdej analizowanej tranzycji decyzja wynikająca z predykcji w oknie była trafna.

Wskaźnik S_{FN} jest warunkową miarą ryzyka — informuje o odsetku niewykrytych powrotów obciążenia, czyli sytuacji, w których model rekomendował możliwość wyłączenia sprężarki, podczas gdy w oknie decyzyjnym okazała się ona ponownie potrzebna. Wartość 0% jest stanem idealnym, oznaczającym, że żadna rekomendacja wyłączenia nie naraziła układu na spadek ciśnienia. Wartość 100% oznacza, że model nie wykrył żadnego z występujących powrotów obciążenia — rekomendacje wyłączeń dla takiej maszyny nie mogą stanowić samodzielnej podstawy decyzji (przypadek charakterystyczny m.in. dla sprężarek wiodących, których powroty do obciążenia są zdarzeniami rzadkimi [73] i trudnymi do przewidzenia). Wyniki na poziomie np. 60% wskazują, że blisko 60% powrotów obciążenia pozostało niewykrytych, co wymaga wprowadzenia regułowych zabezpieczeń po stronie sterownika nadrzędnego.

Wskaźnik S_{FP} określa odsetek decyzji nadmiarowo ostrożnych — sytuacji, w których model przewidział powrót obciążenia, który w oknie decyzyjnym nie nastąpił, przez co sprężarka pozostała utrzymana w gotowości, mimo że mogła zostać wyłączona. Wartość bliska 0% oznacza pełne wykorzystanie dostępnych możliwości oszczędności energii. Wzrost tego wskaźnika nie zagraża bezpieczeństwu dostaw sprężonego powietrza, ogranicza natomiast uzyskiwane oszczędności energii elektrycznej.

Uzupełniająco wprowadzono dwie metryki dopasowania dokładnego, wyznaczane odrębnie dla dwóch populacji zdarzeń — tranzycji odciążenia oraz tranzycji obciążenia:

$$M_U = \frac{N_T^{exact}}{N_T} \cdot 100\% \tag{3.19}$$

$$M_L = \frac{N_L^{exact}}{N_L} \cdot 100\% \tag{3.20}$$

gdzie N_T i N_L oznaczają liczbę wykrytych tranzycji odpowiednio odciążenia ($1 \to 0$) i obciążenia ($0 \to 1$), a N_T^{exact}, N_L^{exact} — liczbę tych tranzycji, dla których przewidziana sekwencja stanów w całym oknie horyzontu jest identyczna z rzeczywistą (dopasowanie dokładne, ang. *exact match*). Metryki M_U i M_L oceniają wierność odwzorowania całego przebiegu stanów, przechodząc próbka po próbce.

Wskaźnik M_U informuje o jakości odwzorowania zachowania układu wokół tranzycji odciążenia — wartość 100% oznacza, że każda przewidziana sekwencja idealnie pokrywa się z rzeczywistym przebiegiem stanów.

Wskaźnik M_L ocenia analogicznie odwzorowanie przebiegów wokół tranzycji obciążenia — niska wartość wskazuje, że model ma trudności z wiernym odwzorowaniem momentu wznowienia tłoczenia, co jest spójne z podwyższonym S_{FN} dla maszyn o trudno przewidywalnych powrotach obciążenia.

3.5.3 Ranking modeli

Dla wyników eksperymentów wprowadzono system dwóch kryteriów oceny, które odpowiadają innym priorytetom w kontekście wdrożenia.

Ranking A oparty jest na sprawdzeniu dopasowania dokładnego M_U (ang. *Exact Match Rate*), badającym wierność odwzorowania pełnej sekwencji stanów opisanych wzorami (3.19) i (3.20). To pierwszy oraz starszy wskaźnik, jakim posługiwano się w ocenie modeli, jednak nie oceniający jakości podejmowanych decyzji.

Ranking B oparty jest na wskaźniku S_{FN} (rosnąco) jako głównym kryterium, z rozstrzygającą rolą S (malejąco). Tutaj oceniana jest jakość decyzji pod względem bezpieczeństwa. Model wybrany według Rankingu B jest modelem rekomendowanym do wdrożenia produkcyjnego z racji skupienia na minimalizacji rekomendacji wyłączenia niezbędnej maszyny.

Na początku posługiwano się Rankingiem A jednak w praktyce okazało się, że brak skupienia na koszcie nieprawidłowych decyzji może prowadzić do sytuacji, w której model o wysokim dopasowaniu, będzie generował błędne rekomendacje. Dlatego zdecydowano się na wykorzystanie do wdrożenia Rankingu B.

3.5.4 Walidacja wieloobiektowa

Proponowana metodyka zweryfikowana została na danych rzeczywistych z 12 obiektów przemysłowych reprezentujących zróżnicowane profile produkcji, konfiguracje sprężarek i charakterystyki zapotrzebowania.

Ze względu na poufność danych klientów firmy Marani Sp. z o.o., obiekty oznaczono kodami anonimizowanymi, jak również usunięto dane pozwalające na identyfikację.

Zróżnicowanie zbioru walidacyjnego obejmuje:

- liczbę sprężarek w układzie: od dwóch do pięciu,
- okresowość próbkowania: 20 s (w wyjątkowych przypadkach 5 s jak w tabeli 5.9),
- profile produkcji: ciągła, zmianowa, z przerwami weekendowymi,
- różne gałęzie przemysłu.

Ponieważ zbiór zakładów jest silnie zróżnicowany, to metodyka może zostać oceniona w różnych warunkach przemysłowych. Szczegółowe wyniki dla każdego obiektu przedstawiono w rozdziale 4.

3.6 Środowisko i narzędzia

Architektura oprogramowania:

- **Python 3.x**, **TensorFlow/Keras**,
- **Optuna** (TPE),
- **NumPy, Pandas**,
- **snap7** (PLC, protokół S7),
- **MariaDB**.

System wdrażany został zarówno na komputerach klasy przetwarzania brzegowego (ang. *edge computing*) bezpośrednio na obiekcie, jak i zdalnie z wykorzystaniem chmurowych rozwiązań firmowych co pozwoliło na walidację systemu w różnych warunkach.

Kod źródłowy stanowi własność intelektualną firmy Marani Sp. z o.o. i nie jest publicznie dostępny. Szczegóły implementacyjne opisano w formie pseudokodu w Dodatku A a konfiguracyjne w Dodatku B.

3.7 Podsumowanie

Rozdział opisał zaproponowaną metodykę budowy modelu predykcyjnego odzwierciedlającego działanie sprężarkowni w świetle dynamiki obiektu. Metodyka obejmuje przygotowanie danych, dobór architektury sieci BiLSTM, dwu oraz czteroskładnikową budowę funkcji straty, aż po metodykę oceny i selekcję modelu produkcyjnego. Uczenie sieci prowadzone jest przez składową regresyjną i klasyfikacyjną, natomiast na wybór modelu produkcyjnego przekłada się składowa trendu oraz asymetryczna kara. Wskaźniki oceny skupione na tranzycjach a nie na całym przebiegu pozwalają dobrać bezpieczny model. W rozdziale 4 przedstawiono wyniki eksperymentów na danych rzeczywistych z wspomnianych 12 obiektów przemysłowych.

4 Eksperymenty i wyniki

Rozdział przedstawia wyniki zastosowania proponowanej metodyki dla 12 wybranych obiektów przemysłowych odzwierciedlających reprezentatywne grupy sprężarkowni występujące w polskim przemyśle. Podano hiperparametry użytych sieci neuronowych oraz założenia dla przyjętych eksperymentów. Na wstępie opisuje studium ablacji, które bada jaka architektura sieci neuronowej w zestawieniu z proponowanymi funkcjami straty jest odpowiednia. Kolejno opisane zostają wyniki zgrupowane wg budowy sprężarkowni. Rozdział kończy się badaniem wpływu horyzontu predykcji na proponowaną ocenę ryzyka.

4.1 Charakterystyka obiektów badawczych

Walidację zaproponowanej w rozdziale 3 metodyki przeprowadzono na rzeczywistych danych pochodzących z 12 obiektów przemysłowych należących do różnych gałęzi przemysłu. Dane zostały podzielone na trzy grupy na bazie rodzaju oraz konfiguracji maszyn wchodzących w skład sprężarkowni. Jak zaznaczono wcześniej, w przypadku oszczędności energii elektrycznej, najważniejsze jest posiadanie w układzie przynajmniej jednej maszyny wyposażonej w przetwornicę częstotliwości VFD (ang. *Variable Frequency Drive*). Ponieważ możliwości potencjalnych oszczędności energii w tym wypadku będą znacząco się różnić, zdecydowano o podzieleniu obiektów na grupy, gdzie kryterium podziału stanowi obecność oraz ilość sprężarek VFD. W związku z tym dane podzielono na trzy grupy:

A 4 obiekty bez sprężarek wyposażonych w przetwornicę częstotliwości VFD,
B 4 obiekty z jedną sprężarką z przetwornicą częstotliwości VFD,
C 4 obiekty, gdzie wszystkie sprężarki wyposażone są w falownik VFD.

Tabela 4.1: Charakterystyka obiektów badawczych. N — łączna liczba sprężarek, N_{VFD} — liczba sprężarek z falownikiem, T_s — okres próbkowania.

Obiekt	Typ	N	N_{VFD}	T_s [s]
A1	A	3	0	20
A2	A	3	0	20
A3	A	4	0	20
A4	A	7	0	20
B1	B	3	1	20
B2	B	3	1	20
B3	B	3	1	20
B4	B	5	1	20
C1	C	3	3	20
C2	C	3	3	20
C3	C	3	3	20
C4	C	5	5	20

W przypadku większości obiektów przeważa konfiguracja trzech maszyn, co jest typową konfiguracją w sprężarkowniach, gdzie założeniem może być działanie dwóch ze sprężarek, a trzecia maszyna jest maszyną rezerwową. Jednak, co będzie widoczne w dalszej części rozdziału, zmienność obiektów w zakresie liczby zdarzeń tranzycji, a co za tym idzie ich dynamika, będzie zdecydowanie inna oraz będzie odzwierciedlać pośrednio charakter obiektu. Ostatnie obiekty w każdej grupie zawierają większą liczbę maszyn, co również daje wgląd w efektywność proponowanej metodyki, gdy sprężarkownia jest bardziej rozbudowana. Ze względu na poufność danych klientów firmy Marani Sp. z o.o. obiekty oznaczono indywidualnymi, niepowiązanymi kodami a dane zostały poddane anonimizacji w sposób zachowujący ich charakterystykę.

4.2 Ocena zmienności obiektów

Celem opracowanego mechanizmu jest adekwatne przewidywanie przyszłych stanów obciążeń maszyn na bazie zdarzeń poprzedzających. W celu podkreślenia zróżnicowania danych pomiarowych wynikających ze zmiennego profilu produkcji zdecydowano się na podejście statystyczne. Wyznaczono długość cykli w trakcie trwania których maszyna pozostaje w stanie obciążenia lub odciążenia, mnożąc liczbę następujących po sobie próbek przez okres próbkowania. W tym miejscu wypada podkreślić, że długość cyklu jest zawsze wielkością nieujemną. Zdecydowano się na zgrupowanie wszystkich danych dla pojedynczych maszyn w ramach jednego zbioru obciążenia lub odciążenia,

dzięki czemu można oddać zmienność układu jako całości. Takie podejście będzie jednak związane z tym, że część obserwowanej asymetrii będzie powiązana z odmienną rolą maszyn. Dla każdego ze zbiorów stanów obciążenia oraz odciążenia wyznaczono statystyczny współczynnik zmienności:

$$CV = \frac{\sigma}{\mu}, \tag{4.1}$$

gdzie σ oznacza odchylenie standardowe długości cykli a μ oznacza średnią długość cykli. Zaproponowany współczynnik zmienności nie jest miarą normalności rozkładu. W rozpatrywanym przypadku badane wielkości są nieujemne (opisują czas trwania), przez co wartości CV znacznie przekraczające jedność wykluczają rozkład normalny, wymagałby on bowiem znacznego udziału wartości ujemnych (przy CV zbliżonym do jedności ponad kilkanaście procent rozkładu leżałoby poniżej zera).

Kolejnym współczynnikiem zaczerpniętym ze statystyki jest współczynnik skośności, który za zadanie będzie miał określenie stopnia i kierunku asymetrii rozkładu. By zbadać asymetrię rozkładu długości cykli zdecydowano się na współczynnik skośności wyznaczony jako trzeci moment standaryzowany:

$$\gamma = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \mu}{\sigma}\right)^3, \tag{4.2}$$

gdzie x_i oznacza długość i-tego cyklu, n liczbę analizowanych cykli, μ ich wartość średnią, a σ odchylenie standardowe. Przy rozkładzie symetrycznym przyjmuje wartość zerową, a dla wartości dodatnich oznacza prawostronną skośność (występowanie długiego ogona po prawej stronie).

Uzupełniająco wprowadzono również stosunek wartości średniej do mediany $\mu/\tilde{x}$ będący prostym wskaźnikiem asymetrii jaki pojawia się w literaturze [93]. Dla rozkładu symetrycznego jest bliski jedności, a wraz z rosnącą asymetrią oddala się od niej. W rozpatrywanym przypadku oznacza, jak bardzo wartość średnia odbiega od typowej długości cyklu.

Ostatnia zaproponowana wartość to wskaźnik $\tau_{<h}$, który wyraża procentowo jaki odsetek okresów biegu jałowego jest krótszy od przyjętego horyzontu predykcji $h = 2$ min.

Cykle produkcyjne różnią się znacząco od siebie, co przekłada się na różną długość cykli biegu jałowego, będąc również powiązane z różnymi zdarzeniami wpływającymi na pracę obiektu w danym momencie (serwisy, przeglądy, zakumulowana objętość układu). Ta specyfika pozwala oczekiwać, że rozkłady długości cykli będą odbiegać od normalnego, co potwierdzają wartości przedstawione w dalszej części. W związku z operowaniem na tysiącach cykli, świadomie zrezygnowano z testów Shapiro-Wilka lub Kołmogorowa-Smirnowa, które miałyby formalnie zweryfikować normalność rozkładu.

Dla tak dużych prób występuje zjawisko nadmiernej mocy testów statystycznych [94], zgodnie z którym nawet niewielkie odstępstwa powodują odrzucenie hipotezy o normalności, nie wskazując przy tym charakteru ani skali odstępstwa. Dlatego zastosowano zaproponowane metryki oceny, które wskazują na charakter i kierunek odstępstwa od rozkładu normalnego. Wyniki oceny przedstawiono w tabeli dotyczącej obiektów badawczych 4.2 oraz obiektów wdrożeniowych 4.3.

TABELA 4.2: Charakterystyka zmienności cykli pracy obiektów badawczych. CV — współczynnik zmienności (odchylenie standardowe odniesione do wartości średniej), γ — współczynnik skośności rozkładu długości okresów biegu jałowego, $\mu/\tilde{x}$ — stosunek wartości średniej do mediany, $\tau_{<h}$ — udział okresów biegu jałowego krótszych od horyzontu predykcji.

Obiekt	CV_{obc}	CV_{luz}	γ_{luz}	$\mu/\tilde{x}$	$\tau_{<h}$ [%] ↓
A1	4,27	2,05	21,75	1,85	4
A2	47,00	0,56	1,82	1,02	99
A3	2,36	4,64	9,89	5,11	0
A4	23,99	21,70	45,15	11,15	92
B1	10,77	8,52	25,86	8,52	45
B2	2,06	3,67	6,36	16,85	25
B3	35,05	9,80	19,80	7,55	85
B4	2,42	1,85	3,31	3,13	9
C1	5,47	2,53	4,09	5,51	1
C2	3,10	3,45	6,33	14,63	13
C3	2,12	2,47	8,16	4,36	0
C4	8,90	4,33	11,24	9,34	36

TABELA 4.3: Charakterystyka zmienności cykli pracy obiektów wdrożeniowych. Oznaczenia jak w tabeli 4.2. Dla obiektu D3 zestawiono dodatkowo wartości uzyskane przed zmianą okresu próbkowania (D3, $T_s = 20$ s) oraz po niej (D3', $T_s = 5$ s).

Obiekt	CV_{obc}	CV_{luz}	γ_{luz}	$\mu/\tilde{x}$	$\tau_{<h}$ [%] ↓
D1	2,41	7,84	25,52	4,60	1
D2	0,19	11,52	37,45	2,52	2
D3	3,85	1,17	1,57	2,03	76
D3'	0,98	1,18	1,06	4,25	9
D4	2,70	3,40	6,93	3,68	3

Pierwszym spostrzeżeniem dotyczącym przedstawionych powyżej tabel jest obecność w nich dużych wartości współczynników. Zauważalny jest współczynnik zmienności, który na prawie wszystkich obiektach badanych oraz wdrożeniowych przekracza wartość jeden, przy maksymalnych wartościach dochodzących do kilkudziesięciu. Wszystkie współczynniki skośności przekraczają dodatnią wartość jeden, również dochodząc do poziomu kilkudziesięciu. Może to oznaczać, że obok kilkudziesięciosekundowych zdarzeń znajdują się zdarzenia trwające przez godziny. Współczynnik $\mu/\tilde{x}$ tylko dla obiektu A2 znajduje się w okolicach jedności, dla obiektów A4, B2, C2 przekracza jedenaście, co oznacza, że wartość średnia jest stanem występującym w procesie sporadycznie. Wypada również wspomnieć o obiekcie D2 o zmienności obciążenia na poziomie 0,19, co oznacza regularną charakterystykę, która rokuje duży potencjał oszczędnościowy, co potwierdziło wdrożenie. Uśrednione wskaźniki nie niosą w sobie informacji o typowym stanie pracy i nie nadają się do porównywania właściwości badanych układów. Z takiego powodu właśnie przyjęto ocenę skuteczności bazującą na faktycznych pojedynczych zdarzeniach decyzyjnych, a nie wartościach uśrednionych.

Wskaźnik $\tau_{<h}$ wskazuje na brak potencjału oszczędności przez wskazanie sytuacji, gdy powrót do tłoczenia następuje w okresie przyjętego horyzontu czasowego. Czyli jeżeli taki powrót następuje, to nie można wyłączyć maszyny. Dla obiektów A3 oraz C3 wynosi zero co oznacza, że taki powrót nie następuje, dla obiektów C1 oraz A1 jest poniżej pięciu, z drugiej strony dla obiektów A2, A4 przekracza 90% pozostawiając niewiele miejsca do optymalizacji.

Wyprzedzając wnioski ze studium przypadku obiektu D3 (sekcja 5.9) warto zauważyć, że dane zapisywane co 5 s w późniejszym okresie miały niższe $\tau_{<h}$ dla sprężarki SP1 niż dane zapisywane co 20 s (24% wobec 60%) przy tym samym podziale ról sprężarek (tabela 5.9). Mediana czasu trwania biegu jałowego sprężarki SP1 wzrosła z 80 do 410 s, a liczba okresów biegu jałowego na dobę zmalała z 374 do 226, jednak zjawisko to związane jest z dryftem koncepcyjnym (sekcja 2.7). Z kolei te same dane 5 s przeskalowane do 20 s dają $\tau_{<h}$ sprężarki SP1 na poziomie 18%, zatem sam okres próbkowania nie tłumaczy tej różnicy. Zdarzenia nie były pomijane, jednak moment ich zarejestrowania nie nadążał za zmianami ciśnienia, przez co sieć nie zauważała właściwego momentu decyzji i konieczne było wprowadzenie opóźnienia wywołania predykcji (sekcja 5.9.6). Wysoka wartość $\tau_{<h}$ w danych z wcześniejszego okresu wynikała zatem ze zmiany profilu pracy układu, a nie z okresu próbkowania.

4.3 Założenia eksperymentów

Proponowana metodyka została przetestowana na wszystkich obiektach przy identycznych założeniach eksperymentalnych, co pozwala na porównywalność wyników pomiędzy nimi. Badania zostały przedstawione do momentu siatki kombinacji 3.4.2, optymalizacja bayesowska [86], [95] daje czasem niewielką poprawę wyników, ale nie zmienia znacząco ogólnego trendu, w związku z czym zrezygnowano tutaj z jej stosowania. Najważniejszy jest punkt wyjściowy obserwacji, dlatego zdecydowano, że w tym wypadku będzie wspólny dla wszystkich obiektów. Punktem wyjściowym dla wszystkich eksperymentów są następujące, stałe założenia:

Tabela 4.4: Hiperparametry modelu BiLSTM, jednakowe dla wszystkich obiektów badawczych.

Parametr	Wartość
Architektura sieci	BiLSTM
Liczba warstw rekurencyjnych	2
Jednostki warstwy 1	100
Jednostki warstwy 2	50
Współczynnik porzucania Dropout	0,20
Optymalizator stochastyczny	Adam
Współczynnik uczenia	0,01
Rozmiar batch	4096
Wczesne zatrzymanie (ang. *Early stopping*) — cierpliwość	15 epok
Wczesne zatrzymanie — min delta	0,001
Ziarno generatora liczb losowych	42 (stałe)
Podział danych	75% trening / 25% walidacja
Próbkowanie	$T_s = 20$ s
Horyzonty predykcji (lookback = horyzont)	
2 min	$L = h = 6$ próbek
3 min	$L = h = 9$ próbek
4 min	$L = h = 12$ próbek
5 min	$L = h = 15$ próbek

Kolejny ważnym aspektem proponowanej metodyki jest założenie symetryczności okna lookback i horyzontu predykcji ($L = h$) co oznacza, że model zawsze bierze pod

uwagę tyle samo próbek wstecz, ile przewiduje w przód, co upraszcza interpretację wyników. Identyczna konfiguracja hiperparametrów dla wszystkich 12 obiektów zapewnia porównywalność warunków eksperymentalnych.

4.4 Studium ablacji

W celu weryfikacji wkładu poszczególnych składowych proponowanej metodyki przeprowadzono studium ablacji na reprezentatywnych obiektach z każdej grupy: A3 (typ A), B1 (typ B) oraz C3 (typ C). Obiekty dobrano jako przedstawicieli grup o średniej liczebności tranzycji, przez co wyeliminowano przypadki skrajne, które mogłyby zaburzyć interpretację wyników (obiekty o marginalnej liczbie zdarzeń oraz obiekty z jedną maszyną prowadzącą) [73]. Problem ewaluacji przy asymetrycznym koszcie omówiono w dedykowanej sekcji 2.5. Użyte warianty funkcji straty przedstawiono w tabeli 4.6. Pozostałe parametry sieci i uczenia są wspólne dla wszystkich wariantów i odpowiadają tabeli 4.4, a zmieniana jest wyłącznie funkcja straty a dla wariantów porównawczych architektura. Wszystkie warianty uczono przy stałym ziarnie generatora liczb losowych, dzięki czemu różnice spośród wyników wynikały z funkcji straty oraz architektury, a nie z losowej inicjalizacji wag.

Tabela 4.5: Wyniki studium ablacji dla obiektów A3, B1 i C3, horyzont 2 min. Wskaźniki policzono łącznie dla wszystkich sprężarek obiektu.

Wariant	Architektura	S_{FN} [%]↓	S [%]↑	RMSE↓
Obiekt A3 (typ A — 4 sprężarki bez falownika)				
$\mathcal{L}_{MSE}$	LSTM	35,7	73,2	0,4477
$\mathcal{L}_{MSE} + \mathcal{L}_{BCE}$	LSTM	0,0	6,0	0,5027
$\mathcal{L}_{MSE} + \mathcal{L}_{BCE}$	GRU	0,0	4,0	0,4797
$\mathcal{L}_{MSE} + \mathcal{L}_{BCE}$	BiLSTM	50,0	99,2	0,4992
$\mathcal{L}_{MSE} + \mathcal{L}_{BCE} + \mathcal{L}_{reload}$	BiLSTM	42,9	99,4	0,4890
`industrial`	BiLSTM	35,7	99,4	0,4752
Obiekt B1 (typ B — 3 sprężarki, 1 z falownikiem)				
$\mathcal{L}_{MSE}$	LSTM	0,0	17,6	0,0763
$\mathcal{L}_{MSE} + \mathcal{L}_{BCE}$	LSTM	31,7	51,2	0,0773
$\mathcal{L}_{MSE} + \mathcal{L}_{BCE}$	GRU	31,7	50,0	0,0772
$\mathcal{L}_{MSE} + \mathcal{L}_{BCE}$	BiLSTM	49,1	87,5	0,0498
$\mathcal{L}_{MSE} + \mathcal{L}_{BCE} + \mathcal{L}_{reload}$	BiLSTM	40,7	88,1	0,0498
`industrial`	BiLSTM	94,0	83,0	0,0520

Tabela 4.5 — ciąg dalszy

Wariant	Architektura	S_{FN} [%]↓	S [%]↑	RMSE↓
Obiekt C3 (typ C — 3 sprężarki z falownikami)				
$\mathcal{L}_{MSE}$	LSTM	0,0	41,7	0,0478
$\mathcal{L}_{MSE} + \mathcal{L}_{BCE}$	LSTM	0,0	32,4	0,0472
$\mathcal{L}_{MSE} + \mathcal{L}_{BCE}$	GRU	0,0	33,7	0,0483
$\mathcal{L}_{MSE} + \mathcal{L}_{BCE}$	BiLSTM	0,0	99,8	0,0306
$\mathcal{L}_{MSE} + \mathcal{L}_{BCE} + \mathcal{L}_{reload}$	BiLSTM	0,0	99,4	0,0314
`industrial`	BiLSTM	0,0	99,6	0,0319

Przy przyjętym progu `min_delta` (tabela 4.4) kryterium wczesnego zatrzymania [87] aktywowane jest praktycznie wyłącznie przez składowe progowe. Wynika to z przyjęcia metodyki, w której krok z niewykrytym powrotem obciążenia wnosi do składowej progowej wysoką karę $w_r \cdot \alpha = 9{,}36$ (tabela 4.6), a składowa regresyjna po uwzględnieniu wagi wnosi mniej niż $w_p = 0{,}15$ (ciśnienie jest standaryzowane, przez co błąd średniokwadratowy modelu nadążającego za jego przebiegiem jest mniejszy od jedności). Przy tej samej konfiguracji niewielka zmiana liczby zdarzeń przesuwa moment zatrzymania uczenia, czego skutkiem jest rozbieżność pomiędzy różnymi treningami.

Tabela 4.6: Konfiguracja wariantów funkcji straty w studium ablacji. Wagi składowych nie są normalizowane. We wszystkich wariantach ze składową $\mathcal{L}_{reload}$ mnożnik kary za FP wynosi $\beta = 1$.

Wariant	Arch.	w_p	w_b	w_r	w_t	α
$\mathcal{L}_{MSE}$	LSTM	1,0*	–	–	–	–
$\mathcal{L}_{MSE} + \mathcal{L}_{BCE}$	LSTM, GRU, BiLSTM	0,30	0,70	–	–	–
$\mathcal{L}_{MSE} + \mathcal{L}_{BCE} + \mathcal{L}_{reload}$	BiLSTM	0,15	0,10	0,78	0	12
`industrial`	BiLSTM	0,15	0,10	0,78	0,15	12

* pierwiastek składowej regresyjnej (`loss_type: rmse`)

Przedstawione wyniki studium ablacji prowadzą do czterech obserwacji (tabela 4.5). Pierwszą obserwacją jest krytyczna rola dwukierunkowości przetwarzania [54]. W przypadku architektur jednokierunkowych z funkcją $\mathcal{L}_{MSE} + \mathcal{L}_{BCE}$ następuje degeneracja prowadząca do uproszczonej strategii działania, co dla obiektu A3 (tabela 4.7) oznacza, że wariant LSTM [47], [96] generuje 2053 zdarzenia FP na 2185 okien decyzyjnych ($S = 6{,}0\%$), a GRU [50] odpowiednio 2097 ($S = 4{,}0\%$). W przypadku C3 występuje analogiczne zachowanie, model przewiduje jedną klasę prawie bezwarunkowo [73], przez co prawie zawsze sygnalizuje powrót obciążenia i nie generuje ryzykownych rekomendacji wyłączeń ($S_{FN} = 0{,}0\%$). Przy tylko dwóch zdarzeniach powrotu obciążenia

oznacza to blokadę wszystkich możliwości oszczędności. Doskonały wskaźnik S_{FN} jest artefaktem degeneracji, a nie dowodem jakości modelu. Takie zachowanie uzasadnia konieczność łącznej analizy kompletu wskaźników wraz z surowymi liczbami zdarzeń (jak również zalecane jest zapoznanie się z wykresami poszczególnych tranzycji decyzyjnych przed przeprowadzeniem wdrożenia). BiLSTM przy identycznej funkcji straty i identycznych hiperparametrach osiąga S na poziomie 87,5–99,8%, co wskazuje, że to kierunkowość przetwarzania, a nie mechanizm bramek, decyduje o zdolności identyfikacji charakterystyki pracy układu. Zachowanie architektur LSTM i GRU pozostaje zgodne z obserwacjami z literatury, gdzie GRU jest wskazana jako uproszczona architektura LSTM osiągająca dla szeregów czasowych wyniki porównywalne [97].

Tabela 4.7: Liczby zdarzeń w studium ablacji dla obiektu A3, horyzont 2 min, suma dla wszystkich sprężarek. N_T — liczba tranzycji.

Wariant	Arch.	N_T	TP	TN	FP	FN	S_{FN} [%]↓	S [%]↑
$\mathcal{L}_{MSE}$	LSTM	2185	9	1591	580	5	35,7	73,2
$\mathcal{L}_{MSE} + \mathcal{L}_{BCE}$	LSTM	2185	14	118	2053	0	0,0	6,0
$\mathcal{L}_{MSE} + \mathcal{L}_{BCE}$	BiLSTM	2185	7	2161	10	7	50,0	99,2
$\mathcal{L}_{MSE} + \mathcal{L}_{BCE}$	GRU	2185	14	74	2097	0	0,0	4,0
$\mathcal{L}_{MSE} + \mathcal{L}_{BCE} + \mathcal{L}_{reload}$	BiLSTM	2185	8	2163	8	6	42,9	99,4
`industrial`	BiLSTM	2185	9	2162	9	5	35,7	99,4

Drugą obserwacją jest powiązanie pomiędzy składową klasyfikacyjną $\mathcal{L}_{BCE}$ a składową regresyjną $\mathcal{L}_{MSE}$. W literaturze często można spotkać się z wykorzystaniem samej składowej regresyjnej [24], [64], która w tym wypadku jest niewystarczająca, dla B1 wariant samej składowej $\mathcal{L}_{MSE}$ osiąga $S = 17{,}6\%$ przy 2897 zdarzeniach FP (tabela 4.8) co wynika z braku uwzględnienia stanów sprężarek o których model nie wie. Może się również zdarzyć, że dodanie składowej BCE do architektury jednokierunkowej nie tylko nie poprawia, ale wręcz pogarsza wynik (A3: spadek S z 73,2% do 6,0%). Składowa klasyfikacyjna $\mathcal{L}_{BCE}$ jest efektywna tylko w połączeniu z architekturą dwukierunkową wzajemnie się uzupełniając.

Tabela 4.8: Liczby zdarzeń w studium ablacji dla obiektu B1, horyzont 2 min, suma dla wszystkich sprężarek. N_T — liczba tranzycji.

Wariant	Arch.	N_T	TP	TN	FP	FN	S_{FN} [%]↓	S [%]↑
$\mathcal{L}_{MSE}$	LSTM	3514	615	2	2897	0	0,0	17,6
$\mathcal{L}_{MSE} + \mathcal{L}_{BCE}$	LSTM	3514	420	1380	1519	195	31,7	51,2
$\mathcal{L}_{MSE} + \mathcal{L}_{BCE}$	BiLSTM	3514	313	2762	137	302	49,1	87,5
$\mathcal{L}_{MSE} + \mathcal{L}_{BCE}$	GRU	3514	420	1337	1562	195	31,7	50,0
$\mathcal{L}_{MSE} + \mathcal{L}_{BCE} + \mathcal{L}_{reload}$	BiLSTM	3514	365	2730	169	250	40,7	88,1
`industrial`	BiLSTM	3514	37	2880	19	578	94,0	83,0

Trzecia obserwacja dotyczy asymetrycznej kary [72], [78] za najbardziej niebezpieczne z punktu widzenia stałości produkcji błędy FN. Wariant pośredni ($\mathcal{L}_{MSE} + \mathcal{L}_{BCE} + \mathcal{L}_{reload}$) przynosi dla A3 i B1 umiarkowaną poprawę S_{FN} względem bazowego MSE+BCE+BiLSTM, dla A3 z 50,0% do 42,9%, dla B1 z 49,1% do 40,7%, przy zachowaniu, a nawet nieznacznej poprawie wskaźnika S. Sama asymetryczna kara stanowi zatem bezpieczne rozszerzenie wariantu bazowego.

Czwarta obserwacja dotyczy pełnego wariantu przemysłowej funkcji straty `industrial` i jest niejednoznaczna. Dla obiektu A3 przynosi najlepszy wynik bezpieczeństwa ($S_{FN} = 35{,}7\%$ przy $S = 99{,}4\%$ oraz oferuje najniższy wśród wariantów BiLSTM błąd RMSE), dla C3 jest neutralny (tabela 4.9), natomiast dla B1 prowadzi do załamania zdolności wykrywania powrotów obciążenia (liczba zdarzeń TP spada z 365 do 37, a FN rośnie z 250 do 578 przy $S_{FN} = 94{,}0\%$). Następuje degeneracja modelu w kierunku przeciwnym niż architektur jednokierunkowych. Pojawia się bezwarunkowe przewidywanie pozostawienia maszyn na biegu jałowym, przez co jego rekomendacje wyłączeń przestają być wiarygodną podstawą decyzji mimo pozornie wysokiego $S = 83{,}0\%$. Należy zauważyć, że wariant pośredni ze składową $\mathcal{L}_{reload}$, lecz bez składowej trendu, działał na B1 poprawnie, a oba warianty różnią się wyłącznie wagą składowej trendu ciśnienia ($w_t = 0{,}0$ wobec $w_t = 0{,}15$). Przyczyną załamania nie jest zatem asymetryczna kara, lecz właśnie ta składowa. Obie oddziałują przez kryterium wyboru modelu, a nie przez krok gradientowy (sekcja 3.2), włączenie składowej trendu zmienia kryterium wczesnego zatrzymania w sposób, który dla obiektu wyposażonego w falownik prowadzi do wyboru modelu bezwarunkowo przewidującego brak powrotu obciążenia. Wariant pełny wymaga zatem indywidualnej weryfikacji dla każdego obiektu i dostrojenia wag, co szerzej omówiono w rozdziale poświęconym wdrożeniu.

Tabela 4.9: Liczby zdarzeń w studium ablacji dla obiektu C3, horyzont 2 min, suma dla wszystkich sprężarek. N_T — liczba tranzycji.

Wariant	Arch.	N_T	TP	TN	FP	FN	S_{FN} [%]↓	S [%]↑
$\mathcal{L}_{MSE}$	LSTM	1146	2	476	668	0	0,0	41,7
$\mathcal{L}_{MSE} + \mathcal{L}_{BCE}$	LSTM	1146	2	369	775	0	0,0	32,4
$\mathcal{L}_{MSE} + \mathcal{L}_{BCE}$	BiLSTM	1146	2	1142	2	0	0,0	99,8
$\mathcal{L}_{MSE} + \mathcal{L}_{BCE}$	GRU	1146	2	384	760	0	0,0	33,7
$\mathcal{L}_{MSE} + \mathcal{L}_{BCE} + \mathcal{L}_{reload}$	BiLSTM	1146	2	1137	7	0	0,0	99,4
`industrial`	BiLSTM	1146	2	1139	5	0	0,0	99,6

Różnice błędów RMSE pomiędzy wariantami BiLSTM są znikome (rzędu 0,001–0,025). Dobór funkcji straty wpływa przede wszystkim na jakość klasyfikacji stanów, praktycznie nie pogarszając predykcji ciśnienia (od strony wdrożeniowej, nadążanie za charakterystyką ciśnienia jest również mile widziane). Zastrzec należy, że studium przeprowadzono na trzech obiektach reprezentatywnych. Wnioski dotyczące kierunku zmian dla wariantu `mse_bce` z architekturą BiLSTM zweryfikowano na wszystkich dwunastu obiektach, natomiast konkretne wartości wskaźników pozostają specyficzne dla każdego obiektu indywidualnie. Możemy tutaj zauważyć tendencję, że charakterystyki pomiędzy obiektami potrafią bardzo się różnić i wymagać dedykowanego, zindywidualizowanego podejścia.

4.5 Wyniki dla grupy A — obiekty bez maszyn z falownikiem

Grupę A stanowią cztery obiekty wyposażone wyłącznie w sprężarki o stałej prędkości obrotowej. Stan na rok 2026, jest taki, że można by powiedzieć, że są to obiekty starego typu, które z różnych przyczyn nie są wyposażone w przynajmniej jedną maszynę VFD. W tego typu układach sterowanie nadrzędne realizowane jest przez sekwencyjne dobieranie liczby pracujących maszyn. Sprężarki pracują w trybie dwustanowym czyli albo z pełną wydajnością albo przechodzą na bieg jałowy (na którym zużywają ok. 30% elektrycznej energii nominalnej). Oszczędność energii wymaga fizycznego odciążenia oraz zatrzymania maszyny tak, aby uniknąć utrzymywania jej przez długi czas na biegu jałowym. Predykcja stanów binarnych ma tutaj zatem bezpośrednie przełożenie na redukcję zużycia energii. Charakterystyczną cechą układów tej grupy jest obecność maszyn wiodących, pracujących niemal ciągle. Może to np. wynikać z różnej mocy rozpatrywanych maszyn. W wynikach jest to zauważalne, przez małą liczbę okien decyzyjnych, w których po odciążeniu następuje powrót obciążenia. Przy nielicznej ilości zdarzeń model nie znajduje w historii wzorca pozwalającego na ich przewidzenie,

co przekłada się na wysokie wartości S_{FN} dla tych sprężarek. Dlatego rekomendacja wyłączenia maszyny wiodącej nie powinna stanowić samodzielnej podstawy decyzji dla sterownika nadrzędnego i powinna zostać dodatkowo zabezpieczona.

4.5.1 Obiekt A1

Wyniki dla obiektu A1 przedstawiono w tabelach 4.10 (zdarzenia w oknach decyzyjnych), 4.11 (wskaźniki bezpieczeństwa) oraz 4.12 (wpływ horyzontu predykcji).

Tabela 4.10: Zdarzenia w oknach decyzyjnych dla obiektu A1, horyzont 2 min, `mse_bce`. N_T — liczba tranzycji.

Sprężarka	N_T	TP	TN	FP	FN
SP1	361	21	297	12	31
SP2	1429	0	1418	1	10
SP3	3584	3	3556	17	8
Suma	5374	24	5271	30	49

Tabela 4.11: Wskaźniki bezpieczeństwa dla obiektu A1, horyzont 2 min, `mse_bce`. Symbol „—” oznacza brak zdarzeń, dla których można wyznaczyć wskaźnik.

Sprężarka	N_T	S_{FN} [%]↓	S_{FP} [%]↓	S [%]↑
SP1	361	59,6	3,3	88,1
SP2	1429	100,0	0,1	99,2
SP3	3584	72,7	0,5	99,3
Suma	5374	67,1	0,6	98,5
			RMSE =	0,0522

Tabela 4.12: Wyniki dla obiektu A1 dla różnych horyzontów predykcji, `mse_bce`.

Horyzont	S_{FN} [%]↓	S_{FP} [%]↓	S [%]↑	RMSE↓
2 min	67,1	0,6	98,5	0,0522
3 min	47,9	0,2	99,2	0,0669
4 min	39,7	0,1	99,3	0,0835
5 min	39,7	0,1	99,3	0,0950

Obiekt A1 to układ trzech sprężarek o stałej prędkości obrotowej. Rozkład tranzycji jest wyraźnie asymetryczny oraz odzwierciedla sekwencyjną pracę systemu, gdzie SP3

dominuje (3 584 zdarzeń, maszyna dobijająca, być może przy projekcie rezerwowa), SP2 ma 1 429 zdarzeń (podstawa zapotrzebowania), natomiast SP1 jedynie 361 (maszyna wiodąca, załączana w pierwszej kolejności). Obiekt cechuje bardzo mała liczba okien decyzyjnych, w których po przejściu na bieg jałowy następuje powrót do obciążenia, występują łącznie 73 zdarzenia klasy $TP+FN$ na 5 374 tranzycje. Dla SP2 jest to tylko 10 zdarzeń i model nie wykrył żadnego z nich ($S_{FN} = 100\%$). Dla SP3 to 11 zdarzeń przy $S_{FN} = 72{,}7\%$. Najlepszy indywidualny wynik uzyskuje SP1 ($S_{FN} = 59{,}6\%$ przy 52 zdarzeniach), będąca zarazem jedyną maszyną o statystycznie istotnej liczbie zdarzeń decyzyjnych. Zatem, żeby sieć mogła się nauczyć wzorca powrotu obciążenia, potrzebna jest odpowiednia liczba zdarzeń w historii.

Wysokie wartości S_{FN} oznaczają, że rekomendacje wyłączenia generowane przez model nie mogą stanowić samodzielnej podstawy decyzji (w części przypadków sprężarka okazuje się ponownie potrzebna). Bardzo niski S_{FP} (0,6% przy 2 min, nie więcej niż 0,2% dla dłuższych horyzontów) świadczy o pełnym wykorzystaniu dostępnych możliwości oszczędności (model prawie nigdy nie blokuje wyłączenia, jeśli byłoby możliwe). Z kolei rozpatrywanie długości horyzontu pokazuje systematyczną poprawę bezpieczeństwa wraz z jego wydłużaniem tzn. S_{FN} maleje z 67,1% (2 min) do 39,7% (4 min) i stabilizuje się na tym poziomie. Dla obiektu A1 można zatem zalecić horyzont 4–5 min, przy którym ryzyko przedwczesnego wyłączenia jest najniższe, co daje potencjał oszczędności energii. Nasuwa się wniosek, że horyzont czasowy jest kwestią indywidualną dla każdego obiektu i tak powinien być rozpatrywany.

4.5.2 Obiekt A2

Wyniki dla obiektu A2 przedstawiono w tabelach 4.13 (zdarzenia w oknach decyzyjnych), 4.14 (wskaźniki bezpieczeństwa) oraz 4.15 (wpływ horyzontu predykcji).

Tabela 4.13: Zdarzenia w oknach decyzyjnych dla obiektu A2, horyzont 2 min, `mse_bce`. N_T — liczba tranzycji.

Sprężarka	N_T	TP	TN	FP	FN
SP1	0	0	0	0	0
SP2	17677	17324	1	309	43
SP3	26	26	0	0	0
Suma	17703	17350	1	309	43

Tabela 4.14: Wskaźniki bezpieczeństwa dla obiektu A2, horyzont 2 min, `mse_bce`. Symbol „—” oznacza brak zdarzeń, dla których można wyznaczyć wskaźnik.

Sprężarka	N_T	S_{FN} [%]↓	S_{FP} [%]↓	S [%]↑
SP1	0	—	—	—
SP2	17677	0,2	1,7	98,0
SP3	26	0,0	0,0	100,0
Suma	17703	0,2	1,7	98,0
			RMSE =	0,1833

Tabela 4.15: Wyniki dla obiektu A2 dla różnych horyzontów predykcji, `mse_bce`.

Horyzont	S_{FN} [%]↓	S_{FP} [%]↓	S [%]↑	RMSE↓
2 min	0,2	1,7	98,0	0,1833
3 min	0,0	1,8	98,2	0,1893
4 min	11,7	1,1	87,5	0,1942
5 min	1,8	1,4	96,8	0,1889

Obiekt A2 to układ gdzie jedna maszyna (SP2) odpowiada za prawie wszystkie tranzycje (17 677 z 17 703), podczas gdy z SP1 nie są związane żadne zdarzenia, a SP3 ma ich jedynie 26. Liczba ta oddaje charakter pracy układu, tak duża liczba tranzycji wskazuje na bardzo krótkie, częste cykle pracy sprężarki rezerwowej/dobijającej. Struktura zdarzeń jest tu odwrotna niż w pozostałych obiektach grupy: w 17 367 z 17 677 okien decyzyjnych (98,2%) po odciążeniu następuje powrót obciążenia, a model rozpoznaje go w 17 324 przypadkach. Dzięki temu osiągnięty zostaje znakomity wskaźnik $S_{FN} = 0{,}2\%$, oznaczający zaledwie 43 niewykryte powroty.

Struktura liczby obserwowanych zdarzeń będzie przekładać się na możliwości osiągnięcia korzyści energetycznych. Okien, w których wyłączenie było realnie możliwe (klasa $TN + FP$), jest tylko 310. Model zdecydował się wykorzystać tylko jedno z nich, a pozostałe 309 zaklasyfikował jako decyzje ostrożne, gdzie lepiej pozostawić maszynę włączoną (FP). Wskaźnik $S_{FP} = 1{,}7\%$, obliczany względem całego zbioru okien, ukrywa ten fakt. W takiej sytuacji wskaźnik względny wymaga zestawienia z liczbami bezwzględnymi oraz rozpatrywania całościowego. Zauważalna jak specyfika dynamiki układu. Przy horyzoncie 2 min sprężarka SP2 prawie zawsze wraca do obciążenia, a przestrzeń do oszczędności jest niewielka. Nie jest to ograniczenie modelu, lecz właściwość obiektu oraz wspomniany już indywidualny charakter.

Przy rozpatrywaniu horyzontów czasowych, warto zauważyć pojawienie się anomalii przy horyzoncie 4 min (S_{FN} = 11,7%, która daje wyraźnie gorszy wynik niż dla 3 min i 5 min). Nieregularność ta może wynikać z powiązania długości horyzontu z okresowością cykli pracy SP2. W takim wypadku wymagana jest ostrożność przy doborze horyzontu. Stosunkowo wysokie wartości RMSE w grupie A ($\approx$ 0,18) odzwierciedlają dynamiczny profil ciśnienia, który wynika z częstych przełączeń. Przy takim dynamicznym charakterze, biorąc pod uwagę bezpieczeństwo, obiekt A2 kwalifikuje się do wdrożenia przy horyzoncie 2–3 min, przy czym oczekiwane korzyści energetyczne w tym oknie czasowym będą niewielkie.

4.5.3 Obiekt A3

Wyniki dla obiektu A3 przedstawiono w tabelach 4.16 (zdarzenia w oknach decyzyjnych), 4.17 (wskaźniki bezpieczeństwa) oraz 4.18 (wpływ horyzontu predykcji).

Tabela 4.16: Zdarzenia w oknach decyzyjnych dla obiektu A3, horyzont 2 min, `mse_bce`. N_T — liczba tranzycji.

Sprężarka	N_T	TP	TN	FP	FN
SP1	117	0	117	0	0
SP2	52	0	52	0	0
SP3	1625	7	1605	9	4
SP4	391	1	386	2	2
Suma	2185	8	2160	11	6

Tabela 4.17: Wskaźniki bezpieczeństwa dla obiektu A3, horyzont 2 min, `mse_bce`. Symbol „—” oznacza brak zdarzeń, dla których można wyznaczyć wskaźnik.

Sprężarka	N_T	S_{FN} [%]↓	S_{FP} [%]↓	S [%]↑
SP1	117	—	0,0	100,0
SP2	52	—	0,0	100,0
SP3	1625	36,4	0,6	99,2
SP4	391	66,7	0,5	99,0
Suma	2185	42,9	0,5	99,2
			RMSE =	0,4900

Tabela 4.18: Wyniki dla obiektu A3 dla różnych horyzontów predykcji, `mse_bce`.

Horyzont	S_{FN} [%]↓	S_{FP} [%]↓	S [%]↑	RMSE↓
2 min	42,9	0,5	99,2	0,4900
3 min	42,9	0,6	99,1	0,5026
4 min	42,9	0,6	99,1	0,5123
5 min	42,9	0,6	99,1	0,5072

Obiekt A3, składający się z czterech sprężarek, to kolejny obiekt o ciekawych, indywidualnych własnościach co potwierdza wynik badania. Jego indywidualizm objawia się wynikiem S_{FN}, który wynosi identycznie 42,9% dla każdego z czterech horyzontów. Wartość ta jest wynikiem bardzo małej liczby zdarzeń decyzyjnych: przy 2 min klasę $TP + FN$ tworzy tylko 14 zdarzeń (8 TP, 6 FN) na 2 185 tranzycji, czyli wskaźnik odpowiada sześciu niewykrytym powrotom obciążenia. Dynamika obiektu, której cechą charakterystyczną jest niezależność od horyzontu, wskazuje na charakter zdarzeń, który jest niemożliwy do przewidzenia w żadnym z badanych okien czasowych.

Kolejną cechą indywidualną obiektu, jest to, że sprężarki SP1 i SP2 nie generują żadnych zdarzeń klasy TP ani FN. Po odciążeniu nie wracają do stanu obciążenia w oknie decyzyjnym, przez co nie można dla nich wyznaczyć wskaźnika S_{FN}. W związku z takim charakterem wszystkie ich okna to trafne rozpoznania braku powrotu (TN), czyli zrealizowane możliwości wyłączenia. Zdawać by się mogło, że sprężarki, które pozwalają na wsparcie decyzji to sprężarki SP3 ($S_{FN} = 36{,}4\%$ przy 11 zdarzeniach) oraz SP4 ($S_{FN} = 66{,}7\%$ jednak przy zaledwie 3 zdarzeniach, wynik jest statystycznie mało wiarygodny).

Ostatnią zauważalnie indywidualną cechą obiektu jest najwyższe w całym zbiorze RMSE ciśnienia ($\approx 0{,}49$). Tak wysoki poziom wskazuje złożoną, dynamikę układu, która przy czterech sprężarkach jest trudna do modelowania regresyjnego. Tutaj wkraczają wskaźniki bezpieczeństwa, które mimo tego, że $S_{FP} \leq 0{,}6\%$ przy wszystkich horyzontach pokazują, że model niemal nie blokuje możliwych wyłączeń. W tym wypadku po prostu potencjał oszczędnościowy jest wykorzystany w pełni, a ryzyko koncentruje się w nielicznych zdarzeniach sprężarek SP3 oraz SP4.

4.5.4 Obiekt A4

Wyniki dla obiektu A4 przedstawiono w tabelach 4.19 (zdarzenia w oknach decyzyjnych), 4.20 (wskaźniki bezpieczeństwa) oraz 4.21 (wpływ horyzontu predykcji).

Tabela 4.19: Zdarzenia w oknach decyzyjnych dla obiektu A4, horyzont 2 min, `mse_bce`. N_T — liczba tranzycji.

Sprężarka	N_T	TP	TN	FP	FN
SP1	1	0	1	0	0
SP2	1	1	0	0	0
SP3	208	13	165	11	19
SP4	4	2	0	2	0
SP5	13376	12539	178	566	93
SP6	101	79	0	22	0
SP7	267	0	267	0	0
Suma	13958	12634	611	601	112

Tabela 4.20: Wskaźniki bezpieczeństwa dla obiektu A4, horyzont 2 min, `mse_bce`. Symbol „—” oznacza brak zdarzeń, dla których można wyznaczyć wskaźnik.

Sprężarka	N_T	S_{FN} [%]↓	S_{FP} [%]↓	S [%]↑
SP1	1	—	0,0	100,0
SP2	1	0,0	0,0	100,0
SP3	208	59,4	5,3	85,6
SP4	4	0,0	50,0	50,0
SP5	13376	0,7	4,2	95,1
SP6	101	0,0	21,8	78,2
SP7	267	—	0,0	100,0
Suma	13958	0,9	4,3	94,9
			RMSE =	0,0765

Tabela 4.21: Wyniki dla obiektu A4 dla różnych horyzontów predykcji, `mse_bce`.

Horyzont	S_{FN} [%]↓	S_{FP} [%]↓	S [%]↑	RMSE↓
2 min	0,9	4,3	94,9	0,0765
3 min	1,1	4,4	94,6	0,0828
4 min	1,7	2,6	95,8	0,0887
5 min	1,9	3,3	94,9	0,0940

Obiekt A4 to największy zespół sprężarek w badanym zbiorze, który składa się aż z siedmiu sprężarek. Drugą cechą charakterystyczną jest wyraźnie widoczna sprężarka

dobijająca SP5 (co jest charakterystyczne dla układów bez VFD), która ma ogromną liczbą tranzycji (13 376 z 13 958). Powrót do obciążenia w jej wypadku nastąpił w 12 632 oknach decyzyjnych, z czego model wykrył 12 539 z nich. W związku z tym metryki bezpieczeństwa dają jedne z najlepszych wyników w całym zbiorze badawczym: $S_{FN} = 0{,}7\%$ i zagregowany wynik 0,9%. W związku z tym poziom ryzyka przedwczesnego wyłączenia jest dla tego obiektu bardzo niski.

W przypadku pozostałych sprężarek zdarzenia są jednak zróżnicowane. Ponieważ obiekt charakteryzuje duża asymetria pomiędzy liczbą zdarzeń decyzyjnych poszczególnych sprężarek, metryki bezpieczeństwa dla nich wykazują ciekawe własności statystyczne. Sprężarka SP3 ma 32 zdarzenia decyzyjne, w czym 19 niewykrytych powrotów obciążenia, co daje wskaźnik $S_{FN} = 59{,}4\%$. W skali obiektu wartość ta jest marginalna, lecz dla samej maszyny istotna. Dla sprężarki SP6 model zachował konserwatywne podejście nie decydując się na wyłączenie maszyny, ignorując 22 okna w których wyłączenie było możliwe. Wskaźniki bezpieczeństwa w tym wypadku wskazują na bezpieczne zachowanie ale z pominięciem jakichkolwiek korzyści energetycznych ($S_{FP} = 21{,}8\%$ przy $S_{FN} = 0\%$). Kolejna sprężarka, SP4 ma tylko 4 tranzycje. Tak mała ich liczba daje wyniki statystycznie nieistotne ($S_{FP} = 50\%$ odpowiada dwóm zdarzeniom), model tutaj też statystycznie wykazał konserwatywne podejście (brak odpowiednich zdarzeń w zbiorze uczącym). W przypadku sprężarek SP1 i SP7 nie pojawiły się zdarzenia powrotu do obciążenia, a dla SP2 pojawiło się pojedyncze zdarzenie, co nie pozwala na miarodajną ocenę ryzyka. Dla sprężarki SP7 model zdecydował, że w przypadku odciążenia, maszyna nie będzie potrzebna i dla każdego z 267 okien decyzyjnych rozpoznanie braku powrotu do stanu obciążenia było trafne.

Dynamika obiektu A4 doskonale pokazuje, że metryki bezpieczeństwa należy rozpatrywać indywidualnie. Doskonały wynik sumaryczny, jest zdominowany przez maszynę dobijającą, podczas gdy wyniki indywidualne poszczególnych sprężarek są inne. W publikacji [15] zauważono, że w przypadku wdrożenia wsparcie modelu może ograniczać się do wybranych maszyn, gdy dla innych może być nieprzewidywalne. W przypadku wdrożenia zatem sterowanie predykcyjne dla A4 powinno obejmować przede wszystkim SP5, która ma największą liczbę zdarzeń i oferuje największe potencjalne oszczędności energetyczne. Co ciekawe obiekt kwalifikuje się do wdrożenia dla każdego horyzontu.

4.6 Wyniki dla grupy B — obiekty z jedną maszyną falownikową

Obiekty grupy B stanowią popularną, ekonomiczną energetycznie grupę w przemyśle, w której jedna sprężarka z każdej grupy jest wyposażona w przetwornicę częstotliwości VFD. Maszyna ta powinna pełnić rolę wiodącą, regulując ciśnienie w sposób płynny,

dobierając obroty silnika na bazie własnego dokładnego pomiaru ciśnienia (w oparciu o najczęściej regulator PI lub PID). Przy prawidłowo działającym układzie, maszyna falownikowa rzadko generuje zdarzenia odciążenia (co potwierdza fakt, że w obiektach B2 i B4 takie nie wystąpiły). W tak zoptymalizowanym układzie, głównym celem predykcji są sprężarki stałoobrotowe. Ich liczba determinowana jest sekwencyjnie na bazie aktualnego zapotrzebowania. Obecność falownika komplikuje jednak zadanie, jego reakcja kompensująca zmiany zapotrzebowania osłabia związek między przebiegiem ciśnienia a nadchodzącymi przełączeniami maszyn stałoobrotowych. Ta własność jednocześnie powoduje uzyskanie oszczędności energetycznych przez minimalizację liczby załączeń pozostałych maszyn. W związku z tym wewnątrz tej grupy występuje większe zróżnicowanie wyników.

4.6.1 Obiekt B1

Wyniki dla obiektu B1 przedstawiono w tabelach 4.22 (zdarzenia w oknach decyzyjnych), 4.23 (wskaźniki bezpieczeństwa) oraz 4.24 (wpływ horyzontu predykcji).

Tabela 4.22: Zdarzenia w oknach decyzyjnych dla obiektu B1, horyzont 2 min, `mse_bce`. N_T — liczba tranzycji.

Sprężarka	N_T	TP	TN	FP	FN
SP1	1726	39	1545	12	130
SP2	916	209	538	132	37
SP3	872	111	632	40	89
Suma	3514	359	2715	184	256

Tabela 4.23: Wskaźniki bezpieczeństwa dla obiektu B1, horyzont 2 min, `mse_bce`. Symbol „—” oznacza brak zdarzeń, dla których można wyznaczyć wskaźnik.

Sprężarka	N_T	S_{FN} [%]↓	S_{FP} [%]↓	S [%]↑
SP1	1726	76,9	0,7	91,8
SP2	916	15,0	14,4	81,6
SP3	872	44,5	4,6	85,2
Suma	3514	41,6	5,2	87,5
			RMSE =	0,0502

Tabela 4.24: Wyniki dla obiektu B1 dla różnych horyzontów predykcji, `mse_bce`.

Horyzont	S_{FN} [%]↓	S_{FP} [%]↓	S [%]↑	RMSE↓
2 min	41,6	5,2	87,5	0,0502
3 min	37,7	3,0	90,4	0,0576
4 min	32,4	4,5	89,8	0,0641
5 min	27,2	7,0	88,2	0,0683

W przypadku obiektu B1, dla każdej sprężarki występuje duża liczba zdarzeń decyzyjnych. Powrót do stanu obciążenia następuje w 615 z 3514 okien, co pozwala na wiarygodną ocenę wszystkich trzech maszyn. Podobnie rozkład tranzycji jest przy tym względnie równomierny (1726/916/872). Użycie tutaj maszyny falownikowej przy dużej dynamice układu nie przełożyło się na znaczne ograniczenie ilości tranzycji przejścia. Najtrudniejsza w modelowaniu okazała się sprężarka SP1, dla której przy 169 zdarzeniach decyzyjnych model wykrywa jedynie 39 powrotów obciążenia ($S_{FN} = 76{,}9\%$). Tutaj należy się słowo wyjaśnienia, ponieważ w związku z niewielkim okresowym zapotrzebowaniem maszyna pracuje jako jedyna i jednocześnie wiodąca będąc wyposażoną w przetwornicę częstotliwości. Jej zachowanie odpowiada profilowi maszyny pracującej w reżimie ciągłej regulacji. Cykle wyraźnie powtarzalne, zostały zidentyfikowane w przypadku sprężarki SP2 i jest sprężarką najlepiej modelowaną ($S_{FN} = 15{,}0\%$ przy 246 zdarzeniach). Jednak należy zwrócić uwagę, że występuje tu podwyższony $S_{FP} = 14{,}4\%$, czyli 132 okna, w których model niepotrzebnie zablokował możliwe wyłączenie. SP3 z kolei wypada pośrodku ($S_{FN} = 44{,}5\%$).

Dla różnych horyzontów predykcji, sumaryczny wskaźnik S_{FN} systematycznie maleje wraz ze wzrostem horyzontu. Z początkowe 41,6% (2 min) obniża się do 27,2% (5 min) dzięki czemu nasuwa się wniosek, że w indywidualnym przypadku obiektu B1, wydłużenie horyzontu predykcji obniża ryzyko przedwczesnego wyłączenia. W tym wypadku model jednak zaczyna być bardziej zachowawczy czego kosztem jest wzrost S_{FP} z 5,2% do 7,0% i możliwa utrata części możliwych oszczędności. Wybór horyzontu stanowi tu zatem czytelny kompromis między bezpieczeństwem a oszczędnością energii. W przypadku tego obiektu horyzont 4 min ($S_{FN} = 32{,}4\%$, $S_{FP} = 4{,}5\%$) wydaje się rozsądnym punktem równowagi.

4.6.2 Obiekt B2

Wyniki dla obiektu B2 przedstawiono w tabelach 4.25 (zdarzenia w oknach decyzyjnych), 4.26 (wskaźniki bezpieczeństwa) oraz 4.27 (wpływ horyzontu predykcji).

Tabela 4.25: Zdarzenia w oknach decyzyjnych dla obiektu B2, horyzont 2 min, `mse_bce`. N_T — liczba tranzycji.

Sprężarka	N_T	TP	TN	FP	FN
SP1	0	0	0	0	0
SP2	85	9	57	10	9
SP3	367	9	261	2	95
Suma	452	18	318	12	104

Tabela 4.26: Wskaźniki bezpieczeństwa dla obiektu B2, horyzont 2 min, `mse_bce`. Symbol „—” oznacza brak zdarzeń, dla których można wyznaczyć wskaźnik.

Sprężarka	N_T	S_{FN} [%]↓	S_{FP} [%]↓	S [%]↑
SP1	0	—	—	—
SP2	85	50,0	11,8	77,6
SP3	367	91,3	0,5	73,6
Suma	452	85,2	2,7	74,3
			RMSE =	0,0558

Tabela 4.27: Wyniki dla obiektu B2 dla różnych horyzontów predykcji, `mse_bce`.

Horyzont	S_{FN} [%]↓	S_{FP} [%]↓	S [%]↑	RMSE↓
2 min	85,2	2,7	74,3	0,0558
3 min	82,0	5,8	72,1	0,0614
4 min	85,2	3,5	73,5	0,0665
5 min	80,3	5,3	73,0	0,0692

W stosunku do obiektu B1, który cechowała duża dynamika i równiomierny rozkład obiekt B2 należy do najtrudniejszych w badanym zbiorze. Maszyna falownikowa SP1 pracuje w sposób ciągły i nie generuje żadnych zdarzeń odciążenia. Liczba dostępnych tranzycji jest niewielka (452). Niebezpieczne okazały się przewidywania zdarzeń dla sprężarki SP3, która miała najwięcej zdarzeń, ale identyfikacja sytuacji gdy była faktycznie potrzebna było prawie niemożliwe co poniekąd jest zrozumiałe dla maszyny dobijającej jako ostatnia, kiedy pojawia się większe zapotrzebowanie. Dla SP3 spośród 104 okien, w których po odciążeniu następował powrót obciążenia, model rozpoznał tylko 9 (S_{FN} = 91,3%). Wskaźnik ten nie poprawia się wraz z długością horyzontu

(sumaryczne S_{FN} oscyluje w wąskim przedziale 80,3–85,2% dla wszystkich czterech okien). Nasuwa się wniosek, że problem nie leży w długości horyzontu predykcji, lecz w charakterystyce tranzycji. Ponieważ najprawdopodobniej mają one charakter nieregularny, który nie jest poprzedzony rozpoznawalnym wzorcem ciśnienia. W takim wypadku prawdopodobnie wymagana byłaby dodatkowa informacja zwrotna ze strony zakładu, która mogłaby zostać wprowadzona jako dodatkowy czynnik.

Niekorzystne będzie również zestawienie wspomnianego wyniku S_{FN} z niskim S_{FP} (2,7% przy 2 min). Uzyskany model rzadko blokuje możliwe wyłączenia, a zarazem w ponad 85% przypadków (przy 2 min) nie rozpoznaje nadchodzącego powrotu obciążenia. Proponuje zatem rekomendacje wyłączeń, które w większości okazują się przedwczesne i niosą ze sobą ryzyko zachwiania ciągłości produkcji. Uzyskany charakter jest odwrotny do pożądanego. System chętnie wskazuje możliwości oszczędności, lecz ich trafność jest niska. Niestety obiekt B2 w obecnej konfiguracji nie kwalifikuje się do wdrożenia produkcyjnego. W tym wypadku kierunkami dalszej analizy byłby inne okresy treningowe, wprowadzenie cech gradientowych oraz weryfikacja, czy cykle SP3 w ogóle zawierają informację predykcyjną w badanych oknach czasowych. Zaleceniem w przypadkach nieprzewidywalności jest sprawdzenie całej konfiguracji układu sterowania, weryfikacja działania maszyny falownikowej, przypadów występowania awarii oraz wykonania prawidłowego serwisu. Może również okazać się, że w tym wypadku próbkowanie 20 s jest zbyt rzadkie i należy je zagęścić.

4.6.3 Obiekt B3

Wyniki dla obiektu B3 przedstawiono w tabelach 4.28 (zdarzenia w oknach decyzyjnych), 4.29 (wskaźniki bezpieczeństwa) oraz 4.30 (wpływ horyzontu predykcji).

Tabela 4.28: Zdarzenia w oknach decyzyjnych dla obiektu B3, horyzont 2 min, `mse_bce`. N_T — liczba tranzycji.

Sprężarka	N_T	TP	TN	FP	FN
SP1	1	1	0	0	0
SP2	2223	1761	230	85	147
SP3	52	2	47	0	3
Suma	2276	1764	277	85	150

Tabela 4.29: Wskaźniki bezpieczeństwa dla obiektu B3, horyzont 2 min, `mse_bce`. Symbol „—” oznacza brak zdarzeń, dla których można wyznaczyć wskaźnik.

Sprężarka	N_T	S_{FN} [%]↓	S_{FP} [%]↓	S [%]↑
SP1	1	0,0	0,0	100,0
SP2	2223	7,7	3,8	89,6
SP3	52	60,0	0,0	94,2
Suma	2276	7,8	3,7	89,7
			RMSE =	0,0554

Tabela 4.30: Wyniki dla obiektu B3 dla różnych horyzontów predykcji, `mse_bce`.

Horyzont	S_{FN} [%]↓	S_{FP} [%]↓	S [%]↑	RMSE↓
2 min	7,8	3,7	89,7	0,0554
3 min	7,9	2,6	90,8	0,0609
4 min	11,4	1,8	88,6	0,0656
5 min	10,0	1,8	89,8	0,0680

Na wstępie warto zwrócić uwagę, że jeżeli sieć radzi sobie z identyfikacją charakterystyki obiektu, oferując większą adekwatność i bezpieczeństwo to obiekt prawdopodobnie będzie nadawał się do wsparcia przez inteligentne decyzje. Rozpatrywany tutaj obiekt B3 osiąga najlepsze wyniki w grupie B. Zagregowana liczba zdarzeń klasy $TP + FN$ liczy 1914 okien. Przy czym największa liczba tranzycji występuje dla sprężarki SP2 (2223 z 2276) uzyskując wskaźnik $S_{FN} = 7{,}7\%$ przy TP = 1761. Model rozpoznaje zatem ponad 92% powrotów obciążenia drugiej najważniejszej maszyny (zaraz za VFD) układu, co czyni jego rekomendacje wyłączeń wiarygodnymi dla systemu sterowania. Ponieważ statystycznie SP3 wypada dużo gorzej dając $S_{FN} = 60{,}0\%$ to należy wziąć pod uwagę, że wskaźnik ten ma do dyspozycji małą liczbę zdarzeń (5 zdarzeń klasy $TP + FN$). Wskaźnik S_{FN} dla SP3 odpowiada trzem niewykrytym powrotom i nie powinien być nadinterpretowywany. Zagregowany wskaźnik S_{FP} pozostaje umiarkowany (3,7% przy 2 min, malejący do 1,8% przy dłuższych horyzontach), co oznacza potencjał dobrego wykorzystania dostępnych możliwości oszczędności energii elektrycznej.

Rozpatrując wskaźniki bezpieczeństwa dla wszystkich horyzontów, zauważalnym jest, że wyniki są stabilne w funkcji horyzontu (S_{FN} w przedziale 7,8–11,4%). Przy braniu pod uwagę tylko S_{FN} najkrótszy horyzont 2 min okazuje się najlepsze z niewielką przewagą. Charakter obiektu jest odzwierciedlony w modelu predykcyjnym i

daje dobre wyniki dla większości horyzontów, a stabilność taka jest pożądaną właściwością wdrożeniową. Obiekt B3 nadaje się do wdrożenia przy dowolnym horyzoncie i stanowi dowód, że obecność maszyny falownikowej nie wyklucza skutecznej predykcji stanów pozostałych sprężarek.

4.6.4 Obiekt B4

Wyniki dla obiektu B4 przedstawiono w tabelach 4.31 (zdarzenia w oknach decyzyjnych), 4.32 (wskaźniki bezpieczeństwa) oraz 4.33 (wpływ horyzontu predykcji).

Tabela 4.31: Zdarzenia w oknach decyzyjnych dla obiektu B4, horyzont 2 min, `mse_bce`. N_T — liczba tranzycji.

Sprężarka	N_T	TP	TN	FP	FN
SP1	0	0	0	0	0
SP2	466	1	391	6	68
SP3	687	5	564	6	112
SP4	721	3	607	0	111
SP5	827	6	670	1	150
Suma	2701	15	2232	13	441

Tabela 4.32: Wskaźniki bezpieczeństwa dla obiektu B4, horyzont 2 min, `mse_bce`. Symbol „—” oznacza brak zdarzeń, dla których można wyznaczyć wskaźnik.

Sprężarka	N_T	S_{FN} [%]↓	S_{FP} [%]↓	S [%]↑
SP1	0	—	—	—
SP2	466	98,6	1,3	84,1
SP3	687	95,7	0,9	82,8
SP4	721	97,4	0,0	84,6
SP5	827	96,2	0,1	81,7
Suma	2701	96,7	0,5	83,2
			RMSE =	0,0944

Tabela 4.33: Wyniki dla obiektu B4 dla różnych horyzontów predykcji, `mse_bce`.

Horyzont	S_{FN} [%]↓	S_{FP} [%]↓	S [%]↑	RMSE↓
2 min	96,7	0,5	83,2	0,0944
3 min	96,3	0,6	83,2	0,1068
4 min	93,9	1,1	83,0	0,1138
5 min	90,4	2,3	82,5	0,1180

Obiekt B4 jest najtrudniejszym przypadkiem z całego zbioru badawczego i jednocześnie stanowi pewną granicę wykorzystania metody. W skład układu wchodzi aż pięć maszyn, z czego wszystkie cztery aktywne sprężarki (SP2–SP5) uzyskują S_{FN} w przedziale 95,7–98,6%. Aktywne sprężarki cechuje też duża liczba tranzycji układająca się narastająco (2701 okien zagregowanych) co świadczy o dołączaniu się kolejnych sprężarek w ramach zapotrzebowania. Jednak pojedynczych zdarzeń dla wszystkich sprężarek jest dużo. Spośród dostępnych 456 okien decyzyjnych, w których po odciążeniu następował powrót obciążenia, model rozpoznał tylko 15. Rekomendacje wyłączeń predykowane dla tego obiektu są w większości nieprawidłowe. Maszyna wskazana jako możliwa do wyłączenia okazuje się potrzebna w ciągu nadchodzących dwóch minut. Co ciekawe towarzyszy temu skrajnie niski wskaźnik $S_{FP} = 0{,}5\%$ (13 zdarzeń na 2701 okien), co oznacza, że model niemal nigdy nie blokuje wyłączenia. Strategia rekomendowana przez sieć jest ryzykowna ponieważ system zawsze sygnalizuje możliwość wyłączenia bez rozróżnienia sytuacji bezpiecznych od niebezpiecznych.

Wydłużenie horyzontu do 5 min przynosi niewielką poprawę ($S_{FN} = 90{,}4\%$). Podobnie jak w przypadku B2 gdzie problem występował dla jednej sprężarki, tutaj równomiernie pojawił się dla wszystkich czterech sprężarek stałoobrotowych. Nieprzewidywalny charakter wskazywałby na prawdopodobnie bardzo dynamiczny, nieregularny trend zapotrzebowania. Prawdopodobnie maszyny wracają do obciążenia szybciej, niż jest to zauważalne przy 20-sekundowym okresie próbkowania. Hipoteza ta jest spójna z obserwacjami z wdrożeń, gdzie dla obiektu o charakterystyce szybkozmiennej wprowadzono gęstsze próbkowanie (rzędu 5 s). Niestety obiekt B4 nie kwalifikuje się do wdrożenia przy użyciu proponowanej architektury i stanowi ważny przykład granic stosowalności metody. Obiekt B4 również dobrze wskazuje potrzebę rozpatrywania wszystkich wskaźników całościowo, ponieważ ocena modelu nie może opierać się na samym wskaźniku S, który przy wartości 83,2% sugeruje wynik akceptowalny.

4.7 Wyniki dla grupy C — obiekty z samymi maszynami VFD

Grupę C stanowią cztery obiekty złożone wyłącznie ze sprężarek wyposażonych w przetwornicę częstotliwości. Są to najnowocześniejsze układy obecnie występujące w przemyśle w których mogłoby się wydawać, że możliwości oszczędności energii są najmniejsze (niewielka liczba tranzycji w wyniku ciągłej adaptacji do poboru). W związku z charakterystyczną dla nich płynną regulacją wydajności występują wyraźnie rzadsze zdarzenia tranzycji niż w pozostałych grupach A i B. Sumaryczna liczba zdarzeń w zbiorach ewaluacyjnych waha się od 186 (C4) do 1146 (C3). Niewielka liczba zdarzeń decyzyjnych będzie prowadzić do dwóch spostrzeżeń. Co zauważono już wcześniej, pojedyncze zdarzenia silnie wpływają na wartości metryk procentowych, co wymaga świadomej, ostrożnej, całościowej interpretacji. Drugim spostrzeżeniem jest to, że potencjał oszczędności energetycznych z tytułu właściwej predykcji stanów pracy jest w tych układach znacznie mniejszy. To właśnie maszyny falownikowe mają na sobie zadanie kompensacji nagłych zmian zapotrzebowania na sprężone powietrze, a przedmiotem predykcji pozostają stosunkowo rzadkie zdarzenia odciążenia i wyłączenia maszyn.

4.7.1 Obiekt C1

Wyniki dla obiektu C1 przedstawiono w tabelach 4.34 (zdarzenia w oknach decyzyjnych), 4.35 (wskaźniki bezpieczeństwa) oraz 4.36 (wpływ horyzontu predykcji).

Tabela 4.34: Zdarzenia w oknach decyzyjnych dla obiektu C1, horyzont 2 min, `mse_bce`. N_T — liczba tranzycji.

Sprężarka	N_T	TP	TN	FP	FN
SP1	105	2	103	0	0
SP2	146	0	138	6	2
SP3	11	2	7	1	1
Suma	262	4	248	7	3

Tabela 4.35: Wskaźniki bezpieczeństwa dla obiektu C1, horyzont 2 min, `mse_bce`. Symbol „—” oznacza brak zdarzeń, dla których można wyznaczyć wskaźnik.

Sprężarka	N_T	S_{FN} [%]↓	S_{FP} [%]↓	S [%]↑
SP1	105	0,0	0,0	100,0
SP2	146	100,0	4,1	94,5
SP3	11	33,3	9,1	81,8
Suma	262	42,9	2,7	96,2
			RMSE =	0,1010

Tabela 4.36: Wyniki dla obiektu C1 dla różnych horyzontów predykcji, `mse_bce`.

Horyzont	S_{FN} [%]↓	S_{FP} [%]↓	S [%]↑	RMSE↓
2 min	42,9	2,7	96,2	0,1010
3 min	28,6	2,3	96,9	0,1145
4 min	42,9	2,3	96,6	0,1257
5 min	28,6	2,7	96,6	0,1325

Pierwszy rozpatrywany obiekt C1 charakteryzuje mała liczba tranzycji (262 łącznie) co potwierdza typową własność dla układów zawierających wyłącznie maszyny falownikowe. Ich obecność ogranicza rozruchy do minimum dzięki możliwości płynnej adaptacji do zmiennego zapotrzebowania. Zbiór zdarzeń decyzyjnych jest niewielki, dla 2 min występuje tylko 7 okien, w których po odciążeniu następuje powrót obciążenia. Należy zauważyć, że każde pojedyncze zdarzenie zmienia wskaźnik o kilkanaście punktów procentowych, co wymaga wspomnianej ostrożności w interpretacji. SP1 rozpoznaje oba wystąpienia powrotu ($S_{FN} = 0{,}0\%$), SP2 nie rozpoznaje żadnego z dwóch ($S_{FN} = 100\%$), a SP3 przy trzech zdarzeniach nie pozwala na wiarygodną ocenę.

Sumaryczny wskaźnik S_{FN} poprawia się z 42,9% (2 min) do 28,6% (3 min i 5 min). Ta wydawałoby się znaczna różnica, w liczbach bezwzględnych oznacza różnicę jednego rozpoznanego zdarzenia, a nie systematyczną przewagę dłuższego horyzontu. Niski i stabilny S_{FP} (2,3–2,7%) oznacza, że model wykorzystuje niemal wszystkie dostępne możliwości wyłączenia. Obiekt może zostać objęty wdrożeniem dla horyzontu 3 min, jednak ze względu na niewielką liczbę zdarzeń decyzyjnych ocena bezpieczeństwa wymaga dłuższego okresu obserwacji produkcyjnej. Wyprzedzając wnioski płynące z rozdziału dotyczącego wdrożenia, w przypadkach niepewności zalecane byłoby wprowadzenie dodatkowego progu ciśnienia poniżej którego, nie są wykonywane decyzje sterujące.

4.7.2 Obiekt C2

Wyniki dla obiektu C2 przedstawiono w tabelach 4.37 (zdarzenia w oknach decyzyjnych), 4.38 (wskaźniki bezpieczeństwa) oraz 4.39 (wpływ horyzontu predykcji).

Tabela 4.37: Zdarzenia w oknach decyzyjnych dla obiektu C2, horyzont 2 min, `mse_bce`. N_T — liczba tranzycji.

Sprężarka	N_T	TP	TN	FP	FN
SP1	0	0	0	0	0
SP2	64	0	55	6	3
SP3	345	8	238	80	19
Suma	409	8	293	86	22

Tabela 4.38: Wskaźniki bezpieczeństwa dla obiektu C2, horyzont 2 min, `mse_bce`. Symbol „—” oznacza brak zdarzeń, dla których można wyznaczyć wskaźnik.

Sprężarka	N_T	S_{FN} [%]↓	S_{FP} [%]↓	S [%]↑
SP1	0	—	—	—
SP2	64	100,0	9,4	85,9
SP3	345	70,4	23,2	71,3
Suma	409	73,3	21,0	73,6
			RMSE =	0,0269

Tabela 4.39: Wyniki dla obiektu C2 dla różnych horyzontów predykcji, `mse_bce`.

Horyzont	S_{FN} [%]↓	S_{FP} [%]↓	S [%]↑	RMSE↓
2 min	73,3	21,0	73,6	0,0269
3 min	70,0	4,9	90,0	0,0306
4 min	90,0	11,0	82,4	0,0343
5 min	80,0	2,9	91,2	0,0372

Obiekt C2 również wykazuje ciekawe indywidualne właściwości. Cechuje go najniższe w całym zbiorze RMSE ciśnienia (0,0269, co oznaczałoby bardzo stabilny jednostajny pobór), który połączony jest ze słabymi wynikami klasyfikacji. Przy proponowanej metodzie, brak sygnału poprzedzającego załączenie i niewielkie wahania ciśnienia będą prowadzić do tego, że przy 2 min model nie rozpoznaje 22 z 30 powrotów obciążenia ($S_{FN} = 73{,}3\%$). Z kolei dla 86 okien blokuje możliwe wyłączenie ($S_{FP} = 21{,}0\%$).

Tutaj oba wskaźniki są jednocześnie wysokie co wyróżnia go spośród pozostałych obiektów zbioru. Przy tak jednostajnej charakterystyce i dynamicznych zmianach zapotrzebowania model myli się w obie strony, przy czym rola SP3, odpowiadająca za większość zdarzeń obu klas, pozostaje nieprzewidywalna. Zauważalne jest zatem stabilne ciśnienie i słaba klasyfikacja. Tranzycje na tym obiekcie nie są poprzedzone wyraźnym trendem w przebiegu ciśnienia. Maszyny falownikowe kompensując płynnie zmiany zapotrzebowania, „ukrywają" nadchodzące przełączenia przed modelem bazującym głównie na ciśnieniu.

Wyniki zmieniają się niemonotonicznie z horyzontem (S_{FP}: 21,0% → 4,9% → 11,0% → 2,9%), co również utrudnia wybór konfiguracji produkcyjnej. Wyraźnie można zauważyć, że model nie ma wystarczających informacji by oferować odpowiednie wsparcie decyzyjne przez co w obecnym stanie nie kwalifikuje się do wdrożenia (chociaż model nauczył się, że gdy włączy się SP2, to raczej nie będzie potrzebna i można ją wyłączyć). Ponownie, być może gdyby posiadać od strony klienckiej dodatkowe informacje np. o liczbie i wielkości aktywnych odbiorców, to wzbogacenie wektora wejściowego o te dane dałyby brakującą składową, pozwalającą na adekwatniejsze predykcje.

4.7.3 Obiekt C3

Wyniki dla obiektu C3 przedstawiono w tabelach 4.40 (zdarzenia w oknach decyzyjnych), 4.41 (wskaźniki bezpieczeństwa) oraz 4.42 (wpływ horyzontu predykcji).

Tabela 4.40: Zdarzenia w oknach decyzyjnych dla obiektu C3, horyzont 2 min, `mse_bce`. N_T — liczba tranzycji.

Sprężarka	N_T	TP	TN	FP	FN
SP1	555	2	553	0	0
SP2	255	0	254	1	0
SP3	336	0	336	0	0
Suma	1146	2	1143	1	0

Tabela 4.41: Wskaźniki bezpieczeństwa dla obiektu C3, horyzont 2 min, `mse_bce`. Symbol „—” oznacza brak zdarzeń, dla których można wyznaczyć wskaźnik.

Sprężarka	N_T	S_{FN} [%]↓	S_{FP} [%]↓	S [%]↑
SP1	555	0,0	0,0	100,0
SP2	255	—	0,4	99,6
SP3	336	—	0,0	100,0
Suma	1146	0,0	0,1	99,9
			RMSE =	0,0315

Tabela 4.42: Wyniki dla obiektu C3 dla różnych horyzontów predykcji, `mse_bce`.

Horyzont	S_{FN} [%]↓	S_{FP} [%]↓	S [%]↑	RMSE↓
2 min	0,0	0,1	99,9	0,0315
3 min	0,0	0,3	99,7	0,0408
4 min	0,0	0,3	99,7	0,0469
5 min	0,0	0,1	99,9	0,0536

Obiekt C3 jest obiektem, którego praca jest zoptymalizowana i nie cechuje go duża dynamika przebiegu ciśnienia. Przede wszystkim osiąga najlepszy wynik wskaźników bezpieczeństwa w całym zbiorze badawczym dla każdego z czterech horyzontów przy $S_{FN} = 0{,}0\%$ oraz $S_{FP} \leq 0{,}3\%$. Model rozpoznaje każdy powrót obciążenia i tylko w pojedynczych oknach niepotrzebnie blokuje możliwe wyłączenie (1–3 zdarzenia na 1146 okien). Model nauczył się, że zwykle w przypadku odciążenia maszyny, nie będzie ona już potrzebna. Tak więc cechą charakterystyczną tego obiektu jest obserwacja, że zróżnicowanie zdarzeń decyzyjnych jest skrajnie małe. Tylko SP1 generuje zdarzenia klasy TP (2 przy każdym horyzoncie), natomiast SP2 i SP3 nie generują ani jednego zdarzenia TP lub FN. Po odciążeniu nie wracają do obciążenia w oknie decyzyjnym co opisuje charakter pracy tego układu i zostało wychwycone przez sieć neuronową. W związku z taką dynamiką praktycznie całość okien tego obiektu stanowią trafne rozpoznania braku powrotu do obciążenia, czyli realizowalne możliwości wyłączenia.

Co pokazono wcześniej stabilność wyników niezależna od horyzontu predykcji wskazuje na regularną, łatwą do modelowania charakterystykę obiektu, który działa w sposób optymalny. Maszyny po odciążeniu pozostają w biegu jałowym przez czas znacznie dłuższy niż okno decyzyjne więc nie ma potrzeby utrzymywania ich długo w stanie biegu jałowego. Jednocześnie zagregowana liczba zdarzeń TN dla wszystkich na poziomie 1143 daje potencjalne oszczędności energii elektrycznej. C3 nadaje się do wdrożenia

przy dowolnym horyzoncie. Rozpatrując wyniki całościowo, można zauważyć, że prawie zerowa liczba zdarzeń ryzykownych oznacza, że sam wynik $S_{FN} = 0\%$ wynika ze sprzyjającej charakterystyki obiektu. Obiekt ten, użyty również w studium ablacji jako reprezentant grupy C, stanowi przykład układu, dla którego już bazowa konfiguracja $\mathcal{L}_{MSE} + \mathcal{L}_{BCE}$ z architekturą BiLSTM jest w pełni wystarczająca. Należy również podkreślić, że jest to układ zoptymalizowany gdzie maszyny jak najlepiej dopasowują się do zapotrzebowania na sprężone powietrze.

4.7.4 Obiekt C4

Wyniki dla obiektu C4 przedstawiono w tabelach 4.43 (zdarzenia w oknach decyzyjnych), 4.44 (wskaźniki bezpieczeństwa) oraz 4.45 (wpływ horyzontu predykcji).

Tabela 4.43: Zdarzenia w oknach decyzyjnych dla obiektu C4, horyzont 2 min, `mse_bce`. N_T — liczba tranzycji.

Sprężarka	N_T	TP	TN	FP	FN
SP1	2	0	0	2	0
SP2	57	8	39	8	2
SP3	99	1	93	5	0
SP4	28	0	25	3	0
SP5	0	0	0	0	0
Suma	186	9	157	18	2

Tabela 4.44: Wskaźniki bezpieczeństwa dla obiektu C4, horyzont 2 min, `mse_bce`. Symbol „—” oznacza brak zdarzeń, dla których można wyznaczyć wskaźnik.

Sprężarka	N_T	S_{FN} [%]↓	S_{FP} [%]↓	S [%]↑
SP1	2	—	100,0	0,0
SP2	57	20,0	14,0	82,5
SP3	99	0,0	5,1	94,9
SP4	28	—	10,7	89,3
SP5	0	—	—	—
Suma	186	18,2	9,7	89,2
			RMSE =	0,0779

Tabela 4.45: Wyniki dla obiektu C4 dla różnych horyzontów predykcji, `mse_bce`.

Horyzont	S_{FN} [%]↓	S_{FP} [%]↓	S [%]↑	RMSE↓
2 min	18,2	9,7	89,2	0,0779
3 min	36,4	18,3	79,6	0,0813
4 min	45,5	9,1	88,2	0,0809
5 min	63,6	16,1	80,1	0,0726

Obiekt C4 składa się z pięciu sprężarek falownikowych, przy czym zauważalne są maszyny wiodące tj. SP1 i SP5. Ich stan pozostaje praktycznie niezmienny (0–2 tranzycje). Maszyny dobijające, na których można oprzeć obserwacje to SP2, SP3 i SP4, przy łącznie zaledwie 186 tranzycjach. Wynik przy 2 min jest akceptowalny ($S_{FN} = 18{,}2\%$, czyli 2 nierozpoznane powroty obciążenia z 11 zdarzeń klasy $TP + FN$; $S_{FP} = 9{,}7\%$), jednak systematycznie pogarsza się z horyzontem. Dla horyzontu 5 min S_{FN} rośnie do poziomu 63,6%. Ta ciekawa własność zachowuje się odwrotnie niż w większości obiektów i wskazuje, że w układach czysto falownikowych o rzadkich tranzycjach dłuższy horyzont predykcji wprowadza więcej zakłóceń niż istotnego sygnału. Płynna regulacja przetwornic częstotliwości sprawia, że stan układu sprzed 4–5 minut niestety nie niesie informacji o nadchodzącym przełączeniu.

Przy bardzo małej populacji zdarzeń spotykamy się tutaj z przypadkiem, który zaburza obraz tabeli i narzuca konieczność ostrożnej interpretacji. Wskaźnik bezpieczeństwa $S_{FP} = 100\%$ dla SP1 odpowiada dokładnie dwóm oknom, w których model zablokował możliwe wyłączenie przy tylko dwóch tranzycjach. Ponieważ dla horyzontu 2 min uzyskano akceptowalne wyniki śledzące dynamikę układu obiekt C4 może być rozważany do wdrożenia wyłącznie przy takim horyzoncie z zaleceniem ciągłej obserwacji produkcyjnej.

4.8 Analiza wpływu horyzontu predykcji

Zmiana horyzontu predykcji nie zmienia długości okna decyzyjnego W, dlatego wskaźniki S, S_{FN} i S_{FP} dla wszystkich horyzontów wyznaczane są na tym samym zbiorze zdarzeń, a horyzont wpływa na nie wyłącznie przez model (tabela 3.1).

Tabela 4.46: Wskaźnik S_{FN} [%] dla wszystkich obiektów i horyzontów predykcji, `mse_bce`. Pogrubiono najniższą wartość dla każdego obiektu.

Obiekt	2 min	3 min	4 min	5 min
A1	67,1	47,9	**39,7**	**39,7**
A2	0,2	**0,0**	11,7	1,8
A3	**42,9**	**42,9**	**42,9**	**42,9**
A4	**0,9**	1,1	1,7	1,9
B1	41,6	37,7	32,4	**27,2**
B2	85,2	82,0	85,2	**80,3**
B3	**7,8**	7,9	11,4	10,0
B4	96,7	96,3	93,9	**90,4**
C1	42,9	**28,6**	42,9	**28,6**
C2	73,3	**70,0**	90,0	80,0
C3	**0,0**	**0,0**	**0,0**	**0,0**
C4	**18,2**	36,4	45,5	63,6

Najpierw należy odnieść się do odmiennych wzorców zależności ryzyka wyłączenia FN od horyzontu predykcji. W rozpatrywanej grupie pojawiły się cztery takie wzorce.

Pierwszy — systematyczny spadek — S_{FN} wraz z wydłużaniem horyzontu. Wystąpił dla obiektów A1 (67,1% → 39,7%), B1 (41,6% → 27,2%) oraz, słabiej, B4. Dla tych układów powrót obciążenia poprzedzony jest wzorcem ciśnienia, który dłuższe okno obserwacji pozwala uchwycić. Nadążanie za charakterystyką wpływa na to, że model rzadziej rekomenduje wyłączenie maszyny, która za chwilę mogłaby okazać się potrzebna.

Drugi — niezależność od horyzontu — który reprezentują A3 (stałe 42,9%, odpowiadające sześciu nierozpoznawalnym zdarzeniom) i C3 (stałe 0,0%). Przyczyną jest struktura zdarzeń oraz charakterystyka obiektu a nie parametry modelu.

Trzeci — wzrost ryzyka z horyzontem — który dotyczy obiektów C4 (18,2% → 63,6%), A4 (0,9% → 1,9%) oraz B3 (7,8% → 10,0%). W układzie czysto falownikowym, o rzadkich tranzycjach, odległa historia ciśnienia nie niesie wystarczających informacji o nadchodzącym powrocie do obciążenia. Wydłużanie okna obserwacji powoduje pojawienie się dodatkowych zakłóceń. Zauważona tendencja jest analogiczna do spostrzeżeń jakie można znaleźć w literaturze [75].

Czwarty — brak wyraźnej tendencji — obejmuje A2, B2, C1 oraz C2. Dla C1 wynik zmienia się naprzemiennie między 42,9% a 28,6%, co przy siedmiu zdarzeniach decyzyjnych oznacza różnicę jednego zdarzenia. Podobnie zachowuje się C2, a B2 wraca przy 4 min do wartości wyjściowej. W obiekcie A2 zdarzeń jest nieporównanie więcej, więc skok do 11,7% przy 4 min nie wynika z małej próby i wymaga osobnego wyjaśnienia.

Kolejnym spostrzeżeniem jest fakt, że okno decyzyjne ma stałą długość, więc zbiór zdarzeń decyzyjnych jest identyczny dla wszystkich horyzontów, a różnice wskaźnika S_{FN} wynikają wyłącznie z modelu, czyli z długości sekwencji wejściowej i prognozy. Horyzont musi być dobierany empirycznie dla każdego obiektu indywidualnie. Dla układów grup A i B dobrym punktem startowym jest zakres 3–5 min. Dla układów czysto falownikowych (grupa C) punktem startowym jest najkrótszy horyzont 2–3 min. Wyjątkiem są obiekty o wzorcu trzecim (A4, B3, C4), dla których wydłużanie horyzontu pogarsza wynik i punktem startowym pozostaje 2–3 min. Wybór ten należy zawsze zestawiać ze wskaźnikiem S_{FP}, ponieważ obniżenie ryzyka może wiązać się z utratą części możliwych oszczędności (przypadek B1). Tutaj również należy podkreślić, że trzeba zapoznać się z rzeczywistymi wynikami pojedynczych tranzycji z przyjrzeniem się szczególnie decyzjom FN. Wskaźniki stanowią doskonały punkt do oceny ryzyka wdrożenia, jednak wykresy pokazują zależności, które statystycznie mogą być nie w pełni uchwycone.

4.9 Podsumowanie

Zaproponowane metoda, jak wykazały przeprowadzone badania na 12 obiektach, jest skuteczna tylko dla określonej charakterystyki układu sprężonego powietrza. W związku z możliwą nieprzewidywalnością procesu dobór architektury i funkcja straty miały znaczenie marginalne. Wysoka wartość $\tau_{<h}$ wyznacza granicę stosowalności metody, ponieważ powrót obciążenia wewnątrz horyzontu, nie zostawia wiele miejsca na optymalizację. Co zbadano, to niezbędność dwukierunkowości przetwarzania, gdzie architektury jednokierunkowe ograniczały się do trafnego przewidywania zachowania tylko jednej maszyny. Wybrana architektura w połączeniu z asymetryczną karą za błędy krytyczne poprawia wynik (chociaż dla obiektu B1 wariant pogorszył wyniki). Wskaźniki oceny powiązane są z horyzontem predykcji, przy czym zaobserwowano cztery różne wzorce, co wyklucza przyjęcie jednego horyzontu dla wszystkich układów. Wnioski te stanowią punkt wyjścia dla doboru obiektów do wdrożenia produkcyjnego, co opisane zostało w rozdziale 5.

5 Wdrożenie produkcyjne

Rozdział skupia się na wdrożeniu produkcyjnym czterech spośród zbadanych wcześniej obiektów oraz kolejnych czterech, które nie występowały w poprzednich rozdziałach. Na początku przedstawiono architekturę modułu SI przeznaczonego do współpracy ze sterownikami PLC oraz opisano zabezpieczenia i zalecany przebieg wdrożenia. Następnie przedstawiono charakterystyki wdrożeń i metryki dla modeli produkcyjnych. Kolejno opisano powstały w związku ze wdrożeniem wskaźnik nadzoru operacyjnego, osiągnięte potencjalne oszczędności energii elektrycznej oraz wyniki 806 dni wdrożenia. Poświęcono również uwagę opisowi czterech zauważonych mechanizmów degradacji oraz opisano rolę dotrenowania podczas eksploatacji. Przed podsumowaniem kończącym rozdział pojawia się obszerne studium przypadku obiektu, w którym domyślna metodyka się nie sprawdziła i wymagała adaptacji. Ten istotny przypadek podkreśla znaczenie indywidualnego rozpatrywania każdego przypadku i jest przewodnikiem działania.

5.1 Architektura modułu SI

Sprężone powietrze ma opinię medium silnie nieprzewidywalnego. O ile publikacje chętnie dzielą się statystycznymi wskaźnikami, gdzie stosunkowo łatwo uzyskać dobre wyniki np. pierwiastka błędu średniokwadratowego [63] czy znaleźć wzorzec określonego zużycia energii w oknach czasowych [14], to działanie w czasie rzeczywistym jest innym wyzwaniem. Proponowana metodyka przede wszystkim używa wskaźników w sposób, który zawęża się do praktycznego wykorzystania metody. Również ewaluacja następuje na danych z następującego miesiąca, na których obiekt nie był trenowany. Jednak nawet najbardziej rygorystyczne wskaźniki oceniające ryzyko wdrożenia nie odpowiadają na pytanie, jak implementacja wsparcia decyzyjnego sprawdzi się w praktyce. W tym wypadku obowiązkowym jest wykroczenie poza warunki laboratoryjne i zweryfikowanie hipotez w warunkach eksploatacji przemysłowej. Niniejszy rozdział przedstawia wyniki wdrożenia produkcyjnego predykcyjnego systemu wsparcia sterowników nadrzędnych sprężarkowni na ośmiu obiektach przemysłowych. Ponieważ w momencie wdrożenia uzyskano więcej doświadczenia i wiedzy co do proponowanego rozwiązania, zdecydowano się na wybranie czterech obiektów, które zostały już przebadane oraz czterech

innych reprezentatywnych. Co istotne jeden z obiektów wdrożeniowych, któremu postanowiono się przyjrzeć w studium przypadku, ukazuje bardzo ciekawą sytuację, gdzie nastąpiło zderzenie doskonałych wskaźników z katastrofalnym wdrożeniem produkcyjnym.

W rozdziale 1 opisano ewolucję przemysłową odejścia od układów przekaźnikowych, przejścia do sterowników programowalnych PLC [34] oraz systemów wizualizacji, zbierania i przetwarzania danych pomiarowych SCADA [39], której kolejnym krokiem jest zaproponowana w rozprawie predykcyjna warstwa decyzyjna wykorzystująca uczenie maszynowe. Takie rozwiązanie, pozwala na integrację z istniejącą infrastrukturą przemysłową a prowadzone wdrożenie nie wymaga modyfikacji sprzętowej ani rozbudowy opomiarowania obiektu (co byłoby istotnym dodatkowym kosztem). Program PLC również łatwo przygotować zarówno do współpracy z systemem predykcyjnym jak i sterowania sprężarkami. Dużym atutem takiego rozwiązania jest skupienie się na tym co najistotniejsze, czyli rekomendacji wyłączeń z warstwy predykcyjnej, a deterministyczne warunki eksploatacyjne i bezpieczeństwa sterowania maszyn pozostają po stronie nadrzędnego sterownika PLC (oczywiście z pewnymi ograniczeniami co można zauważyć w dodatku B). Wypada zaznaczyć, że nadal pozostaje ryzyko zachwiania dostaw przy decyzji błędnej rekomendacji wyłączenia. Sterownik PLC, nawet wyposażony w prosty algorytm predykcyjny bazujący np. na regresji liniowej i gradiencie, nie niesie w sobie wiedzy o działaniu obiektu w okresie długofalowym. Dlatego nadal układ predykcyjny musi być jak najbardziej adekwatny, jednak wyposażony w dodatkowe reguły zabezpieczające wynikające ze specyfiki pracy danego obiektu (np. nie wyłączaj maszyny, gdy włączenie nastąpiło w celu dogrzania jej w związku z niską temperaturą na hali w okresie zimowym). Podsumowując: najważniejszym założeniem architektonicznym jest założenie bycia warstwą nadrzędną nad istniejącym systemem sterowania, a nie zastępowanie jego roli, ponieważ warstwa predykcyjna nie przejmuje sterowania zespołem sprężarek i roli sterownika nadrzędnego, lecz dostarcza mu rekomendacji wcześniejszego wyłączenia.

Ze względu na jak najłatwiejsze utrzymanie w ruchu pracujących układów oraz możliwości wprowadzenia szybkich zmian, po wcześniejszych testach wdrożeniowych na zasadach przetwarzania brzegowego (ang. *edge computing*) wymagających cyklicznych wyjazdów na obiekt, zdecydowano się na zrealizowanie przetwarzania danych w oddalonym centralnym punkcie (poniekąd odgrywającym rolę przytoczonej w przeglądzie literatury sterowni [36]) bez konieczności instalacji sprzętu obliczeniowego na sprężarkowni. Moduł decyzyjny sztucznej inteligencji (SI) odczytuje w 5-sekundowym cyklu bieżący stan układu. Dane zostały zminimalizowane do jak najmniejszej paczki, która zawiera w sobie najistotniejsze informacje — ciśnienie w sieci, stany binarne sprężarek (obciążenia, awarie, rola w układzie sterowania lub bity związane ze specyfiką danego obiektu). Dane odczytywane są bezpośrednio ze sterownika PLC przy pomocy

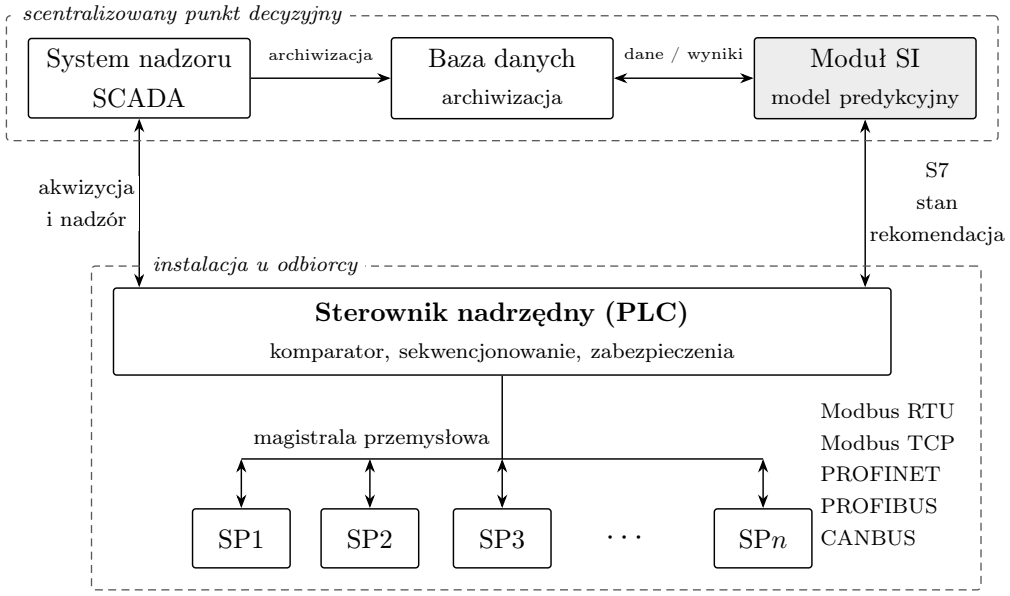
protokołu S7. Dane historyczne dotyczące poprzednich stanów, niezbędne dla uczenia maszynowego, pobierane są bezpośrednio z bazy danych systemu SCADA.

Moduł decyzyjny SI śledzi na bieżąco charakterystykę obiektu, w momencie, gdy wykryje stan odciążenia danej maszyny, następuje przygotowanie sekwencji wejściowej dla sieci neuronowej. Należy zwrócić uwagę na bardzo istotną rolę próbkowania, ponieważ modele uczone są na danych z próbkowaniem 20-sekundowym, będą podejmować adekwatne decyzje tuż po 20-sekundowym stanie odciążenia. Ta informacja jest przekazywana w sposób niejawny, ale sprawdzono, że przyspieszenie decyzji może prowadzić do destabilizacji całego modułu decyzyjnego SI. Czas, po którym następuje wysłanie rekomendacji, jest kluczowym parametrem i powinien być świadomie dobrany. Następnie program sprawdza ograniczające warunki regułowe, które mogą powstrzymać uruchomienie sieci neuronowej. Jeżeli nie ma przeciwwskazań do wykonania predykcji, następuje uruchomienie modelu oraz predykcja. Jej wynik sprawdzany jest z regułami wyjściowymi i jeżeli osiągnięty jest stan, w którym nie jest przewidywane obciążenie maszyny, następuje wysłanie rekomendacji wyłączenia do bloku danych sterownika PLC. Wszystkie decyzje, wraz z ich późniejszą weryfikacją, zapisywane są zarówno w lokalnych plikach dziennika danych, jak i w bazie danych systemu monitoringu. Dzięki centralnemu punktowi zbierania danych pełna historia decyzji jest dostępna dla systemu nadrzędnego.

Wybór architektury scentralizowanej wiąże się z możliwymi przerwami z dostępem do obiektów oddalonych. Jednak w takim wypadku lokalny sterownik nadrzędny będzie nadal pracował w pełni autonomicznie według swoich deterministycznych reguł. Jedynym problemem w takim wypadku będzie utrata potencjalnych oszczędności bez zagrożenia zachwiania stałości dostaw sprężonego powietrza.

Od strony programistycznej oprogramowanie zostało przygotowane zgodnie z dobrą praktyką inżynierską, ze świadomością możliwości utraty komunikacji z systemem bazodanowym czy sterownikiem PLC. Utworzony program umożliwia działanie bez bazy danych w oparciu o buforowanie pomiarów dla potrzeb decyzji.

Na rysunku 5.1 przedstawiono schemat blokowy architektury wdrożeniowej. W systemie SCADA archiwizowane są dane procesowe oraz informacje o działaniu układu. Moduł SI korzysta z danych historycznych, za pomocą protokołu S7 komunikuje się z nadrzędnym sterownikiem PLC odczytując aktualne stany i wysyłając rekomendacje sterowania. Sterownik PLC realizuje funkcje komparatora poziomu ciśnienia, sekwencjonowania pracy maszyn oraz zabezpieczenia technologiczne. Dane ze sprężarek są dostarczone sterownikowi PLC przy wykorzystaniu magistrali przemysłowych.



Rysunek 5.1: Schemat blokowy architektury wdrożeniowej.

5.2 Zabezpieczenia i przebieg wdrożenia

5.2.1 Parametry zabezpieczeń

Wypada podkeślić, że sprężarki mimo podobnych zasad działania mogą się od siebie znacznie różnić konstrukcyjnie. Dobrą analogią tutaj jest liczba dostępnych obecnie samochodów osobowych i fakt, że co roku wychodzą nowe, ulepszone modele. Ta analogia odnosi się również do sprężarek śrubowych. Druga kwestia to fakt, że dynamika układu pomiędzy obiektami, jak pokazano w poprzednim rozdziale, jest zróżnicowana. W związku z tym należy wprowadzić ograniczenia regułowe, które odzwierciedlają inną konstrukcję oraz zmienność obiektu, przez co rekomendacja modelu nie będzie realizowana bezwarunkowo. Pomiędzy predykcją a decyzją sterującą wprowadzono warstwę następujących reguł deterministycznych:

- **limit wyłączeń w jednostce czasu** — indywidualne ograniczenie liczby wyłączeń danej sprężarki w ciągu godziny, zgodnie z wytycznymi producenta (parametr związany z wielkością oraz rodzajem użytego silnika trójfazowego stopnia sprężającego),
- **próg ciśnienia decyzyjnego** — blokada rekomendacji przy ciśnieniu poniżej wartości granicznej, niezależnie od predykcji modelu, podejmowanie decyzji np. odroczonej w czasie poniżej progu, na którym był uczony będzie nieadekwatne, próg może odzwierciedlać próg ciśnienia, przy którym sterownik nadrzędny obciąża kolejne maszyny, lub być zabezpieczeniem przy bardzo szybkich spadkach ciśnienia, tak aby wyeliminować sytuacje, gdy następuje nagły spadek ciśnienia a model zdążył już rekomendować wyłączenie maszyny,

- **weryfikacja stanu operacyjnego maszyn** — kontrola gotowości dostępnych maszyn, poprawności komunikacji, integracji w systemie sterowania przed dopuszczeniem decyzji o wyłączeniu.

Wprowadzona warstwa bezpieczeństwa pełni podwójną, bardzo istotną rolę. Przede wszystkim chroni proces przed skutkami błędnej rekomendacji. Druga jej funkcja nie jest widoczna od razu, a zostanie opisana w sekcji 5.8: zapewnia uniknięcie procesu autozapętlenia procesu uczenia, gdzie sieć uczy się np. własnych nieprawidłowych reakcji na zmienność poboru. Dzięki takiemu rozwiązaniu proces dotrenowania modelu, nie jest w pełni zdeterminowany jego samodzielnym działaniem, tylko niesie w sobie wiedzę o dynamice obiektu (natomiast reguła ograniczająca musi zostać dobrana empirycznie).

5.2.2 Opóźnienie realizacji decyzji

Wcześniej wspomniano już o istotności kwestii próbkowania w perspektywie decyzji predykcyjnych, jednak niezależnie kwestia opóźnienia realizacji decyzji sterującej, wymaga wprowadzenia dodatkowego parametru konfiguracyjnego. Parametr ten jest niezależny od okresu próbkowania oraz horyzontu predykcji. Rekomendacja utworzona przez model nie jest przekazywana do sterownika natychmiastowo, lecz po upływie zadanego czasu, w którym stan układu podlega ponownej weryfikacji.

Takie obwarowanie wynika z obserwacji związanych ze wdrożeniem: predykcja może poprawnie wskazywać moment możliwego wyłączenia, jednak w chwili jej wygenerowania charakterystyka procesu przykładowo ulega gwałtownej zmianie (co oznacza, że odbyła się w stanie nieustalonym obiektu). Decyzja podjęta na podstawie stanu przejściowego prowadzi wówczas do błędu, mimo poprawnej predykcji.

W większości obiektów dynamika nie wymagała aktywowania opóźnienia, tym bardziej dobór jego wartości jest zagadnieniem empirycznym i zależy od dynamiki konkretnego obiektu. Jako przykład takiego doboru przedstawiono studium przypadku w sekcji 5.9. Należy podkreślić, że okres próbkowania oraz opóźnienie realizacji decyzji sterującej odpowiadają za odmienne własności systemu: pierwszy określa rozdzielczość, z jaką obserwowany jest proces, drugi natomiast odporność decyzji na stany nieustalone. Parametry te muszą być zatem dobierane niezależnie i są niezbędne do działania, szczególnie w przypadkach trudnych.

5.2.3 Weryfikacja decyzji

Przy wdrożeniu systemu z modułem SI niezbędne jest wprowadzenie mechanizmu weryfikacji następczej. Jego celem jest, aby każda wysłana rekomendacja wyłączenia sprężarki była rejestrowana i, po upływie horyzontu predykcji, została porównana z rzeczywistym zachowaniem układu. Jeżeli w oknie decyzyjnym maszyna powróciła do

obciążenia, decyzja klasyfikowana jest jako operacyjnie błędna i rejestrowana w dzienniku działania systemu.

Przyjęte kryterium wymaga rozwinięcia. Po pierwsze spadek ciśnienia procesowego jest najbardziej niebezpieczny i może prowadzić do zatrzymania procesu. W związku z tym powrót maszyny do stanu obciążenia nie jest miarą szkody, a jest wskaźnikiem zastępczym. Rzeczywisty skutek błędnej decyzji to utrzymywanie się spadku ciśnienia procesowego poniżej dolnej granicy roboczej sterownika nadrzędnego. W takim przypadku sprężarkownia nie nadąża z pokryciem zapotrzebowania. Gdy wyłączona maszyna wraca do stanu tłoczenia, sygnalizuje to, że była potrzebna, jednak efekt ten mógł nie wyrządzić szkody procesowej. Jeżeli układ dysponował maszyną rezerwową lub odpowiednią pojemnością sieci, doszło do kompensacji niedoboru przed osiągnięciem wartości krytycznej. Przy wdrożeniu zaobserwowano, że nie doszło do krytycznych spadków ciśnienia w żadnym przypadku poza obiektem D3. Gdyby wprowadzić taki wskaźnik dla wdrożonych obiektów to wartość w większości wynosiłaby zero, przez co nie miałby żadnej funkcji diagnostycznej. Dla obiektu D3 w całym okresie eksploatacji odnotowano 16 decyzji sklasyfikowanych jako błędne, którym towarzyszyły jedynie dwa zdarzenia rzeczywistego spadku ciśnienia poniżej granicy bezpieczeństwa (jednak należy pokreślić, że do niewiele mniejszego spadku doszłoby prawdopodobnie również w ramach normalnej pracy układu).

W związku z tym decyzja o rejestracji zdarzeń potencjalnie niebezpiecznych jest przyjęta świadomie, nawet jeśli tak zakwalifikowane zdarzenie nie prowadzi do niebezpiecznych skutków. Wskaźnik oparty na powrotach obciążenia pozwolił wykryć narastającą degradację modelu na obiekcie D3 na wiele dni przed wystąpieniem realnych zakłóceń. Założeniem było, by rejestracja zdarzeń poprzedzających szkodę pozwalała na reakcję wyprzedzającą (standardowe podejście przy analizie prewencyjnej). Mechanizm weryfikacji decyzji na produkcji realizuje więc koncepcję wprowadzenia błędu krytycznego, która została użyta w funkcji straty (sekcja 3.2) oraz w metrykach ewaluacji (sekcja 3.5). Takie rozwiązanie zapewnia spójność metodyki na poziomie: uczenia, oceny offline (rozumianej jako nie na produkcji) oraz nadzoru eksploatacyjnego.

5.2.4 Schemat przygotowania danych

Modele produkcyjne uczone były według jednolitego schematu, zgodnego z przedstawioną metodyką. Zbiór treningowy stanowią trzy tygodnie danych, zbiór walidacyjny to tydzień następujący bezpośrednio po nim (podział 75%/25% w obrębie miesięcznego okna). Ewaluacja prowadzona jest na danych z kolejnego miesiąca, czyli danych, których model nie poznał w procesie uczenia. Zastosowana ocena walidacji modelu uczenia maszynowego to OOT (ang. *out-of-time*), gdzie dane są oddalone od okresu treningowego, spodziewając się, że taka metoda będzie odpowiadać rzeczywistym warunkom

produkcyjnym i weryfikować praktyczną użyteczność modelu.

5.2.5 Dopuszczenie obiektu do wdrożenia

W przypadku wdrożenia produkcyjnego, ryzyka, które może powodować zakłócenia produkcji, zdecydowano się na wdrożenie dwuetapowe:

- pierwszy etap to **ocena ilościowa** — analiza metryk bezpieczeństwa oraz liczebności zdarzeń decyzyjnych,
- drugi etap to **inspekcja ekspercka** przebiegów predykcji w sytuacjach standardowych oraz ryzykownych. Przegląd obejmował nadążanie za charakterystyką pracy układu, a następnie zawężenie analizy tylko do momentów decyzyjnych ze szczególną uwagą na błędne decyzje. Następowała weryfikacja czy na bazie wcześniejszych przebiegów sieć miała podłoże do wykonanych decyzji, czy proces pozostał nieprzewidywalny. Procedura ta nie jest zbędna. Przykładem może być obiekt B4, który w badaniach porównawczych (rozdział 4) uzyskał najgorsze wskaźniki bezpieczeństwa w całym zbiorze ($S_{FN} = 96{,}7\%$), inspekcja przebiegów potwierdziła ocenę ilościową i podjęto decyzję o rezygnacji z wdrożenia. Metryki wskazały kandydata do odrzucenia, a weryfikacja wykresów przez inżyniera niezależnie tę ocenę potwierdziła.

5.3 Charakterystyka wdrożeń

Ponieważ charakterystyka obiektów sklasyfikowanych w ramach grup odbiega od siebie, zdecydowano o wybraniu jednego reprezentatywnego obiektu z każdej grupy przedstawionej w rozdziale 4: są to obiekty A1, B1 oraz C1. Dodatkowo został wdrożony obiekt A3, który stanowił inspirację do rozpoczęcia badań w ramach rozprawy doktorskiej. Obiekty wdrożeniowe zostały oznaczone literą D, gdyby jednak należało je przypisać do poszczególnych grup to obiekt D1, D2, D4 należałyby do grupy B, która jest często spotykanym doborem maszyn w sprężarkowni (wiodąca maszyna falownikowa oraz pozostałe maszyny stałoobrotowe). Do grupy A należy obiekt D3, który składa się z dwóch maszyn, z których obie są potrzebne podczas produkcji i żadna z nich nie jest maszyną rezerwową. Podsumowując system uruchomiono na ośmiu obiektach a zestawienie zbiorcze przedstawiono w tabeli 5.1. Horyzonty wdrożeniowe nie zawsze pokrywają się z zaleceniami z rozdziału 4. Przed każdym wdrożeniem, z wyjątkiem obiektu A3 omówionego w dalszej części, model uczono na danych bezpośrednio je poprzedzających. Horyzont wybierano ponownie na podstawie wskaźników dla wszystkich horyzontów oraz manualnej oceny wykresów z decyzjami modelu.

Tabela 5.1: Charakterystyka wdrożeń produkcyjnych. N — liczba sprężarek w układzie, h — horyzont predykcji, strojenie — optymalizacja bayesowska hiperparametrów.

Obiekt	Grupa	N	h	Funkcja straty	Strojenie	Dni pracy
A1	A	3	2 min	`mse_bce`	nie	134
A3	A	4	2 min	`industrial`	nie	65
B1	B	3	3 min	`industrial`	tak	48
C1	C	3	2 min	`mse_bce`	nie	128
D1	D	3	3 min	`mse_bce`	tak	135
D2	D	2	3 min	`mse_bce`	tak	134
D3	D	2	2 min	`industrial`	tak	35
D4	D	4	2 min	`mse_bce`	nie	127
					Łącznie	**806**

Rozruch systemu odbył się w okresie ośmiu tygodni. Łączny czas pracy systemu w warunkach produkcyjnych, analizowany na potrzeby rozprawy wyniósł 806 dni. Dwa obiekty: B1 oraz D3 zostały w trakcie eksploatacji wyłączone a przyczyny obu decyzji zostały omówione odpowiednio w sekcjach 5.6 oraz 5.7.2. Obiekt A3 z kolei został wykluczony po dwóch miesiącach ze względu na prace modernizacyjne 5.7.1.

W publikacjach własnych zaproponowano pierwotny wariant funkcji straty [68], [69], która przy wdrożeniu produkcyjnym okazała się na tyle efektywna, że pięć z ośmiu wdrożeń wykorzystuje bazowy wariant funkcji straty `mse_bce` z domyślnymi hiperparametrami, bez indywidualnego strojenia. Potwierdziło to obserwację z rozdziału 4, zgodnie z którą konfiguracja startowa stanowi skuteczny punkt wyjścia dla większości obiektów (pamiętając o siatce stanów).

5.4 Metryki modeli produkcyjnych

Zgodnie z metodyką opisaną w sekcji 3.5 oraz założeniami OOT modele produkcyjne poddano ewaluacji, na danych z miesiąca następującego po okresie treningowym. Wyniki przedstawione zostały w tabeli 5.2.

Tabela 5.2: Wskaźniki bezpieczeństwa modeli produkcyjnych (ocena OOT). N_W — liczba okien decyzyjnych, N_R — liczba powrotów obciążenia ($TP + FN$), N_O — liczba okien bez powrotu ($TN + FP$). Dla D3 ($T_s = 5$ s) oceniono dobę z tygodnia uczącego (sekcja 5.9).

Obiekt	N_W	N_R	N_O	S_{FN} [%]↓	S_{FP} [%]↓	S [%]↑	RMSE↓
A1	1688	176	1512	42,0	3,5	92,1	0,076
A3	2185	14	2171	42,9	0,5	99,2	0,444
B1	3961	1148	2813	**13,5**	6,3	89,8	0,048
C1	426	38	388	73,7	0,7	92,7	0,108
D1	2003	24	1979	100,0	0,0	98,8	0,062
D2	6983	574	6409	70,0	4,9	89,3	0,059
D3	532	387	145	22,2	0,2	83,6	0,169
D4	3303	216	3087	86,6	0,6	93,7	0,074

W związku z działaniem na rzeczywistych obiektach, gdzie mogą również pojawić się czynniki nieprzewidywalne mające wpływ na wyniki badań oraz ewaluację, należy wprowadzić trzy zastrzeżenia związane z zaproponowaną metodyką.

Pierwszym z nich jest pojawienie się prac instalacyjnych prowadzonych po stronie klienta dla obiektu A3. Prace obejmowały m.in. zmiany dotyczące instalacji elektrycznej, więc do momentu wdrożenia, nie było odpowiedniego zbioru do walidacji, a chciano wykorzystać moment, gdy układ mógł pracować ze wsparciem predykcyjnym. W związku z tym ewaluacja oparta jest na tym samym zbiorze, który był wykorzystany przy badaniach porównawczych, przez co nie stanowi niezależnej weryfikacji.

Drugim zastrzeżeniem, jest wynikający z procesu optymalizacji dla obiektu D3, pracującego z okresem próbkowania 5 s, zbiór ewaluacyjny obejmujący jedną dobę. Obiekt ten charakteryzuje intensywna, szybka zmienność cyklu produkcyjnego gdzie dłuższe okno okazało się nie być reprezentatywne.

Trzecim zastrzeżeniem jest różna dynamika układów, gdzie liczebność zdarzeń decyzyjnych N_R różni się pomiędzy obiektami o dwa rzędy wielkości (od 14 dla A3 do 1148 dla B1). W związku z tym interpretacja na podstawie samego wskaźnika S_{FN} byłaby niewiarygodna statystycznie. Należy więc wskaźnik rozpatrywać przy uwzględnieniu rzędu wielkości opisywanych zdarzeń.

Najniższy wskaźnik S_{FN} w całym zestawieniu uzyskał obiekt B1 (13,5%) przy odsetku decyzji nadmiarowo ostrożnych na poziomie $S_{FP} = 6{,}3\%$ oraz największej liczbie zdarzeń ryzykownych w całym zbiorze (1148), co stanowi doskonałą podstawę statystyczną. Podczas eksperymentów wariant `mse_bce` dla 2 min otrzymał wysokie $S_{FN} = 41{,}6\%$, które spadało wraz ze wzrostem horyzontu predykcji do $S_{FN} = 27{,}2\%$ dla 5 min. Jako reprezentant obiektów grupy B w studium ablacji, w niezależnym przebiegu treningowym oraz wykorzystaniu funkcji `industrial` (kara za wysokie FN,

parametry domyślne) pogorszył wariant domyślny wskaźnika z S_{FN} z 49,1% do 94,0%. Wariant samej kary za zdarzenia FN obniża wskaźnik do poziomu 40,7%. Za pogarszanie wskaźnika nie odpowiada asymetria kosztów, lecz składowa trendu ciśnienia, która dla tego obiektu przesunęła kryterium wyboru modelu w stronę rozwiązań nadmiernie zachowawczych.

Interesującą obserwacją jest zatem, że na produkcji uzyskano efekt odwrotny, gdzie najbardziej odpowiednią funkcją okazała się funkcja `industrial`. Doświadczalnie zatem stwierdzono, że funkcja okazała się efektywna po dobraniu jej parametrów do danego przypadku. Wspomniane wartości $S_{FN} = 13{,}5\%$ dla ewaluacji modelu produkcyjnego oraz 41,6% związane są z różnymi zbiorami danych, przez co nie można ich bezpośrednio porównać, w przeciwieństwie do studium ablacji operującego na jednym zestawie danych.

5.5 Nadzór operacyjny — wskaźnik R_{FN}

5.5.1 Definicja

Podczas wdrożenia niedostępne są dane z przyszłości (używane w metodzie OOT), jednak wymagany jest mechanizm, który może z wyprzedzeniem informować o tym, że podejmowane decyzje zaczynają być nieadekwatne, a moduł SI zaczyna się degradować. Zdecydowano się na aktywne podejście w rozumieniu taksonomii z literatury [89]. W związku z występowaniem dryftu koncepcyjnego opisanego w dedykowanej sekcji 2.7 oraz przyjętych założeń bezpieczeństwa opisanych w sekcji 5.2 zdecydowano się na rozwinięcie mechanizmu weryfikacji następczej przez wprowadzenie wskaźnika odzwierciedlającego jakość decyzji podejmowanych przez moduł SI (wskaźnik informuje o zadziałaniu zabezpieczenia regułowego):

$$R_{FN} = \frac{FN_{op}}{FN_{op} + N_D} \cdot 100\% \tag{5.1}$$

gdzie FN_{op} oznacza liczbę rekomendacji wyłączenia, po których sprężarka powróciła pod obciążenie w oknie weryfikacji, które jest skutkiem zadziałania reguły zabezpieczającej, a N_D — liczbę rekomendacji, dla których powrót nie nastąpił.

Okno weryfikacji obejmuje okres bezpośrednio po wydaniu rekomendacji i podczas prowadzonych badań zostało ustalone na poziomie powiązanym z 5 s okresowością odczytu danych przy użyciu protokołu S7 oraz parametrem definiującym liczbę próbek obserwacji polecenia. Przyjęto 6 próbek, co odpowiada okresowi 30 s. Tak ustalony poziom sprawdza zablokowanie polecenia wyłączenia maszyny przez warstwę zabezpieczeń. Powrót maszyny do obciążenia w obserwowanym czasie oznacza, że ciśnienie w

sieci osiągnęło poziom granicy zabezpieczającej stałość dostaw (np. dolna robocza granica ciśnienia sterownika nadrzędnego), a sterownik nadrzędny wymusił ponowne obciążenie sprężarki. Wskaźnik R_{FN} zauważa więc przypadki, w których deterministyczna, regułowa warstwa zabezpieczeń musiała skorygować decyzję warstwy predykcyjnej.

Zaletą wskaźnika jest jego jednoznaczność: nie wymaga interpretacji przebiegu procesu oraz odzwierciedla faktyczną interwencję układu regulacji. Należy zauważyć, że przez wykorzystanie określonego horyzontu, nie oddaje wszystkich sytuacji, gdy maszyna była potrzebna. Priorytetem było sprawdzenie, czy decyzja została podważona w zakresie horyzontu lub sterownik nadrzędny oczekiwał na włączenie maszyny. By zrealizować stałość dostaw sprężonego powietrza przy współpracy z modułem SI, ze względu na bezpieczeństwo procesu, założeniem projektowym było, by predykcja była jak najbardziej adekwatna przez czas rozprężania maszyny. Wówczas gdy dojdzie do wyłączenia, a nastąpi spadek ciśnienia zgodny z predykcją, to będzie gotowa do rozruchu. Czasem dynamika procesu pozwala na dłuższy horyzont czasowy niż czas rozprężania. Sytuacja powrotu do obciążenia po okresie predykcyjnym odzwierciedla charakter procesu, a nie pomyłkę systemu predykcyjnego. Z punktu widzenia diagnostycznej funkcji wskaźnika, zapis zdarzeń jednoznacznie informuje o pracy układu i umożliwia działania prewencyjne. Od strony operatora procesu narastanie wskaźnika R_{FN} świadczy o pogorszeniu jakości decyzji, która może wiązać się ze zmianami w charakterystyce zapotrzebowania.

Sytuacje, których wskaźnik nie uwzględnia to:

- utrzymywanie się ciśnienia tuż poniżej dolnej granicy roboczej, gdy sterownik oczekuje na możliwość obciążenia maszyny oraz ponownego rozruchu — taka sytuacja nie jest niebezpieczna procesowo, ale operator procesu uznałby ją za błędną,
- nagły spadek ciśnienia i załączenie kolejnej dostępnej maszyny w miejsce wyłączonej — poprzedzające zdarzenie wyłączenia zostanie zinterpretowane jako poprawne,
- spadek ciśnienia poniżej ustalonego niebezpiecznego poziomu — tutaj rozgraniczenie pomiędzy tymczasowymi spadkami ciśnienia w związku z większym zapotrzebowaniem wymagałoby rozróżnienia wielu przypadków, w niniejszej rozprawie oceniono wszystkie wykresy pracy układów po wdrożeniu i poza próbnym rozruchem obiektu D3, nie doszło do spadku ciśnienia procesowego.

Kolejną własnością kryterium R_{FN} jest jego prostota, przez co nie wymaga przeglądania wykresów ani tworzenia dedykowanych reguł, zapewnia też pełną porównywalność pomiędzy różnymi obiektami. W związku ze wspomnianymi ograniczeniami wskaźnik nie jest pełną miarą ryzyka, jest oszacowaniem umożliwiającym obserwację trendu w czasie, ale rzeczywista liczba decyzji nieprawidłowych będzie zaniżona.

5.5.2 Wyniki eksploatacyjne oraz wpływ warstwy zabezpieczeń

W tabeli 5.3 przedstawiono weryfikację 133 507 decyzji podjętych przez system sterowania (w tym 133 394 decyzji N_D bez powrotu obciążenia oraz 113 błędów FN_{op}). Wprowadzono również wskaźnik średniej liczby decyzji na dzień pracy systemu D/dz, co pozwala zauważyć różnicę w dynamice pomiędzy poszczególnymi układami. Kolejną istotną informacją jest liczba zablokowanych decyzji N_B co pozwala przeanalizować wpływ zabezpieczeń na pracę układu.

Tabela 5.3: Wyniki nadzoru operacyjnego. N_D — decyzje bez powrotu obciążenia, N_B — zablokowane decyzje (w nawiasie udział w sumie $N_B + N_D$), FN_{op} — błędne decyzje, D/dz — średnia liczba decyzji na dobę. S_{FN} z tabeli 5.2 liczono w oknie 2 min, a R_{FN} w oknie 30 s.

Obiekt	N_D	N_B	FN_{op}	R_{FN} [%]	D/dz	S_{FN} [%]
A1	34 925	394 (1,1%)	7	0,020	261	42,0
A3	9 433	0 (0,0%)	0	**0,000**	145	42,9
B1	2 050	444 (17,8%)	0	**0,000**	43	13,5
C1	2 412	0 (0,0%)	88	**3,520**	19	73,7
D1	22 149	104 (0,5%)	1	0,005	164	100,0
D2	28 395	7 154 (20,1%)	0	**0,000**	212	70,0
D3	5 514	340 (5,8%)	16	0,289	158	22,2
D4	28 516	3 (0,0%)	1	0,004	225	86,6
Łącznie	**133 394**	**8 439**	**113**	**0,085**	—	—

Powtórzony wskaźnik S_{FN} wyznaczony podczas uczenia (offline), przedstawiony porównawczo w zestawieniu ze wskaźnikiem nadzoru operacyjnego R_{FN} prowadzi do ciekawej obserwacji. Gwoli przypomnienia S_{FN} ocenia jakość predykcji względem przebiegu, który jest już dostępny w momencie oceny na zasadach OOT. R_{FN} nie dysponuje takimi informacjami i dokonuje oceny na rzeczywistym układzie o nieznanym przebiegu zapotrzebowania, przez co odzwierciedla rzeczywiste konsekwencje decyzji. Zauważono, że wysoka wartość S_{FN} nie przekłada się jednoznacznie na ryzyko operacyjne. Dla obiektu D1 ($S_{FN} = 100\%$), sieć neuronowa nie miała wystarczającej ilości informacji, która pozwoliłaby na nauczenie się wszystkich zależności dotyczących powrotu do obciążenia. W związku z tym żaden z 24 powrotów obciążenia w zbiorze ewaluacyjnym nie zostało rozpoznane natomiast podczas eksploatacji wystąpił tylko jeden błąd FN_{op} na podjęte 22 150 decyzji. Ta zależność nie jest skutkiem użycia reguł zabezpieczających działanie układu, które zadziałały 104 razy, co stanowi tylko 0,5% podjętych decyzji. Drugim obiektem, gdzie występuje podobna relacja to obiekt D4 (86,6% wobec 0,004% przy tylko 3 blokadach i 28 516 decyzjach).

Rozbieżność pomiędzy wskaźnikami wynika z odmiennej definicji ocenianego zdarzenia. Ewaluacja offline sprawdza każdy powrót obciążenia pojawiający się w oknie decyzyjnym, jednak nie ma świadomości o tym, czy pociąga on za sobą skutki procesowe. Pobór powietrza pozostaje w jakimś stopniu nieidentyfikowalny, nie wiemy, czy w danym momencie, gdy nastąpi błędna decyzja, zakład produkcyjny postanowi np. załączać kolejne urządzenia. Weryfikacja następcza natomiast rejestruje wyłącznie powroty natychmiastowe, wymuszone spadkiem ciśnienia poniżej granicy roboczej. To jest rzeczywiste zachowanie obiektu, które nie jest możliwe do zbadania i modelowania w ramach proponowanej metodyki. Dla obiektu D1 oznacza to, że model istotnie nie rozpoznał żadnego z dwudziestu czterech powrotów obciążenia w zbiorze ewaluacyjnym, jednak nie jest wiadome, czy późniejsze powroty zagrażały ciągłości dostaw. Jeżeli pozostała maszyna rezerwowa była dostępna, można założyć, że układ dysponował wystarczającym zapasem wydajności. Wysoka wartość S_{FN} wskazuje zatem na ograniczoną zdolność modelu do przewidywania momentu powrotu, ale model nie zna konsekwencji procesowych. Niska wartość R_{FN} niesie już w sobie informację o dynamice układu, w przypadku tego konkretnego obiektu błędne decyzje nie miały konsekwencji procesowych.

Podsumowując: ocena ryzyka wdrożenia nie może opierać się wyłącznie na metrykach wyznaczonych offline, ponieważ nie są one w stanie przewidzieć rzeczywistego zachowania sieci odbiorczej, ani działania warstwy zabezpieczeń (której parametry muszą być dobrane empirycznie). W tym miejscu wypada jeszcze podkreślić, że przez pierwsze tygodnie działania układów, operowano bez warstwy zabezpieczeń, nie powodując przy tym istotnego spadku ciśnienia. Jednak zauważano nieznaczne spadki ciśnienia i podejmowanie decyzji przy dolnej granicy ciśnienia, przez co zdecydowano się na jej wprowadzenie. Udział wstrzymanych rekomendacji B jest samodzielnym wskaźnikiem diagnostycznym informującym o tym jak bardzo moduł SI podczas pracy bazuje na zabezpieczeniach. Przy dobrze przeprowadzonej identyfikacji układu będzie niewielka. Dla obiektów A3, C1, D4 osiągnięto wartości zerowe, dla A1, D1 bliskie zeru. Oznacza to zgodność rekomendacji modelu z ograniczeniami eksploatacyjnymi układu. Wypada również podkreślić w związku z inną dynamiką obiektu, obiekt B1 oraz D2 ma ten wskaźnik na poziomie około dwudziestu procent, co oznacza, że co piąta decyzja jest blokowana. W ich wypadku zdecydowano się na taki poziom blokady, by zapewnić wyższy poziom ciśnienia w instalacji odbiorczej, dzięki czemu wdrożenie sieci neuronowej od strony operacyjnej praktycznie przeszło niezauważone.

Bardzo ciekawym wskaźnikiem jest średni wskaźnik gęstości decyzji D/dz, który poza tym, że niesie w sobie informację o potencjale optymalizacyjnym i charakterze pracy obiektu jest istotny przy rozpatrywaniu wskaźnika R_{FN}. Na poparcie tej tezy można przytoczyć obiekt C1, który uzyskuje najwyższą wartość w zbiorze (3,520%), lecz przy 19 decyzjach dziennie, odpowiada to jednemu błędowi na jeden i pół dnia

pracy. Dla tego obiektu nie odnotowano ani jednej wstrzymanej rekomendacji. Charakter tego obiektu powoduje, że moduł SI proponuje wyłączenia rzadko, lecz gdy pojawia się rekomendacja to dotyczy przypadków granicznych. Dla obiektu D4 wartość 0,004% przy 225 decyzjach dziennie oznacza jeden błąd w ciągu czterech miesięcy eksploatacji. Przytoczone przypadki są jakościowo odmienne, mimo że oba obiekty pracują poprawnie.

5.6 Oszczędność energii elektrycznej

5.6.1 Potencjał oszczędnościowy

Podczas prowadzenia badań nie dysponowano danymi ani z liczników energii elektrycznej, ani z analizatorów parametrów sieci, które miarodajnie pokazują wypracowane oszczędności. Zgodnie z przytoczonymi we wstępie normami takie wskaźniki są niezbędne ze strony zakładu produkcyjnego dla wykazania efektów wdrożonego systemu, natomiast w przypadku sprężarkowni często nie są dostępne. Dlatego założenia proponowanej metody, zakładały ich niedostępność. Przyjęto, że maszyny w rozpatrywanym okresie pracują prawidłowo i nie ma sytuacji, które mogłyby zweryfikować tylko liczniki energii elektrycznej (np. maszyna ma sklejoną cewkę zaworu sprężania, przez co tłoczy, ale w praktyce nie dochodzi do obciążenia).

W takim wypadku efekt wdrożenia został oceniony na podstawie procentowego udziału biegu jałowego podczas sumarycznego czasu pracy sprężarek na bazie przepracowanych roboczogodzin poszczególnych maszyn. Zgodnie z zaleceniami ISO 11011 [27] spisanie tych wartości jest miarodajnym wyznacznikiem potencjału oszczędnościowego. Od strony praktycznej, każda godzina przeznaczona na bieg jałowy, oznacza stratę energii elektrycznej, która nie jest przeznaczona na sprężanie. Jednak tutaj należy zastrzec, że często konfiguracja układu będzie wymagała pozostawienia maszyn na biegu jałowym w oczekiwaniu na spadek ciśnienia, odwdzięczając się za to stabilniejszym ciśnieniem (co zostanie pokazane w przypadku obiektu D3).

Należy wspomnieć, że przy dysponowaniu licznikami energii elektrycznej oraz pomiarami przepływu możliwe jest utworzenie wskaźnika efektywności energetycznej dla każdej maszyny sprężarkowni [27], a następnie w ramach eksperymentów na bazie aktualnego profilu zapotrzebowania zakładu (np. produkcja określonego typu) ustalenie takiej sekwencji sprężarek, która będzie najkorzystniejsza energetycznie. Tutaj należy podkreślić, że niezbędne jest właśnie dysponowanie rzeczywistymi, a nie teoretycznymi danymi. Niestety, w ramach niniejszej rozprawy, nie uzyskano takowych, ale warto wspomnieć o tak intuicyjnej metodzie.

Wracając do wdrożenia, należy zauważyć kolejne ograniczenie. Zestawienie okresów przed i po wdrożeniu pokazuje, że profil produkcji zmienia się w czasie, a zmiany te

są niezależne od działania systemu. W większości przypadków miesiąc danych jest wystarczający do wytrenowania sieci, jednak zapotrzebowanie miesiąc do miesiąca może się znacząco różnić. By ograniczyć wpływ tego czynnika zmienności, jako okres bazowy przyjęto trzy pełne miesiące poprzedzające wdrożenie, a przy interpretacji wyników sprawdzano łączny czas pracy układu. Ze względu na poufność danych produkcyjnych klientów w zestawieniu podano wyłącznie wartości procentowe.

5.6.2 Wynik 806 dni wdrożenia

W tabeli 5.4 przedstawiono procentowy udział biegu jałowego w ramach prowadzonego wdrożenia. Na wstępie wypada zaznaczyć, że co zdarza się w przypadku rzeczywistych układów, wystąpiły sytuacje niespodziewane, a w kilku wypadkach statystyczna ocena jest niejednoznaczna i wymaga indywidualnego rozpatrzenia. Kolumny oznaczają kolejne miesiące eksploatacji względem uruchomienia systemu: M_{-3} do M_{-1} to okres bazowy, M_0 oznacza miesiąc uruchomienia, M_{+1} do M_{+3} to kolejne miesiące pracy. Symbol „X” oznacza okres, w którym system nie pracował lub jego praca została zakłócona przez czynniki zewnętrzne.

Tabela 5.4: Procentowy udział biegu jałowego w czasie pracy sprężarek.

Obiekt	M_{-3}	M_{-2}	M_{-1}	M_0	M_{+1}	M_{+2}	M_{+3}
A1	22,8	15,2	16,1	**7,7**	15,4	12,5	15,0
A3[c]	19,6	17,1	17,3	**6,7**	**6,2**	X[a]	X[a]
B1	47,6	50,3	34,6	33,9	35,8[b]	X	X
C1	1,9	2,2	1,8	1,6	1,9	3,8	2,2
D1	32,7	24,4	34,5	35,2	38,1	35,5	33,4
D2	63,5	69,9	67,6	48,0	43,9	**33,5**	41,8
D3	34,1	33,6	36,6	**27,6**	28,5[b]	X	X
D4	43,3	42,4	43,6	15,7	**7,5**	**7,9**	**6,6**

[a] zmiany instalacyjne i brak dostępu do maszyn,

[b] zakończenie pracy systemu w trakcie miesiąca,

[c] system uruchomiono w ostatnich dniach M_{-1}, dlatego M_0 jest pierwszym pełnym miesiącem eksploatacji

Tabela 5.5 przedstawia procentową poprawę oszczędności energii na bazie roboczogodzin. Okres bazowy stanowi średnia z miesięcy od M_{-3} do M_{-1}. Okres po wdrożeniu to średnia z miesięcy, w których system pracował bez zakłóceń rozpocząwszy od miesiąca M_{+1}. Zrezygnowano z uwzględniania w średniej miesiąca M_0 ze względu na różny

dzień, w którym wykonywano wdrożenie, przez co wyniki byłyby zaburzone. Tabela zawiera również kolumnę nadzoru operacyjnego R_{FN}, który nie jest uśredniony a odnosi się do całego okresu wdrożenia.

Tabela 5.5: Udział biegu jałowego w okresie bazowym i po wdrożeniu oraz jego względna zmiana.

Obiekt	Baza [%]	Po wdrożeniu [%]	Zmiana [%]	R_{FN} [%]
D4	43,1	7,3	**−83**	0,004
A3[a]	18,0	6,2	**−66**	**0,000**
D2	67,0	39,7	**−41**	**0,000**
A1	18,0	14,3	−21	0,020
B1[b]	44,2	35,8	−19	**0,000**
D3[b]	34,8	28,5	−18	0,289
D1	30,5	35,7	+17	0,005
C1	2,0	2,6	+34	3,520

Baza — średnia z miesięcy M_{-3} do M_{-1}; okres po wdrożeniu — średnia z miesięcy M_{+1} do M_{+3}, z pominięciem miesiąca rozruchu M_0.

[a] okres po wdrożeniu obejmuje wyłącznie miesiąc M_{+1} — w kolejnych miesiącach brak dostępu do maszyn w związku z pracami instalacyjnymi (sekcja 5.7.1).

[b] okres po wdrożeniu obejmuje wyłącznie miesiąc M_{+1}, w trakcie którego zakończono pracę systemu; wartość nie reprezentuje pełnego miesiąca eksploatacji.

Wcześniej wspomniano o subiektywności oceny statystycznej, która wymaga indywidualnego rozpatrzenia. Takim przypadkiem jest obiekt D1, gdzie użyty wskaźnik nie oddaje rzeczywistego działania systemu. Tutaj dokonano szczegółowej analizy, przy której okazało się, że większość biegu jałowego generuje w tym układzie maszyna falownikowa. Jej pracę charakteryzuje kilkunastokrotne wyłączanie się w ciągu godziny przy braku zapotrzebowania. To zachowanie wynika z charakterystyki układu, a nie z decyzji warstwy predykcyjnej. Analiza trendów oraz wyniki nadzoru operacyjnego (jeden błąd na 22 150 decyzji) oznaczają prawidłowe działanie systemu. Na wykresach pracy maszyn uzyskane oszczędności energii są widoczne (szczególnie w zestawieniu z okresem bez wsparcia modułem SI). Przypadek ten stanowi przykład ograniczenia statystycznego wskaźnika opartego na licznikach motogodzin. Przy hipotetycznym wyłączeniu wsparcia predykcyjnego, po analizie przebiegów czasowych zauważono, że uzyskany procent biegu jałowego byłby znacznie wyższy niż ten będący nowym punktem odniesienia.

Najlepiej wskaźnik wypada dla obiektu D4. W jego wypadku nastąpiła redukcja udziału biegu jałowego z 43,1% do 7,3%, czyli względem stanu bazowego poprawa

wyniosła 83%. Wynik został utrzymany przez kolejne trzy miesiące. Co istotne łączny czas pracy układu pozostał porównywalny z okresem bazowym i różnica między nimi wyniosła 1,3%, co wyklucza wyjaśnienie zmianą wolumenu produkcji. Należy również zauważyć zbieżność z wynikami nadzoru operacyjnego, obiekt ten odnotował jeden błąd operacyjny na 28 516 decyzji.

Obiekt A3, który stanowił inspirację do rozpoczęcia badań, uzyskał poprawę na poziomie 66%, jednocześnie nie popełnił żadnych błędnych decyzji i pokazał duże pole do potencjalnych oszczędności. Wynik taki też wynika z charakterystyki pracy układu, w której zaproponowana sieć doskonale się sprawdza. Działanie układu nie było jednak kontynuowane, ze względu na prace instalacyjne prowadzące do utraty dostępu do maszyn co opisano w sekcji 5.7.1.

Dla obiektu D2 redukcja wyniosła 41%. Zestawiając miesiące o zbliżonym obciążeniu przyjmując dla M_{-3} 100% roboczogodzin jako punkt odniesienia (63,5% biegu jałowego) dla miesiąca M_{+2} otrzymamy 88% przepracowanych roboczogodzin (udział biegu jałowego spadł do 33,5%).

Dwa obiekty nie wykazały statystycznej poprawy w tej mierze, a przyczyny są w każdym przypadku odmienne i wymagają odrębnego omówienia.

Najgorszy wynik uzyskuje obiekt C1 z udziałem biegu jałowego przed wdrożeniem na niskim poziomie 1,8–2,2%. Jest to nowoczesny, zoptymalizowany obiekt z zaawansowanymi konstrukcyjnie maszynami doskonale adaptującymi się do zmiennego poboru (należy wspomnieć, że jest to również zasługa sterowania nadrzędnego). Potencjał optymalizacyjny był zatem niewielki, prawie bliski zeru. Zjawisko to omówiono szerzej w sekcji 5.7.4.

Przy przygotowaniu obiektu B1 do wdrożenia zauważono dalekie od optymalnego działanie maszyny VFD. W związku z tym najpierw przystąpiono do optymalizacji układu, uzyskując poprawę w miesiącu M_{-1}. Osiągnięty poziom utrzymywał się przez cały okres pracy warstwy predykcyjnej. W kolejnych dwóch miesiącach, które nie zostały przedstawione w tabeli zbiorczej, po zakończeniu jej pracy udział biegu jałowego wzrastał systematycznie, powracając do wartości sprzed okresu optymalizacji (wyniósł odpowiednio 38,9% i 45,8%). Obserwacja ta stanowi przesłankę wskazującą na udział systemu w utrzymaniu osiągniętego poziomu. W tym wypadku predykcja została wyłączona w związku ze zmianą profilu produkcji i potrzebą nauczenia się nowej charakterystyki. W związku z tym statystyka przedstawiona nie może być traktowana jako dowód. Dane z okresu po zakończeniu pracy systemu nie były przedmiotem systematycznej analizy.

5.7 Zauważone mechanizmy degradacji

Działając na rzeczywistych obiektach przemysłowych, podczas wielomiesięcznej eksploatacji zauważono cztery jakościowo odmienne mechanizmy, prowadzące do utraty efektywności systemu. Identyczna wartość wskaźnika R_{FN} może w różnych przypadkach wiązać się z innymi przyczynami, a rozstrzygnięciem będzie *przebieg* wskaźnika w czasie, a niekoniecznie jego poziom.

5.7.1 Czynniki nieprzewidywalne — obiekt A3

Obiekt A3 był pierwszym obiektem, gdzie zauważono, że można uzyskać największe korzyści energetyczne. Potwierdziły to dwa miesiące eksploatacji, gdy obiekt nie podjął ani jednej błędnej decyzji. Owszem, wskaźnik S_{FN} nie jest idealny, ponieważ w zbiorze uczącym nie wystąpiły sytuacje, które pozwoliłyby nauczyć się dokładnie roli każdej maszyny natomiast zauważalny jest zerowy wskaźnik R_{FN}. Z czasem na obiekcie zostały przeprowadzane prace instalacyjne, które związane zostały m.in. z rozłączeniem linii transmisyjnej, przełączeniem części maszyn w tryb lokalny, a zatem możliwości jakiegokolwiek wsparcia predykcyjnego zostały w pełni ograniczone. Na etapie metodyki podkreślono, że warunkiem rozpoczęcia prac jest prawidłowa praca układu w okresie treningowym, co nie mogło zostać spełnione.

Odnosząc się jeszcze do samego wskaźnika R_{FN} pełni on funkcję alarmującą, gdy jego gwałtowny wzrost będzie prawdopodobnie wskazywał potrzebę interwencji serwisowej, a nie ponownego trenowania modelu. Przyczyna wzrostu musi zostać sprawdzona indywidualnie. Możliwe są dwie sytuacje. Pierwsza, związana z wyłączeniem danej maszyny, przy czym model nie był trenowany na sytuacjach awaryjnych, więc nie zna nowej dynamiki układu, a jego decyzje będą odzwierciedlać sytuacje, gdy wszystkie maszyny pracowały. Druga sytuacja to znacząca zmiana profilu produkcji, która wymaga dotrenowania sieci neuronowej.

5.7.2 Dryft koncepcyjny — zmiana profilu produkcji — obiekt D3

Obiekt D3 omówiony został szczegółowo w studium przypadku 5.9. Obiekt pracował poprawnie przez pierwsze dwa tygodnie eksploatacji (przy trzech niegroźnych błędach operacyjnych), jednak prawdopodobnie w związku ze zmianą profilu produkcji, wskaźnik R_{FN} zaczął narastać (tabela 5.6) co oznacza pojawienie się dryftu inkrementalnego [89].

Tabela 5.6: Tygodniowy przebieg wskaźnika R_{FN} dla obiektu D3 w pierwszych czterech pełnych tygodniach pracy. W pozostałych dniach nie wystąpiły błędy FN_{op}.

Tydzień	Decyzje	FN_{op}	R_{FN} [%]
1	1231	0	0,00
2	1403	3	0,21
3	1412	5	0,35
4	1187	8	0,67

Moduł SI wspierał system przez 35 dni i nie doprowadził do zatrzymania produkcji. Zauważalny stał się systematyczny wzrost wskaźnika R_{FN} w kolejnych tygodniach pracy, z wyraźnym przyspieszeniem w ostatnim tygodniu, któremu towarzyszył spadek liczby wydawanych decyzji. Przyczynę takiego zachowania można znaleźć po sprawdzeniu rozkładu godzin obciążenia obiektu (aczkolwiek przy analizie trendów również zauważalna jest intensyfikacja pracy maszyn). Całkowity czas pracy sprężarek narastał w stosunku do miesiąca uruchomienia M_0 o kolejno 24%, 32% oraz 36%. W tym okresie udział sprężarki SP1 w łącznym czasie pracy sprężarkowni wzrósł z 31% w miesiącu treningowym aż do 46% w kolejnym. Dynamika układu zmieniła się na tyle, że wytrenowany model nie odpowiadał już aktualnemu procesowi. Wizualnie w tygodniach poprzedzających wyłączenie również było widoczne rozszerzenie widełek regulacji ciśnienia. O fakcie zmiany profilu produkcji klient również poinformował i poprosił o taki dobór nastaw, który zapewni stabilność ciśnienia, nawet jeżeli odbędzie się to kosztem wysokiego poziomu biegu jałowego. Warto tutaj podkreślić, że wskaźnik R_{FN} doskonale oddał to, co było widoczne na wykresach i pełnił funkcję wczesnego ostrzegania. Ostatecznie wsparcie modułu SI zostało wyłączone w 13 dniu drugiego miesiąca rozruchu, a liczba blokad bezpieczeństwa na poziomie 5,8% świadczy o nadążaniu za dynamiką procesu jednak bez sprostania nowym wymaganiom. Szczegółowo przypadek został omówiony w osobnej sekcji studium przypadku 5.9.

Jedną z największych obaw przy przystępowaniu do wdrożenia, była właśnie długofalowa eksploatacja systemu predykcyjnego na obiekcie i zmiana profilu produkcji. Właśnie ten przypadek stanowi potwierdzenie znaczenia dryftu koncepcyjnego [55], [56] zauważonego przy przeglądzie literatury w rozdziale 2.

5.7.3 Zanikanie efektywności mimo dotrenowania — obiekt A1

W przypadku obiektu A1 nastąpiła korelacja dwóch czynników: dryftu koncepcyjnego inkrementalnego [88] oraz zanikania efektywności predykcji mimo dotrenowywania sieci. Miesiąc wdrożenia otrzymał najlepszy wynik na poziomie 7,7% w porównaniu z

bazowym 18,0%. Następnie w kolejnych miesiącach procent biegu jałowego utrzymywał się na poziomie odpowiednio 15,4%, 12,5% oraz 15,0% przy 1,1% zablokowanych decyzji. Na obiekcie wystąpiły również zdarzenia związane z awariami maszyn, przez co przez jakiś czas w układzie brakowało jednej z nich a charakter pracy układu zmienił się. Mimo gorszych efektów oraz widocznej większej ilości możliwych sytuacji oszczędności, nie można powiedzieć, że nastąpiło całkowite wygaszenie działania nauczonej sieci.

Przy analizie rozkładu godzin poszczególnych maszyn okazało się, że o ile liczba decyzji podejmowanych miesięcznie jest podobna, to zmienia się sekwencja pracy sprężarek. W takich momentach działanie układu odbiegało od działania w okresie zbioru treningowego. Zauważonym efektem było to, że w przypadku takich sytuacji nie pojawiały się rekomendacje wyłączenia maszyn, przez co maszyna była niepotrzebnie utrzymywana na biegu jałowym. Efekt zauważono stosunkowo szybko i na obiekcie wykonano dotrenowanie w cyklach co 12 godzin, jednak nie udało się w ten sposób uzyskać znaczącej poprawy. Być może rozdzielenie trenowania sieci na okres tygodniowy oraz weekendowy (na obiekcie występuje zauważalna cykliczność produkcji) doprowadziłoby do powrotu do wyników miesiąca M_0 natomiast 12-godzinne dotrenowanie nie przynosiło spodziewanych efektów.

5.7.4 Brak potencjału oszczędnościowego — obiekt C1

Obiekty grupy C, to obiekty zbudowane wyłącznie z maszyn z przetwornicą częstotliwości, często również wyposażonych w silniki z magnesami trwałymi, które zapewniają szerokie pole adaptacji. Wystarczy wspomnieć, że maszyna z przetwornicą częstotliwości oferuje zakres pracy pod obciążeniem o rozpiętości np. 35 Hz, podczas gdy maszyna z magnesami trwałymi może mieć do dyspozycji aż 95 Hz. Wdrożenie obiektu C1 miało za zadanie sprawdzenie, czy jest w jego wypadku jakiekolwiek miejsce na optymalizację.

Już stan wejściowy przed przystąpieniem do wsparcia modułem AI dawał sumaryczny bieg luzem na poziomie od 1,8 do 2,2%, przy czym po wdrożeniu i pominięciu ekstremum (3,8%) wyniki znajdowały się na poziomie od 1,6 do 2,2%. Przy liczbie około 19 decyzji na dzień, pole do optymalizacji jest niewielkie. Można by powiedzieć, że w tym przypadku mamy do czynienia z nietypowym, skrajnym zachowaniem układu. Również co istotne, przy takim wolumenie, mimo braku zablokowanych zdarzeń, często następowały niebezpieczne zdarzenia FN_{op}. Jest to również zauważalne dzięki wskaźnikowi R_{FN}, który jest najwyższy ze wszystkich wdrożonych obiektów (3,520%). Wniosek jest jednoznaczny — układ pracuje optymalnie wykorzystując w pełni swoje możliwości. Takie działanie systemu na etapie treningu, nie daje modelowi odpowiedniej przestrzeni uczącej, by wiedzieć, jakie jest znaczenie każdej maszyny na sprężarkowni z

wykorzystaniem proponowanej metodyki.

Mając udział procentu biegu jałowego na poziomie poniżej 2,5% wdrożenie systemu predykcyjnego mija się z celem. W taki wypadku należy sprawdzić moc znamionową maszyn wchodzących w skład systemu i wyliczyć, jakie będą potencjalne oszczędności. W wypadku takich układów, pole do optymalizacji dotyczyłoby odpowiedniego nadążania za charakterystyką ciśnienia, co pośrednio również minimalizuje załączenia kolejnych maszyn. Ten aspekt powinien zostać uwzględniony przy wyznaczaniu obiektów do wdrożenia. W przeglądzie literatury wspomniano o tym, że można by tutaj radzić sobie przez wprowadzenie dodatkowej wagi dla maszyn, których zachowania nie znamy, jednak nie było to przedmiotem eksperymentu w niniejszej rozprawie.

5.8 Rola dotrenowania modelu podczas eksploatacji

Przystępując do badań zakładano, że gdy dojdzie do wdrożenia modułu SI na produkcji, to cykliczne dotrenowanie [23] powinno doskonale nadążać za zmianami charakterystyki układu i minimalizować dryft koncepcyjny. W związku z tym moduł SI został wyposażony w mechanizm cyklicznego dotrenowywania na bazie ustalonego horyzontu. Trening uruchamiany jest z zadaną częstotliwością, która została dobrana doświadczalnie podczas pracy modelu, jednak nie przeprowadzono pełnych testów, które mogłyby pokazać neutralną skalę porównawczą. Mimo to nasuwa się kilka interesujących wniosków wdrożeniowych. Wyprzedzając wnioski tej sekcji należy zaznaczyć, że opisane zostały rady czego nie robić oraz na co należy zwrócić uwagę przy wdrożeniu na bazie własnych błędów. Parametry dotrenowania zostały przedstawione w tabeli 5.7. Kolumna „okno” oznacza zakres danych przekazywanych w pojedynczym cyklu, FN_{op} to liczba wydanych błędnych decyzji, wprowadzono również kolumny opisujące uzyskane oszczędności energii: „śr.” to średnia zmiana udziału biegu jałowego w całym okresie pracy systemu w stosunku do okresu bazowego, a „koniec” to efekt oszczędnościowy uzyskany w ostatnim miesiącu eksploatacji.

Tabela 5.7: Konfiguracja mechanizmu dotrenowania, nadzór operacyjny oraz oszczędności energii (średnia liczona z pominięciem miesiąca M_0).

						Zmiana [%]	
Obiekt	Harmonogram	Okno	Epoki	FN_{op}	R_{FN} [%]	śr.	koniec
A3	wyłączone	—	—	0	**0,000**	−66	−66
B1	wyłączone	—	—	0	**0,000**	−19	−19
D2	wyłączone	—	—	0	**0,000**	−41	−38
D3	wyłączone	—	—	16	0,289	−18	−18
A1	2 razy na dobę	12 h	20	7	0,020	−21	−17
C1	co godzinę	24 h	20	88	**3,520**	+34	+12
D4	co godzinę	24 h	20	1	0,004	−83	−85
D1	raz w tygodniu	21 dni	50	1	0,005	+17	+9

Pod względem eksperymentalnym analizowane dotrenowanie miało charakter reaktywny i nie było z góry zaplanowane co do przebiegu. Na bazie bieżących obserwacji działania układu, zauważono, że następuje pogarszanie wskaźników i spodziewano się, że dostarczenie modelowi informacji o nowej charakterystyce pracy przełoży się na lepsze działanie co jest zgodne z przeglądem literatury w tym zakresie [89]. Przy czym wprowadzony mechanizm zabezpieczający, dopuszcza do modelu informacje o tym, gdy było potrzebne działanie korygujące. Ogranicza to sytuacje, gdy model uczyłby się na własnych, błędnych decyzjach, a zatem dopuszcza pewne informacje o nowym charakterze pracy układu.

Zdano się też na mechanizm uczenia głębokiego z założeniem, że nie będzie potrzeby dodatkowego nadzoru, co okazało się błędne. Można wyróżnić trzy założenia, które okazały się niezalecane, dotyczące wprowadzonego podejścia:

- zbyt krótki horyzont douczania — douczanie na bazie dwunastu godzin przy trenowaniu modelu na okresie trzytygodniowym nie niesie w sobie informacji o odpowiedniej wadze zdarzeń, pojawiające się zdarzenie obciążenia, jeśli będzie pojedyncze przy założonej metodyki, nie będzie miało wystarczającego znaczenia,
- brak weryfikacji wystąpienia istotnych dla uczenia zdarzeń — może wystąpić sytuacja alarmowa, która spowoduje, że model będzie uczył się nieprawidłowej sytuacji, bez weryfikacji stanu, taki model znajdzie się na produkcji,
- brak kryterium dopuszczenia modelu do ruchu — obecnie model jest zapisywany automatycznie, w skrajnych przypadkach weryfikacja funkcji straty może znacznie pogorszyć istniejący model.

Mechanizm ten został wprowadzony bez wspomnianej weryfikacji, ponieważ założono, że trzeba jak najszybciej reagować na zmiany, jakie pojawiają się w układzie, a dopiero przyszłość zweryfikuje czy były słuszne.

Dla czterech obiektów nie zdecydowano się na wprowadzanie dotrenowania. Testy offline skutkowały pogorszeniem predykcji modelu więc pozostano przy modelu bazowym. Tylko dla jednego z nich, obiektu D3, zaobserwowano pogarszanie się wskaźnika R_{FN} przy utrzymaniu jakości predykcji. Uzyskane oszczędności energii utrzymują się w ich wypadku na zbliżonym poziomie (różnica pomiędzy średnią a wartością końcową oszczędności jest mniejsza niż 4%). Tutaj jeszcze należy przypomnieć, że obiekty A3, B1 oraz D3 pracowały krócej od pozostałych modeli, przez co ich końcowy miesiąc pracy jest pierwszym pełnym miesiącem od uruchomienia.

Dla obiektów C1 oraz D4 założeniem było, że dotrenowanie będzie pomocne w reagowaniu na bardzo zmienny pobór, tak by jak najszybciej zaadaptować się do zmian, czemu miał odpowiadać horyzont uczący 24 h przy dotrenowaniu co godzinę. Dla obiektu A1 dobrano wolniejsze dotrenowanie, dwa razy na dobę, na horyzoncie 12 h. Pojawiły się przy tym nieprawidłowe decyzje fałszywie negatywne: $FN_{op} = 7$ dla A1, $FN_{op} = 88$ dla C1 oraz $FN_{op} = 1$ dla D4, co po odniesieniu do liczby decyzji daje R_{FN} na poziomie odpowiednio 0,020%, 3,520% i 0,004%. Osiągnięty efekt redukcji biegu jałowego ze wszystkich obiektów jest najlepszy dla D4 i najgorszy dla C1. Tak więc dla nich przy dokładnie tych samych parametrach dotrenowywania uzyskano przeciwny rezultat zarówno dla jakości decyzji jak i oszczędności energii. Przyczyną jest odmienna charakterystyka obiektu C1, który, jak wykazano w sekcji 5.7.4, nie ma wspomnianego potencjału oszczędnościowego.

Regularne dotrenowywanie, nie pomogło też wystarczająco obiektowi A1, którego przypadek omówiono w podpunkcie 5.7.3. Obiekt cechuje stosunkowo duża liczba zdarzeń niezależnych, które powodują co jakiś czas wyłączenie danej maszyny z systemu sterowania w związku z potrzebną interwencją serwisu. Ostatecznie również zdecydowano się, już po okresie testowym, na wymianę jednej z maszyn na maszynę VFD przez co w przyszłości obiekt będzie obiektem typu B. W miesiącu rozruchu uzyskane początkowo znaczne oszczędności energii (-57%) zmalały w ostatnim miesiącu eksploatacji do poziomu -17%. Regularne dotrenowywanie nie przełożyło się na zatrzymanie tego procesu, ale też nieznany jest możliwy potencjał optymalizacyjny, nie spowodowano również spadków ciśnienia w całym rozpatrywanym okresie. Niestety przyjmowanie stanu niepełnej funkcjonalności jako stanu normalnego systemu sterowania zaowocowało właśnie utratą możliwych oszczędności. W obiekcie A1 zeszły się właśnie negatywne skutki wspomnianych wcześniej trzech błędnych założeń dotrenowywania.

Testy dotrenowywania zakończono na obiekcie D1 5.6, który zmienił znacząco charakterystykę pracy podczas badania. W jego wypadku zdecydowano się na dotrenowanie analogiczne jak w zbiorze uczącym wykonywane tylko raz w tygodniu przy

liczbie epok sięgającej pięćdziesięciu. Takie zachowanie poniekąd adaptuje model do nowych warunków. Co ciekawe, obiekt charakteryzuje tylko jedno zdarzenie fałszywie negatywne przez co stabilność dostaw sprężonego powietrza została utrzymana. Gwoli przypomnienia w tym przypadku większość biegu jałowego była skutkiem działania maszyny falownikowej, która również na minimalnych obrotach, okazała się zbyt wydajna, by znaleźć punkt pracy. W przypadku takiego doboru maszyn również nie było znacznego potencjału do optymalizacji.

Zdobyte przy dotrenowywaniu doświadczenie pozwala na przygotowanie wytycznych dotyczących produkcyjnego utrzymania modeli, co opisano w rozdziale 7. Biorąc pod uwagę, że wskaźnik R_{FN} jest bardzo dobrym sygnalistą pogarszania sytuacji to on powinien wywoływać zdarzenie dotrenowywania zgodnie z przyjętymi założenia uczenia. Harmonogramowanie zdarzenia dotrenowywania w tym wypadku wydaje się nieprawidłowe.

5.9 Studium przypadku — obiekt D3

Zdecydowano się na wybór do studium przypadku obiektu D3 ze względu na przejście w ciągu zaledwie 35 dni przez cały cykl życia wdrażanego systemu. Inspiracją było również to, że przy pierwszej próbie wdrożenia, obiekt został wyłączony w ciągu pierwszych dziesięciu minut, gdy podjął dwie niebezpieczne decyzje skutkujące spadkiem ciśnienia. Od razu wówczas pojawiła się ciekawość każąca zbadać przyczynę niepowodzenia. Doświadczenie procesowe wskazało na potrzebę zwiększenia częstotliwości i odejścia od bazowych 20 s okresu próbkowania, na jakich dotychczas pracowały modele. W trakcie pracy zmienił się charakter pracy układu. Przeprowadzono 257 eksperymentów w różnych konfiguracjach, by znaleźć najbardziej adekwatne rozwiązanie. Kolejne sekcje przedstawią wdrożenie opisanej w rozdziale 3 metodyki z opisaniem zmian, które umożliwiły praktyczne wdrożenie.

5.9.1 Konfiguracja startowa

Obiekt D3 należy do grupy obiektów A czyli nie posiadających maszyny falownikowej i będących obecnie archaiczną, jednak nadal spotykaną, konfiguracją sprężarkowni. Składa się tylko z dwóch maszyn, które są ciągle potrzebne i co istotne nie ma trzeciej maszyny rezerwowej. Jedna maszyna nie jest w stanie zapewnić wystarczającego ciśnienia do pracy zakładu produkcyjnego, w związku z czym w przypadku awarii, ciśnienie w zakładzie spadnie do zera. Kolejną cechą charakterystyczną zapotrzebowania jest bardzo zmienny pobór, wymagający częstego obciążania obu maszyn. Czas rozprężania maszyn wynosi ok. 90 s należy jednak zachować nieprzekraczalne 10 startów na godzinę, przez co nie można maszyn ciągle wyłączać. Dla okresu próbkowania

20 s oraz horyzontu 2 min przeprowadzono aż 62 eksperymenty przeszukiwania siatki kombinacji. Wyniki najlepszego wyszukanego modelu z nich zawiera tabela 5.8.

Tabela 5.8: Metryki bezpieczeństwa najlepszego modelu obiektu D3 w konfiguracji bazowej ($T_s = 20$ s, horyzont 2 min, funkcja `mse_bce`).

Sprężarka	N_T	TP	TN	FP	FN	S_{FN} [%]	S [%]
SP1	11 841	10 408	488	446	499	4,58	92,02
SP2	8 223	5 264	426	2 470	63	1,18	69,20
Suma	20 064	15 672	914	2 916	562	3,46	82,67

Przede wszystkim duża zmienność widoczna jest, przez pojawienie się pewnej ilości zdarzeń FN, jednak możemy założyć, że reguły zabezpieczające powinny ograniczyć ich liczbę na produkcji. Wskaźniki bezpieczeństwa pokazują, że układ zidentyfikował warunki załączenia sprężarki SP1 na poziomie S_{FN} wynoszącym 4,58%, a SP2 było dużo bardziej przewidywalne (maszyna wiodąca) uzyskując wynik 1,18%. Czyli nawet gdyby z systemu wyłączyć wsparcie jednej z maszyn z większym ryzykiem, układ powinien mieć potencjał optymalizacyjny przy niewielkim ryzyku (FN wynoszące 63 dla SP2). Zagregowany wskaźnik S_{FN} wynoszący 3,46% i wysoki wskaźnik bezpieczeństwa S na poziomie 82,67% sugerowały prawidłowe działanie układu. Zweryfikowano również proces uczenia, funkcja straty malała i krzywa uczenia była zbieżna.

Po wdrożeniu proponowanego modelu, okazało się, że bardzo szybko doszło do spadku ciśnienia. Model został wycofany w ciągu pierwszych dziesięciu minut w ciągu których, zdążył podjąć dwie ryzykowne decyzje. Warunki rzeczywiste okazały się miażdżące, dla zdawać by się mogło, bezpiecznego modelu.

5.9.2 Gradient ciśnienia

Przy metodyce wspomniano 3.1, że funkcję straty można łatwo rozszerzyć o dodatkowe cechy, które mogłyby być odpowiedzialne za ukryte zależności. Również można wprowadzić dodatkowe dane wejściowe, co spowoduje większe skupienie na przekazanych danych. Sprawdzono, czy nie należałoby sieci wskazać jednoznacznie kolejnego czynnika wejściowego obok stanów ciśnienia i pracy, będącego przyrostem trendu ciśnienia $\Delta p(t)$. Gdy ciśnienia spada, przy określonym nachyleniu charakterystyki będzie należało obciążyć kolejną maszynę. Przeprowadzono dziewięć eksperymentów, które wykazały, że rozszerzenie nie przynosi znaczącej poprawy. Zauważono natomiast przy przeglądaniu wykresów, że pojawiają się niepełne zdarzenia, co do których nie da się przewidzieć kolejnego stanu. Oznacza to, że bieżąca częstotliwość zapisu nie jest wystarczająca, by wykryć wszystkie zdarzenia. Próbkowanie odzwierciedlające rzeczywisty proces jest niezbędne dla prawidłowego działania modułu SI a omijanie istotnych

stanów jest niedopuszczalne, podobnie zaobserwowany przebieg ciśnienia musi odzwierciedlać jak najwierniej istotne zmiany trendu.

5.9.3 Znaczenie okresu próbkowania

Po wykonaniu inspekcji porównującej działanie w czasie rzeczywistym systemu SCADA z zestawieniem bieżących trendów na HMI sterownika nadrzędnego okazało się, że rejestrowana charakterystyka jest niepełna. Stan zapisany w systemie akwizycji nie oddaje wystarczająco dokładnie rzeczywistego przebiegu w związku ze znaczną dynamiką procesu. Podobnie w tym momencie, obserwowany stan obciążenia potrafi nie zostać zauważony w odpowiednim momencie i powiązany z narastającą lub opadającą charakterystyką.

Postanowiono sprawdzić, jak wiele informacji jest pomijanych, przy bieżącej 20 s częstotliwości zapisu. W tym celu zwiększono częstotliwość zapisu do 5 s aby uzyskać rzeczywiste dane porównawcze. Następnie sprawdzono rozkład cykli obciążenia sprężarek, gdy znajdują się one w stanie tłoczenia, oraz przerw pomiędzy nimi, gdy sprężanie nie następuje. Przy bieżących założeniach metodyki stan pracy maszyny nie jest brany pod uwagę, tzn. oszczędność energii elektrycznej jest widoczna na bazie liczników roboczogodzin, natomiast brak informacji czy pomiędzy cyklami sprężarka pracuje, czy jest zatrzymana.

Dane 5 s były danymi domyślnymi, na których wykonano symulację działania systemu SCADA, który zapisuje dane co 20 s. W związku z tym dane przedstawione w poniższej tabeli odnoszą się do tego samego zbioru. Wynik porównania median obciążenia przedstawiono w tabeli 5.9. Kolumna *Próbki* oznacza liczbę próbek, jaka przypada na medianę wyznaczonego biegu jałowego.

Tabela 5.9: Mediany czasu obciążenia i biegu jałowego, liczba okresów biegu jałowego na dobę oraz wskaźnik $\tau_{<h}$ obiektu D3 dla danych 5 s, tych samych danych przeskalowanych do 20 s (sym.) oraz danych archiwalnych 20 s z wcześniejszego okresu (arch.). Symbolem * oznaczono doby, w których SP2 była sprężarką wiodącą.

Dane	Sprężarka	Med. obc.	Med. brak obc.	Próbek	Okr./dobę	$\tau_{<h}$
5 s	SP1	40 s	410 s	82	226	23,9%
5 s	SP2	235 s	30 s	6	285	99,9%
20 s sym.	SP1	40 s	420 s	21	226	18,2%
20 s sym.	SP2	240 s	40 s	2	285	99,9%
20 s arch.	SP1	40 s	60 s	3	379	68,6%
20 s arch.	SP2	40 s	60 s	3	229	87,0%
20 s arch.*	SP1	40 s	80 s	4	374	60,0%
20 s arch.*	SP2	200 s	40 s	2	136	96,3%

Powyższe porównanie ukazuje inną własność, która wynika z rozpatrywanego zbioru, a mianowicie zmienną rolę sprężarek w układzie. W danych archiwalnych z całego okresu $\tau_{<h}$ sprężarki SP1 wynosi 68,6%, a w dobach, w których SP2 była sprężarką wiodącą, a SP1 dobijającą, tak jak w tygodniu danych 5 s, wynosi 60,0%. Rozbieżność względem danych 5 s (23,9%) nie wynika zatem ze zmiany ról, lecz ze zmiany profilu pracy układu pomiędzy okresami rejestracji, a wartości dla pojedynczej maszyny z całego wcześniejszego okresu nie można porównać bezpośrednio z tygodniem danych 5 s.

Na bazie liczby zebranych próbek wyłania się charakter pracy układu, w którym jedna maszyna jest wiodąca i rzadko się odciąża, a następnie wraca do obciążenia. Przy porównaniu liczby wykrytych cykli pozostała identyczna oraz wszystkie zdarzenia zostały wykryte, ale istotny jest moment ich wykrycia. Zauważalna jest różnica okresów bez obciążenia. Mediana powrotu do obciążenia dla sprężarki SP2 jest krótsza o 25% przy zwiększonym próbkowaniu co jednoznacznie informuje o bardziej adekwatnym odwzorowaniu przebiegu procesu. Pozostałe wartości statystyczne natomiast nie oddają w pełni intensywnej dynamiki procesu, gdzie 10 s różnica w obserwacji jest znacząca. Na co jeszcze należy zwrócić uwagę, to fakt, że domyślny okres próbkowania powoduje w tym momencie „zaokrąglenie" momentu zdarzenia do pełnego cyklu, która w zestawieniu z nieprawidłowym próbkowaniem przebiegu ciśnienia nie daje informacji co do podstawowych własności procesu.

Przy manualnej inspekcji wykresów proponowanych decyzji również zauważane były „braki" w trendach, gdy w ciągu trwania jednego cyklu pomiędzy próbkowaniem pojawiało się nielogiczne zachowanie układu. Model nie był w stanie znaleźć wzorców

tranzycji poprzez ich niepoprawne odwzorowanie w danych treningowych. Jak zakłada zaproponowana funkcja straty, musi następować korelacja pomiędzy wartościami analogowymi oraz dwustanowymi, która opisuje układ. Przy domyślnym próbkowaniu moment odciążenia oraz późniejsze zachowanie układu jest kluczowe dla właściwego odwzorowania. Mając do dyspozycji większą częstotliwość zapisu łatwiej znaleźć moment, który nadąża za charakterystyką oraz jest doskonałym punktem startowym dla systemu predykcji.

W związku z opisanym zjawiskiem pojawia się sprzeczność, której skutkiem były dobre metryki bezpieczeństwa w tabeli 5.8 i ich niewielka użyteczność eksploatacyjna. Dane przekazane na potrzeby trenowania nie odzwierciedlały po prostu rzeczywistej dynamiki procesu. Jest to bardzo ważna obserwacja oraz własność zaproponowanych metod oceny. Wskaźniki wyliczone na danych, nieodwzorowujących procesu, będą odnosić się do zgodności pomiędzy modelem a danymi treningowymi. Wymagane jest zatem sprawdzenie nauczonego modelu w rzeczywistych warunkach produkcyjnych, gdzie proces weryfikowany jest wg schematu „dane → model → wskaźniki → wdrożenie". Swego rodzaju autozapętlenie decyzji offline w skrajnych przypadkach, jak tutaj, nie weryfikuje w pełni użyteczności wdrożeniowej.

Mimo tego, że wskaźnik S_{FN} dla modelu nauczonego ze standardową rozdzielczością 20 s był najmniejszy spośród wszystkich wdrożeń (tabela 5.8), w rzeczywistości został wycofany w ciągu pierwszych dziesięciu minut wdrożenia. Dwie błędne rekomendacje doprowadziły do spadku ciśnienia w sieci, przez co trzeba było zatrzymać moduł SI. Przy wytrenowaniu modelu z rozdzielczością 5 s mimo, że uzyskano gorsze wyniki ewaluacji w tabeli 5.2, model lepiej reagował na nieprzewidywalność układu, podejmując adekwatne decyzje. W stosunku do 10 minut, których nie wytrzymała pierwotna sieć, tutaj utrzymała się na produkcji przez 35 dni.

Eksperyment nie został zrealizowany w sposób neutralnie porównawczy, ponieważ zmiany dotyczą zarówno okresu próbkowania, rozpiętości danych treningowych oraz ewaluacyjnych. Priorytetem było jak najszybsze wdrożenie funkcjonalnego rozwiązania. Spostrzeżeniem podsumowującym niniejszą sekcję niech będzie fakt, że wskaźniki wyznaczone offline na danych nieodzwierciedlających procesu nie dostarczają wystarczających rekomendacji wstrzymania rozruchu.

5.9.4 Zmiana okresu próbkowania

Zwiększony do 5 s okres próbkowania oznacza w praktyce czterokrotnie większą liczbę danych. Ponieważ obiekt cechuje duża zmienność zachowano 2-minutowy horyzont predykcji. W takim wypadku zamiast dostępnych wcześniej sześciu próbek konfiguracji startowej do dyspozycji w ramach jednego okna osiągalne są 24 próbki. Dla nowej konfiguracji wykonano 186 eksperymentów. W poprzednim podpunkcie mówiono

o rozkładzie mediany obciążenia, jednak nie oddaje ono rzeczywistej zmienności układu i w tym celu przedstawiono osobną tabelę 5.10, która zawiera rozkład cykli obciążenia.

Uzyskując czterokrotnie więcej danych zdecydowano się na czterokrotne „przyspieszenie" procesu uczenia. Do teraz trening wykonywany był na 3 tygodniach danych przy 1 tygodniu dla weryfikacji (aczkolwiek wskaźniki budowane były na zasadzie OOT). Przy dużej dynamice układu zdecydowano się na zachowanie proporcji i po zebraniu danych z tygodnia, 75% z nich przeznaczono na uczenie, a 25% do weryfikacji. Też do wstępnego rozruchu zrezygnowano z protokołu bezpieczeństwa i dokonano odejścia od zasad OOT by zweryfikować zdarzenia na całym tygodniu, na którym wykonano uczenie. Po tygodniu wrócono do nowego wyznaczenia wskaźników, ale chcąc przyspieszyć rozruch zdecydowano się na opisane odstępstwa od metodyki.

Tabela 5.10: Rozkład długości okresów obciążenia i biegu luzem [%] obiektu D3 (dane 5 s z tygodnia uczącego). Pogrubiono przedziały o największym udziale.

Sprężarka	Stan	< 60 s	60–120 s	120–180 s	180–300 s	300–600 s	> 600 s
SP1	obc.	**98,9**	0,9	0,2	0,0	0,0	0,0
SP1	luz	2,5	**21,4**	9,7	3,8	**58,7**	3,8
SP2	obc.	1,2	15,9	16,4	**31,3**	**34,3**	1,0
SP2	luz	**99,8**	0,1	0,1	0,1	0,0	0,0

Wysoka wartość S_{FP} (tabela 5.8) dla sprężarki SP2 potwierdza jej rolę jako sprężarki wiodącej. Bieg jałowy w jej wypadku w 99,8% trwa poniżej minuty. Jest to maszyna potrzebna do pokrycia bazowego zapotrzebowania, jednak zdarzają się okresy, gdy dochodzi do jej odciążenia. Model słusznie zauważa, że nie powinno się jej wyłączać. Okresy pomiędzy obciążeniami będą zależały od aktualnego zapotrzebowania układu.

Potencjał oszczędnościowy umożliwia sprężarka SP1, która jest maszyną dobijającą. W jej wypadku 98,9% okresów obciążenia trwa poniżej minuty co oznacza wysoką dynamikę układu. Z kolei przestrzeń cyklu biegu luzem rozkłada się w sposób, gdzie prawie 24% przerw wynosi poniżej dwóch minut, 62,5% powyżej pięciu minut. Okres w wyliczeniu od 180 do 300s stanowi tylko 3,8% przypadków. W związku z tym zadaniem sieci jest rozpoznanie okresów, gdy trwa intensywne, cykliczne dobijanie do górnej granicy ciśnienia, gdy przerwy w postojach są mniejsze od czasu rozprężania sprężarki. W pozostałych okresach możliwe jest zatrzymanie sprężarki i uzyskanie pełnego potencjału oszczędnościowego.

W przypadku układów o tak dużej dynamice, wskazane jest wykonanie również pełnego rozkładu długości cykli oraz dobranie takiego okresu próbkowania, które umożliwi adekwatną obserwację procesu.

5.9.5 Optymalizacja hiperparametrów

W przypadku badanego obiektu i dobranego okresu próbkowania najlepiej sprawdziła się funkcja `industrial` z wykorzystaniem optymalizacji bayesowskiej doboru hiperparametrów opisanej w sekcji 3.4. Sieć znaleziona przy pomocy przeszukiwania siatki kombinacji stała się punktem wyjścia do optymalizacji. Optymalizacja bayesowska dotyczyła architektury sieci, współczynnika regularyzacji oraz parametrów uczenia a dobrane parametry sieci przedstawiono w tabeli 5.11.

Tabela 5.11: Konfiguracja modelu produkcyjnego obiektu D3 po optymalizacji bayesowskiej.

Parametr	Wartość
Okres próbkowania T_s	5 s
Lookback / horyzont $L = h$	24 próbki (2 min)
Architektura	BiLSTM
Jednostki warstw	96 / 32
Dropout	0,2608
Współczynnik uczenia	0,010248
Rozmiar batch	8192
Funkcja straty	`industrial`
w_{reload}	0,72
Mnożnik kary FN	20,0

Optymalizacja doprowadziła do redukcji wielkości sieci ze startowych 100/50 do 96/32. Warstwa dropout została zwiększona z domyślnych 0,2 do 0,2608. Ocena modelu w oknie decyzyjnym 2 min wykazała, że dla sprężarki SP1 model nie przewiduje powrotu obciążenia, przez co wszystkie błędy FN w tabeli 5.2 dotyczą tej maszyny.

5.9.6 Opóźnienie predykcji

W poprzednich sekcjach wspomniano o istotności opóźnienia wywołania predykcji. Z punktu widzenia współpracy z systemem SCADA, który zapisuje domyślnie dane co 20 s, pojawienie się informacji o wystąpieniu tranzycji odciążenia niesie w sobie, wraz z aktualnym pomiarem ciśnienia niejawną informację. Jest nią przyrost charakterystyki ciśnienia $\Delta p(t)$, który pojawia się również w wariancie użytej funkcji straty `industrial`.

Przy domyślnym opóźnieniu 20 s nie zwrócono uwagi na istotność tej informacji. Przy przyspieszeniu do 5 s okazało się, że decyzje predykcyjne są nieadekwatne. Przy zwiększonym okresie próbkowania pierwszy pomiar często zauważał stan odciążenia, ale charakterystyka ciśnienia utrzymywała się na zbliżonym poziomie — nie było wiadomo

czy będzie narastać czy opadać. W związku z tym niezbędne było dodanie parametru, który będzie realizował opóźnienie predykcji.

W wypadku rozpatrywanego obiektu D3 przetestowano cztery wartości opóźnienia: 5, 10, 15 oraz 20s. Okazało się, że najlepiej sprawdziło się opóźnienie o cztery próbki. Dotrzymanie takiego okresu niosło w sobie wszystkie istotne informacje, które pozwoliły na wsparcie decyzji sterujących.

Jest to zauważony dodatkowy mechanizm, który jest niezbędny dla obiektów o dużej zmienności i pozwala na zapewnienie sobie odpowiedniej obserwowalności procesu. Parametr powinien zostać dobrany indywidualnie dla każdego obiektu w którym domyślna przedstawiona metodyka nie sprawdza się.

5.9.7 Opóźnienie realizacji decyzji sekwencyjnych

Parametr opóźnienia realizacji decyzji sekwencyjnych, został dobrany jako zabezpieczenie procesu dla obiektów, które cechują nagłe spadki ciśnienia jako uzupełniający do parametru progu ciśnienia, poniżej którego decyzje są zablokowane. Domyślnie próg ciśnienia powinien być na poziomie dolnej kaskadowej granicy ciśnienia sterownika nadrzędnego.

W związku z wystąpieniem kilku nieprzewidywalnych zdarzeń, które nie mają wzorca, dodanie parametru czasowego nie uzależnia od jednego poziomu ciśnienia, tylko umożliwia wstrzymanie decyzji wyłączenia dla nagłych spadków ciśnienia.

Drugą rolą parametru jest zabezpieczenie decyzji kaskadowych. Spotkano się z sytuacjami gdy dla obiektów o dużej zmienności następowało sekwencyjne odciążanie maszyn w krótkich (np. 20 s) cyklach. Wówczas charakterystyka ciśnienia może się jeszcze nie ustabilizować i konieczne będzie sekwencyjne wydawanie decyzji by uniknąć sytuacji, w których w jednym cyklu np. nastąpi wysłanie decyzji wyłączenia dla kilku maszyn. Nawet jeżeli taka decyzja okazałaby się prawidłowa, to nadal istnieje ryzyko nagłego spadku ciśnienia, więc tutaj proponuje się zaakceptowanie straty energetycznej i poczekania przez czas opóźnienia do stabilizacji.

5.9.8 Eksploatacja i zakończenie pracy

Po pierwszych problemach moduł SI został uruchomiony ósmego dnia miesiąca startowego M_0. Podczas pierwszych dwóch tygodni pracy wystąpiły trzy błędy operacyjne, jednak od trzeciego tygodnia wskaźnik R_{FN} (tabela 5.6) zaczął narastać osiągając w czwartym tygodniu wartość 0,67%.

Przy inspekcji wzrokowej zauważano coraz większe wahania charakterystyki wynoszące ok ±0,1 bar w stosunku do dotychczasowej pracy. Charakterystyka zapotrzebowania również zaczęła się zmieniać i po czterech tygodniach pracy maszyny pracowały

więcej niż w miesiącu startowym o około 24%. Co istotne tendencja wzrostowa utrzymała się i w kolejnych dwóch miesiącach wzrosła o ok. 32% oraz 36%. Praca maszyny dobijającej została zintensyfikowana. Analizując zbiór treningowy, okazało się, że od połowy okresu treningowego te zmiany były już widoczne.

Dotrenowanie sieci w oparciu o nowe warunki nie było już możliwe. 13 dnia miesiąca M_{+1}, po 35 dniach pracy, 5530 potwierdzonych decyzjach moduł SI został wyłączony. Przyczyną był wzrost wskaźnika R_{FN} gdzie postanowiono przyjrzeć się nowej charakterystyce układu bez ingerencji modułu SI. Decydujący był jednak drugi czynnik.

Kilka dni od wyłączenia, klient zgłosił potrzebę zmiany zasad pracy obecnego układu. Przy spadkach zapotrzebowania jedna maszyna wyłączała się po odliczeniu czasu pracy na luzie co powodowało co jakiś czas spadek ciśnienia podczas rozprężania przed kolejnym startem. Nowy profil produkcji nie mogł pozwolić sobie na spadki ciśnienia w związku z czym zdecydowano o znacznym wydłużeniu czasu biegu jałowego.

Ze względu na niewielką liczbę danych w wypadku tego obiektu zdecydowano się również na wzięcie pod uwagę miesiąca rozruchu. Podczas pracy w miesiącu realizacji wdrożenia bieg jałowy obniżył się do poziomu 27,6% co stanowi poprawę o 20,7% w stosunku do wartości bazowej a w kolejnych miesiącach biegu luzem powrócił w okolicę poziomu przed optymalizacją (wartość została znacznie zwiększona, ale nie ustawiona tak wysoko by maszyny były zawsze w stanie ciągłego oczekiwania). Należy zwrócić uwagę na użyteczność wskaźnika R_{FN}, który zadziałał prewencyjnie (w momencie kontaktu ze strony odbiorcy moduł SI był wyłączony w celu obserwacji nowej charakterystyki). O zakończeniu pracy modelu zdecydował również odbiorca. Być może w ramach badania kolejnych kierunków rozwoju nastąpi ponowne podejście do problemu optymalizacji tego obiektu.

5.9.9 Wnioski ze studium przypadku

Przypadek obiektu D3 doskonale odzwierciedla potrzebę indywidualnego, świadomego, a nawet krytycznego podejścia do analizowanego obiektu. Podczas eksperymentu i użycia modułu SI nasuwa się kilka wniosków, które opisane zostały w kolejnych akapitach.

W pierwszej kolejności należy dokonać oceny dynamiki obiektu w celu dobrania odpowiedniego okresu próbkowania (jeżeli jest taka opcja od strony technicznej). W przytoczonych wdrożeniach, w większości przypadków proces był wolnozmienny jednak przypadek D3 musiał zostać potraktowany indywidualnie, inaczej od pozostałych obiektów. W pierwszej kolejności należy zapoznać się z danymi zbieranymi z częstotliwością np. 1 s na HMI sterownika nadrzędnego i porównać je z danymi zapisywanymi dla uczenia. Wykrywanie wszystkich zdarzeń obciążenia nie jest wystarczające jeżeli nie jest powiązane ze zmianami charakterystyki ciśnienia.

W przypadku prawidłowego okresu próbkowania unikamy sytuacji, którą wykazało to wdrożenie, czyli przypadku gdy wartości wskaźników bezpieczeństwa nie są dowodem faktycznego bezpieczeństwa potencjalnego wdrożenia. Bez wcześniejszego zapoznania się z przebiegiem indywidualnym, na domyślnej 20 s okresowości zapisu otrzymano najniższy ze wszystkich obiektów wskaźnik $S_{FN} = 3{,}46\%$. Mimo tego, obiekt nie nadawał się produkcyjnie. Założono tutaj również, że otrzymana sieć jest zbieżna i proces uczenia nie budzi wątpliwości.

W przypadku zmian okresu próbkowania, należy również zbadać, jaki jest najlepszy moment wykonywania predykcji. Taki eksperyment najlepiej przeprowadzić offline i zbadać, które opóźnienie daje najlepsze wyniki. Dynamika układu może wymuszać określone opóźnienie, które najlepiej sprawdza się przy danej dynamice. Z dużym prawdopodobieństwem można powiedzieć, że zbyt wczesne wykonanie predykcji, będzie niosło w sobie zbyt małą informację co do przewidywanych kolejnych kroków.

Niezbędny jest ciągły monitoring wskaźnika R_{FN}, który jest systemem wczesnego ostrzegania, gdy charakterystyka obiektu oraz modelu zaczyna się rozmijać. Jest to normalne zjawisko w przypadku gdy zakład zaczyna zmieniać swój profil produkcyjny. W przypadku D3 coraz większe obciążenie maszyn doprowadziło do tego, że model nie był już odpowiedni dla nowych wymagań. Rzeczywisty układ może ciągle zmieniać swoje zapotrzebowanie przez co tego czynnika nie można sprawdzić przed wdrożeniem. Ostatecznie, obiekt pracował poprawnie przez 35 dni, a pogarszający się wskaźnik R_{FN} zdecydował o jego wyłączeniu.

Ostatni aspekt to czynniki nieprzewidywalne, które pojawiają się w instalacjach przemysłowych. Odbiorca zgodził się na wyższy koszt energii elektrycznej i utrzymywanie maszyn na biegu jałowym by w zamian otrzymać stabilność dostaw sprężonego powietrza. Podobnie mogą wystąpić zdarzenia niezależne, które będą wymagały wyłączenia niektórych maszyn a wówczas model nie będzie adekwatny. Podstawowym zaleceniem przed wdrożeniem systemu predykcyjnego jest to, by najpierw doprowadzić system sterowania nadrzędnego do optymalnej pracy, a potem dopiero przystępować do procesu uczenia.

5.10 Podsumowanie

Pierwsze wdrożone obiekty, działały przez kilka początkowych dni bez dodatkowych zabezpieczeń. Podjęto takie ryzyko, by sprawdzić, czy proponowana metodyka jest adekwatna. Nie spowodowano żadnego spadku ciśnienia, natomiast obserwując kolejne decyzje, zauważono pewne zdarzenia nieprzewidziane, których waga podczas uczenia mogła być zbyt mała. Jednym z takich zdarzeń może być np. intensywny spadek ciśnienia związany z nagłym zapotrzebowaniem tuż po odciążeniu i wysłaniu decyzji

wyłączenia. Bez powstrzymania takiej decyzji może dojść do wyłączenia maszyny poniżej dolnej granicy ciśnienia. Z czasem działania systemu rozpoczęto profilaktyczne dodawanie kolejnych zabezpieczeń, które okazały się potrzebne na innych obiektach. W trakcie działania system podjął ponad 133 tysiące decyzji w ciągu rozpatrywanych 806 dni i osiem obiektów. Dzięki mechanizmom bezpieczeństwa w ramach zaproponowanej metodyki odsetek błędnych decyzji stanowił 0,085% i nie doprowadził do spadków ciśnienia. Tutaj należy podkreślić, że gdyby nie nastąpiło wdrożenie, proponowane reguły bezpieczeństwa nigdy by nie powstały, co podkreśla potrzebę zweryfikowania wyników badań. W związku z tym należy przedstawić kilka wniosków podsumowania.

Proponowana metoda ewaluacji OOT bazująca na rzeczywistych tranzycjach oraz oparte o nią metryki bezpieczeństwa okazały się skuteczne. Ze względu na gorsze wyniki obiektu B4 nie zdecydowano się na jego wdrożenie. By zapewnić jak największe bezpieczeństwo produkcyjne wprowadzono dodatkową regułową warstwę zabezpieczeń, która filtruje decyzje warstwy predykcyjnej. Aktywność warstwy zabezpieczeń będzie zależała przede wszystkim od dynamiki obiektu. W przypadku dotrenowywania, na podstawie przeprowadzonych eksperymentów, nie da się jednoznacznie powiedzieć, że przynosi wymierne korzyści, jednak nie pogarsza wstępnie nauczonego modelu. Przy przeanalizowaniu wykresów oraz podejmowanych decyzji, można zauważyć, że układy pracowałby wymiernie gorzej niż bez modułu SI generując większe straty energii. Dotrenowywany obiekt D4 wykazuje statystyczną poprawę, ale może to wynikać ze zmienności profilu produkcji.

Uzyskane oszczędności energetyczne w odniesieniu do średniej z trzech miesięcy poprzedzających wdrożenie są zróżnicowane i powiązane z dynamiką obiektu oraz silnie z profilem produkcji. Najlepsze uzyskane wyniki redukują bieg jałowy o 83% (obiekt D4) i 66% (obiekt A3). W przypadku obiektu silnie zoptymalizowanego jak C1 różnica jest statystyczna, ponieważ zauważono pogorszenie wskaźnika z poziomu 2,0% na 2,6% co dalej stanowi niewielką wartość. W przypadku drugiego najgorszego wyniku dla obiektu D1 otrzymano pogorszenie na poziomie 17%, ale w tym wypadku bez modułu SI wynik byłby gorszy. W każdym z tych przypadków system działał prawidłowo, a uzyskany efekt zależał od aktualnego zapotrzebowania na sprężone powietrze charakteryzującego pracę danego zakładu produkcyjnego.

Przy wdrożeniu zauważono również cztery niepożądane mechanizmy degradacji modułu SI: brak możliwości optymalizacji (obiekty grupy C), awarie maszyn lub powiązanej infrastruktury (wyłączenie jakiejś maszyny z eksploatacji), zmiany profilu produkcji (dryft koncepcyjny) oraz stopniowy zanik korzyści energetycznych mimo dotrenowywania. Każdy z efektów wymaga innego podejścia, by umożliwić ciągłą pracę modułu SI w nowych warunkach. Jednocześnie nie zdecydowano się na wyłączenie modułu SI w związku z awariami pojedynczych maszyn.

Proponowanym zaleceniem jest ciągły monitoring wskaźnika R_{FN}, który pogarszając się zachęca do zbadania z którym mechanizmem degradacji w danym wypadku ma się do czynienia. Uzyskuje się w ten sposób system wczesnego ostrzegania, który pozwala na podjęcie odpowiednich kroków prewencyjnych.

Kolejny istotny aspekt, który trzeba wziąć pod uwagę to potencjalne oszczędności energii a ryzyko nieprawidłowych decyzji FN. Maszyny są ograniczone konstrukcyjnie, mają określoną liczbę startów na godzinę, przy odciążeniu nie zaleca się, by od razu je wyłączać, tylko powinny przepracować jakiś minimalny czas na biegu jałowym. Po zatrzymaniu następuje czas rozprężania — w przypadku złej decyzji, w tym momencie następuje spadek ciśnienia przez czas, kiedy maszyna nie jest gotowa do startu. W pewnych przypadkach trzeba zachować kompromis pomiędzy długością biegu jałowego a stabilnością dostaw. W przypadku obiektu D3 moduł SI mógłby pracować nadal, ale zmieniające się priorytety odbiorcy doprowadziły do wyłączenia modułu. W skrajnych przypadkach pole oszczędności może być niewielkie.

W chwili pisania niniejszej rozprawy wszystkie wspomniane obiekty, poza wyłączonymi prewencyjnie, pracują nadal. Nie doszło do zachwiania stabilności dostaw sprężonego powietrza. W przypadku jednego obiektu doszło do przejścia z grupy A do grupy B gdzie moduł SI w połączeniu z nowoczesną maszyną osiągnął poziom biegu luzem całej stacji na poziomie 1%.

6 Dyskusja

Rozdział zawiera podsumowanie wszystkich prac, jakie zostały wykonane w ramach części badawczej oraz wdrożeniowej niniejszej rozprawy. Na początku podjęto się interpretacji wyników badawczych, roli funkcji straty, wpływu architektury sieci i znaczenia horyzontu predykcji. W kolejnej części opisano wnioski płynące ze wdrożenia dotyczące zaproponowanych wskaźników oceny oraz ich rozbieżności, podkreślono rolę ciągłego nadzoru operacyjnego, zwrócono uwagę na wierność odwzorowania procesu i zaprojektowanych parametrów czasowych. Podano również wytyczne co do możliwych do osiągnięcia oszczędności energii i metryk mogących je określić. Podano również wskazówki, kiedy użycie uczenia maszynowego/głębokiego jest uzasadnione, a kiedy można pozostać przy istniejących rozwiązaniach. Kolejnym zagadnieniem jest jednoznaczne wskazanie, jakie elementy pracy pochodzą z literatury, a jakie stanowią wkład autorski. By umożliwić właściwą interpretację wyników, omówiono ograniczenia badań. Rozdział zamyka umiejscowienie proponowanego rozwiązania w kontekście założeń Przemysłu 4.0.

6.1 Interpretacja wyników badawczych

6.1.1 Rola funkcji straty

Na wstępie należy podkreślić duże zróżnicowanie dynamiki układów, z jakimi można się spotkać w sektorze sprężonego powietrza w niniejszej rozprawie jak i w rzeczywistości. Reguły, jakie można by przyjąć za pewnik to te, że w momencie, gdy ciśnienie spada trzeba będzie obciążyć maszynę, a gdy rośnie powyżej ustalonego poziomu to nastąpi odciążenie. W praktyce te momenty będą zależeć silnie od zmiennego zapotrzebowania, a park maszynowy może składać się z kilkudziesięciu lub nawet kilkuset odbiorców. Zamiast statystycznego podejścia poszukiwano metody, która pozwoli na wsparcie pracujących układów w czasie rzeczywistym. W tym celu w publikacjach własnych zaproponowano bazowy wariant funkcji straty `mse_bce` bazujący na sieci LSTM [68], [69], która miała nieść w sobie informację o składowej poboru i część regresyjną zmiennego zapotrzebowania. Studium ablacji wykazało, że bazowa funkcja straty rozszerzona o architekturę dwukierunkową jest doskonałym punktem wyjścia dla wszystkich grup obiektów. Przy dostarczeniu sieci informacji o koszcie straty FN oczekiwana była dalsza poprawa wyników, jednak okazało się, że poprawa jest kwestią indywidualną silnie

związaną z określonym przypadkiem. Należy przytoczyć obiekt A3 (sprężarkownia bez przetwornic częstotliwości), który w wariancie `industrial` wykazał poprawę obniżając wskaźnik S_{FN} z 50,0% do 35,7%, przy czym liczba nieprawidłowo wykrytych zdarzeń była minimalna. Wariant MSE miał również wskaźnik S_{FN} na tym samym poziomie, ale przyjął strategię przewidywania zakładającą często pozostawanie maszyn pod obciążeniem uzyskując niższy wskaźnik bezpieczeństwa S na poziomie 73,2%. Przypadek, w którym wynik się pogarsza to B1 (grupa obiektów z jedną przetwornicą częstotliwości), gdzie dla `industrial` S_{FN} wzrasta do 94,0% w stosunku do bazowego wariantu `mse_bce` i poziomu 49,1%. Kolejną ciekawą własnością wynikającą z dynamiki obiektu jest pojawienie się największej w całym zbiorze ablacji B1 zdarzeń FN (578) co oznacza, że model przyjął odwrotną strategię do zakładanej gdzie, zamiast pozostawiać maszyny pod obciążeniem, model zakładał pozostawienie ich na biegu jałowym.

Różne warianty funkcji straty pozwalają dobrać określoną funkcję straty, która najlepiej sprawdzi się w przypadku danego obiektu. Jej składowe są powiązane w sposób niejasny, a proponowana metodyka pozwala zweryfikować jej stosowalność w konkretnym przypadku. Być może dysponowanie dodatkowymi informacjami od strony odbiorcy umożliwiłoby przebudowę funkcji straty lub zbioru uczącego w sposób pozwalający na adekwatniejsze predykcje. Zaproponowana metodyka weryfikuje w sposób zbliżony do wdrożenia, użyteczność wsparcia modułem SI. Pozwala ocenić czy dla danego obiektu warto podjąć się optymalizacji, co jest cenną informacją praktyczną. Wdrożenie pokazało, że pięć z ośmiu wdrożeń produkcyjnych pozostało przy wariancie bazowym z domyślnymi parametrami. Natomiast w pierwszej trójce największych uzyskanych oszczędności występują dwa warianty bazowe oraz jeden `industrial`, który uplasował się na drugim miejscu.

Proponowana funkcja `industrial` sprawdziła się w eksploatacji obiektu D3. To obiekt wysokiego ryzyka, z dużą nieprzewidywalnością, gdzie każda błędna decyzja może prowadzić do spadku ciśnienia. W takim przypadku zwrócenie uwagi na wysoki koszt FN jest ważniejsze niż pominięcie możliwych oszczędności energetycznych FP. Dla takich przypadków zaprojektowano asymetryczną funkcję straty. Tutaj założeniem projektowym było, że dla obiektów bardziej wymagających, gdzie należy zwrócić uwagę na koszt błędnych decyzji, taka funkcja byłaby bardziej adekwatna od wariantu bazowego.

6.1.2 Wpływ architektury sieci

Najbardziej jednoznacznym wynikiem studium ablacji okazała się przewaga architektury dwukierunkowej nad rozwiązaniami jednokierunkowymi. Warianty LSTM oraz GRU, przy identycznej funkcji straty i identycznych pozostałych hiperparametrach, degenerowały się — dla obiektu A3 wskaźnik S wyniósł 6,0% wobec 99,2% dla wariantu

dwukierunkowego, a dla obiektu C3 odpowiednio 32,4% wobec 99,8%. Model jednokierunkowy niemal bezwarunkowo sygnalizował powrót obciążenia, przez co osiągał pozornie doskonały wskaźnik S_{FN} równy zeru. Wskaźnik jest artefaktem degeneracji, a nie dowodem jakości modelu. Model, który nigdy nie rekomenduje wyłączenia, jest bezpieczny wg wskaźników i jednocześnie bezużyteczny. Dlatego nie zaleca się oceny modelu na podstawie pojedynczego wskaźnika.

Wyjaśnienie przewagi architektury dwukierunkowej wynika z przebiegu samego procesu technologicznego. Tranzycja obciążenia sprężarki ma rozpoznawalny kształt przebiegu ciśnienia zarówno przed (ciśnienie narasta do wartościo progowej), jak i po zdarzeniu (ciśnienie opada lub stabilizuje się na nowym poziomie). Sieć jednokierunkowa dysponuje wyłącznie kontekstem poprzedzającym i musi rozstrzygnąć, czy obserwowany wzrost ciśnienia zapowiada zmianę stanu, czy jest jedynie wahaniem zapotrzebowania. Sieć dwukierunkowa, przetwarzając sekwencję w obu kierunkach w oknie obserwacji, dysponuje pełnym wzorcem zdarzenia. Zbliżone wyniki LSTM oraz GRU wskazują przy tym, że decydująca jest kierunkowość przetwarzania, nie zaś szczegóły mechanizmu bramkowania.

6.1.3 Znaczenie horyzontu predykcji

Za wprowadzeniem horyzontu predykcji stała ciekawość poznania jak bardzo przeszłość wpływa na przyszłość. Pojawiło się pytanie: czy ucząc sieć na zbiorze wielu sytuacji powiązanych ze sobą, gdzie mamy pewne zapętlenie pomiędzy chwilami poprzedzającymi dane zdarzenia pojawia się wzorzec czy taki nie występuje? Dodatkowym założeniem również była niejawna informacja o bezpieczeństwie. Założono, że jeżeli w danym horyzoncie maszyna nie jest potrzebna, należy zapewnić jej możliwość wyłączenia, by po czasie rozprężania była gotowa do startu. Wówczas w przypadku zbyt szybkiego wykorzystania wszystkich startów maszyna będzie musiała pozostać na biegu jałowym przez czas umożliwiający jej ponowny rozruch, by nie przekroczyć liczby startów zdefiniowanych przez producenta.

Dobór horyzontu predykcji będzie zależał od charakterystyki danego obiektu. Spośród zbadanych obiektów można wyróżnić trzy przypadki zachowania wskaźnika S_{FN}:

- obiekty A1, B1, B2, B4, C1 — spadek wartości wskaźnika wraz z wydłużaniem okna, największa dla obiektu A1 z 67,1% do 39,7%, co oznaczałoby obiekty z mniejszą zmiennością, wymagające dłuższego horyzontu czasowego by uchwywcić kontekst,

- obiekt A3, C3 — brak wpływu horyzontu na wyniki, co sugeruje powtarzalne zdarzenia, łatwe do uchwycenia niezależnie od horyzontu,

- pozostałe obiekty — wzrost wskaźnika wraz z wydłużaniem horyzontu, w skrajnym przypadku obiektu C4, wskaźnik wzrasta z 18,2% do 63,6% (znaczny wzrost ryzyka), co jest zrozumiałe dla układów falownikowych gdzie ważniejszy jest kontekst bezpośrednio poprzedzający zdarzenie odciążenia a większa liczba danych wpływa na osłabienie predykcji.

Wybór horyzontu predykcji będzie również zależał od pozostałych wskaźników i adekwatności odwzorowania charakterystyki obiektu. Dlatego należy dobrać taki punkt pracy, który zapewni niewielkie ryzyko zdarzeń S_{FN} przy utrzymaniu potencjału oszczędnościowego S_{FP} oraz błędu RMSE nadążającego za charakterystyką ciśnienia. Również pewną kwestią dotyczącą potencjalnych oszczędności może być zgoda na większe ryzyko kosztem potencjalnych błędów natomiast tutaj mogą pomóc regułowe zabezpieczenia ustalone przy wdrożeniu.

6.2 Wnioski z wdrożenia

6.2.1 Rozbieżność wskaźników oceny offline a wskaźnikiem oceny operacyjnej

Zaproponowana metodyka sugerowała potencjał wystąpienia pewnej liczby zdarzeń FN po dokonaniu wdrożenia. Zdarzenia wyliczane były w momencie wystąpienia tranzycji odciążenia, co miało symulować rzeczywiste wsparcie sterownika nadrzędnego. W praktyce uzyskano sytuację gdzie dla obiektu np. D1 wskaźnik S_{FN} wyniósł 100%, czyli nie udało się rozpoznać żadnego stanu powrotu do obciążenia (24) co jest związane ze sposobem budowania wskaźnika i jest zagadnieniem pojawiającym się w literaturze [81]. Podczas działań operacyjnych pojawiła się jednak tylko jedna błędna decyzja FN_{op} przy ponad dwudziestu dwóch tysiącach decyzji. Analogiczny przypadek dotyczy również obiektu D4. W obu tych przypadkach nastąpiło jednak zadziałanie warstwy zabezpieczeń, które nawet przy niewielkim udziale wstrzymując odpowiednio 0,5% oraz poniżej 0,1% rekomendacji wyłączenia pozwoliło uniknąć niebezpiecznych stanów.

Wprowadzony wskaźnik nadzoru operacyjnego R_{FN} odnosi się do decyzji podjętych względem rzeczywistej instalacji, przy układzie, którego model dynamiczny nie jest znany. Wskaźnik S_{FN} bazuje na danych zapisanych w pamięci i nie zawiera żadnych zabezpieczeń, które są obecne na wdrożonych instalacjach. W danym momencie nie są znane zmiany zapotrzebowania od strony odbiorcy, przez co nie jest wiadome, jakie przełożenie na produkcję będzie miała negatywna decyzja. Np. wyłączenie maszyny może skutkować utrzymaniem ciśnienia tuż poniżej dolnej granicy ciśnienia, ale nie doprowadzić do spadku i zachwiania dostaw. Gdyby jednak zakład produkcyjny uruchomił w tym momencie wiele kolejnych odbiorników, to bez rezerwy operacyjnej dojdzie do spadku ciśnienia.

Podsumowując ocena ryzyka poza metodami offline wymaga również obserwacji ruchowej oraz nadzoru operacyjnego śledzącego degradację (lub brak degradacji) wdrożonego modelu.

6.2.2 Znaczenie nadzoru operacyjnego

Projektując proponowane rozwiązanie, mając świadomość zmienności odbiorców, oraz dryft koncepcyjny należało wprowadzić mechanizm ciągłego monitoringu adekwatności podejmowanych decyzji. W tym celu zaprojektowany został wskaźnik R_{FN}. Jego użyteczność okazała się dużo wyższa, niż przypuszczano na etapie projektowania. Wskaźnik pozwala bez wnikania w logikę zabezpieczeń, na bazie okna czasowego zweryfikować czy nie doszło do obciążenia maszyny po rekomendacji wyłączenia. W tym momencie nie potrzeba weryfikacji przebiegu czy innych metod ewaluacji, ponieważ działanie odbywa się na podjętych decyzjach oraz blokadach. W ten sposób osiągnięta została kontrola jakości decyzji. Obserwacja przebiegu wskaźnika w czasie ułatwia rozpoznania omówionych mechanizmów degradacji. Warto tutaj przytoczyć trzy przypadki:

- obiekt A3 — nagły skok wskaźnika wykrył awarię infrastruktury, gdzie została zmodyfikowana konfiguracja maszyn,
- obiekt C1 — w związku z niewielką liczbą decyzji, wysoki poziom wskaźnika pokazuje własność obiektu, którą jest brak miejsca dla dalszej optymalizacji,
- obiekt D3 — systematyczne narastanie wskaźnika, było skutkiem zmiany profilu produkcji.

Przy sprawdzaniu wskaźnika R_{FN} należy wziąć również pod uwagę liczbę podejmowanych decyzji. Tutaj doskonałym przykładem jest obiekt C1, który ma najwyższą wartość wskaźnika w całym zbiorze, jednak rekomendacje wyłączeń na dobę nie przekraczają dwudziestu. Nieprawidłowe wyłączenie pojawia się jeden raz na kilka dni pracy. Z drugiej strony obiekt D4 mający ponad dwieście decyzji dziennie odnotował jeden błąd na kilka miesięcy. Oba obiekty pracują poprawnie, a różnica pomiędzy nimi wynika z budowy sprężarkowni oraz zapotrzebowania. W jednym wypadku system jest zoptymalizowany wyposażony w nowoczesne maszyny, a w drugim pracuje w bezpiecznych warunkach.

6.2.3 Wierność odwzorowania procesu

Studium przypadku D3 zostało przedstawione bardziej szczegółowo, by pokazać przypadek, w którym domyślne założenia się nie sprawdziły. Zwraca uwagę na bardzo istotny aspekt, który trzeba wziąć pod uwagę. W związku z ograniczeniami dostępnych

systemów przy realizacji archiwizacji narzucony był domyślny próg zbierania danych wynoszący 20 s. Próg ten dał domyślny wskaźnik S_{FN} na poziomie 3,5%, umieszczając go w czołówce całego zbioru badawczego. W praktyce obiekt nie wytrzymał dziesięciu minut od uruchomienia, gdzie po podjęciu dwóch błędnych decyzji został wstrzymany. W takim wypadku dokonano szczegółowej inspekcji, dzięki której okazało się, że zapisany przebieg odbiega znacznie od rzeczywistego przebiegu. Zapisane dane treningowe nie nadążały za dynamiką procesu, mimo że zdarzenia nie były pomijane, i wszystkie się zgadzały, to moment ich zapisu nie niósł w sobie potrzebnej informacji. By sieć działała efektywnie, stan maszyn musi być powiązany ze zmianą ciśnienia. Zbadano również, że sieć neuronowa musi podjąć się predykcji nie od razu, gdy tylko zauważono odciążenie, tylko po kilku próbkach, gdy charakterystyka obiektu stabilizuje się w określonej tendencji.

Najważniejszym podsumowaniem tego zagadnienia, jest to, że by wskaźniki oceny były adekwatne, dane treningowe muszą odzwierciedlać dynamikę obiektu. W innym wypadku metryki sprawdzają owszem zgodność modelu z zebranymi danymi, ale nie przekładają się na produkcję. W związku z tym zauważone niebezpieczeństwo jest kolejnym czynnikiem, który należy wziąć pod uwagę w proponowanej metodzie. Bezwzględnie zebrane dane muszą odpowiadać przebiegowi procesu. W związku z tym w przypadkach obiektów o większej dynamice liczba próbek związana z pojedynczym zdarzeniem również jest kluczowa.

6.2.4 Parametry czasowe realizacji decyzji

Dzięki obiektowi D3 również wykryto pewną własność szeregów czasowych, która była przekazywana w sposób niejawny i poniekąd pozostała parametrem domyślnym. Tutaj jeszcze należy przypomnieć, że na produkcji układy wspierania decyzji pracują z 5 s częstotliwością i domyślnie, było przyjęte założenie, że jeżeli uczenie odbywało się na 20 s danych to moment predykcji również musi się odbyć po upłynięciu czasu próbkowania zbioru treningowego. Wdrożenie pokazało, że jest to bardzo istotne założenie i niesie w sobie ukrytą informację o kierunku zmian ciśnienia. Przy domyślnym uruchomieniu predykcji z okresem 5 s okazało się, że decyzje predykcyjne znacząco rozmijają się z rzeczywistością. Przy wprowadzeniu opóźnienia 20 s predykcje były już prawidłowe.

Kolejny parametr pojawiający się wiąże się z układami wielosprężarkowymi. Pierwotnym założeniem było uniknięcie sytuacji, w której próbka po próbce pojawiłyby się odciążenia kolejnych maszyn. Wówczas układ znajduje się w stanie nieustalonym i otrzymanie od sieci rekomendacji wyłączenia kolejnych maszyn wiązałoby się z niepotrzebnym ryzykiem. W związku z tym podjęto decyzje wprowadzenia dodatkowego

zabezpieczenia sekwencyjnego. Jego zadaniem jest ochrona przed nagłym spadkiem ciśnienia. Ceną za to będzie utrzymywanie kolejnych maszyn na biegu jałowym w oczekiwaniu na zapotrzebowanie. Ta niewielka strata energetyczna będzie jednak gwarancją stałości dostaw sprężonego powietrza.

Dwa wyłaniające się parametry tzn. jeden odpowiedzialny za obserwowalność procesu oraz drugi uodparniający na stany nieustalone powinny zostać dobrane zgodnie z zaproponowanymi założeniami lub doświadczalnie, jeżeli dynamika układu odbiegałaby od tej przedstawionej w rozprawie.

6.2.5 Oszczędność energii

Możliwości oszczędności energii elektrycznej będzie w dużym stopniu zależeć od zmienności oraz powtarzalności profilu produkcji danego zakładu. Dany cykl produkcyjny zakładu może przekładać się na mniejsze lub większe zapotrzebowanie, inne pracujące maszyny czy zmiany w poziomie ciśnienia (im niższe ciśnienie, tym większe oszczędności). Przyjęto trzymiesięczny okres poprzedzający wdrożenie jako punkt odniesienia dla wprowadzonych zmian, jednak nie jest znany pełny profil produkcyjny w tych miesiącach od strony popytowej. Przy niektórych porównaniach rozpatrywano okresy o zbliżonej liczbie godzin, ale również tutaj brakuje informacji od strony klienckiej. Rezultaty statystycznie wykazują jakąś poprawę i zachęcają do wdrożenia w niektórych przypadkach jednak nie dysponowano rzeczywistymi danymi z liczników energii elektrycznej oraz przepływomierzy, które w sposób deterministyczny przedstawiłyby wypracowane oszczędności. Mniejszy udział biegu jałowego we wszystkich godzinach przepracowanych przez stację jest mierzalnym, podstawowym wskaźnikiem oszczędności. Przy analizie wykresów działającego układu również widać wyraźnie, że nawet w przypadkach obiektów, gdy bieg jałowy wydaje się być gorszym (obiekt D1), bez modułu SI straty energetyczne byłyby większe. Niewielka liczba wstrzymanych wyłączeń i niski wskaźnik R_{FN} będzie świadczyć o prawidłowym działaniu modułu SI.

Osiągnięte oszczędności energii zależą w dużym stopniu od budowy samego układu i chwilowego zapotrzebowania w trakcie badania. W niektórych przypadkach osiągnięto redukcję aż o 83% (obiekt D4), przez efekty w granicach od 66% do 18%, po dwa obiekty bez statystycznej poprawy. Tutaj ponownie najlepiej zweryfikować wykresy pracy przed i po wdrożeniu, by zobaczyć pozytywne zmiany w działaniu, których statystyka nie oddaje.

Na koniec odnosząc się do obiektu D1 oraz jego najgorszego wyniku, gdzie można by podejrzewać, że moduł SI pogorszył bieg jałowy o 17%, przy braku błędów operacyjnych, w praktyce po analizie trendów stwierdzono, że układ działał prawidłowo. W układzie prowadząca maszyna z przetwornicą częstotliwości na minimalnych obrotach

dostarczała zbyt dużo powietrza i często nie była potrzebna. Układ rekomendował wyłączenie, a gdyby tego nie robił, bieg jałowy stacji osiągnąłby większe wartości. Moduł SI nie doprowadził przy tym do spadku ciśnienia.

Ponieważ zakłady dysponują często własnym systemem pomiarowym, najbardziej miarodajnym sposobem wykazania oszczędności byłyby dwa czynniki badane przed i po wdrożeniu:

- w większości wypadków — zestawienie comiesięcznego procentu biegu jałowego dla całej stacji,
- w przypadku pełnego opomiarowania sprężarkowni — wielkość kWh energii elektrycznej zużytych na wyprodukowanie np. 1000 Nm^3, tylko tutaj należy również uwzględnić czy współczynnik powinien brać pod uwagę też cykle pracy systemu np. osuszania.

Również dobrą praktyką byłoby wykonanie rzetelnych pomiarów dostarczanego sprężonego powietrza przez każdą z maszyn i zestawienie ich z pobieraną energią elektryczną, tak by upewnić się co do ich rzeczywistej wydajności. Zauważone tutaj odstępstwa będą wpływać negatywnie na cały system sterowania

6.3 Kiedy zastosowanie uczenia maszynowego/głębokiego jest uzasadnione

W ostatnich latach w związku z coraz większymi wymaganiami dotyczącymi efektywności energetycznej oraz narastającą świadomością zakładów produkcyjnych budowa układów sprężonego powietrza zaczęła ewoluować. Coraz częściej w układach pojawia się przynajmniej jedna maszyna wyposażona w przetwornicę częstotliwości często bazująca na silniku z magnesami trwałymi (grupa B). Częściej można się również spotkać z układami bazującymi na samych sprężarkach falownikowych (grupa C). Natomiast nadal można spotkać układy oparte wyłącznie o maszyny z rozruchem bezpośrednim. W związku z takimi zmianami pojawia się pytanie: czy dla danego obiektu użycie wsparcia uczenia głębokiego może przynieść jakiekolwiek korzyści. Odpowiedź wymaga przeanalizowania przypadku, którego dotyczy pytanie, więc podstawowa odpowiedź zależy od rozpatrywanego przypadku.

W przypadku układów grupy C możliwości oszczędności energii są w znacznym stopniu wykorzystane. Przystępując do optymalizacji startowy poziom biegu jałowego utrzymywał się na poziomie ok. 2%, po wprowadzeniu systemu wsparcia w związku ze zmianą profilu produkcji, wzrósł do 2,6%. Układ zbudowany wyłącznie z maszyn z silnikami z magnesami trwałymi wykorzystywał szerokie spektrum regulacji, minimalizując liczbę niepotrzebnych startów i adaptując się do zmiennego zapotrzebowania. Jedynym

polem gdzie, można było wspomóc sprężarkownię, to momenty mniejszego zapotrzebowania, gdy sprężarka pracując na minimalnych obrotach się odciążała. Przełożyło się to na rekomendacje wydawane przez moduł SI, który podejmował średnio tylko dziewiętnaście decyzji dziennie (ok. dziesięciokrotnie mniej niż w przypadku innych obiektów). Ponieważ sytuacje odciążeń były rzadkie, pojawiały się błędy FN spośród wydawanych rekomendacji (ok. jedna błędna decyzja na jeden i pół dnia rekomendacji). W związku z tym przed wdrożeniem należy sprawdzić, czy energia tracona na bieg jałowy uzasadnia koszt wdrożenia modułu SI. Zakładając, że układ pracuje prawidłowo, to potencjał optymalizacyjny w takim wypadku może być niewielki. Przy tak niewielkiej liczbie decyzji wskaźniki bezpieczeństwa będą sygnalizować niebezpieczeństwo, które będzie bazowało na niewielkiej liczbie sytuacji decyzyjnych. Analizując działanie poszczególnych maszyn, można prawdopodobnie znaleźć pole do innej optymalizacji, ale nie stanowi ona przedmiotu niniejszej rozprawy.

Kolejną popularną w przemyśle grupą sprężarkowni są układy z jedną maszyną z przetwornicą częstotliwości. Rozpatrywane obiekty cechuje duża dynamika i w niektórych przypadkach użycie modułu SI daje wymierne korzyści. Obiekty D4 oraz D2 przed wdrożeniem cechowała duża składowa biegu jałowego w czasie pracy stacji, co związane jest z dużą zmiennością zapotrzebowania. Sterownik nadrzędny nie dysponuje wystarczającą pamięcią poprzednich zachowań, więc reaktywne sterowanie regułowe nie jest w stanie odkryć wzorców zapotrzebowania i pamiętać o nich w przyszłości. Wspomniane dwa obiekty są dużymi beneficjentami zaproponowanego rozwiązania, ale osiągnięte korzyści wynikają z profilu produkcji. Najłatwiej sieci dostrzec wzorce, gdy cykl pracy jest powtarzalny, a dana maszyna jest prowadząca. Następnym krokiem w celu większych oszczędności byłaby wymiana kolejnej maszyny na nowoczesną sprężarkę z silnikiem z magnesami trwałymi i płynną regulacją.

W przypadku obiektu B4 nie zdecydowano się na wdrożenie w związku z niedopasowaniem proponowanej metodyki. Obiekt składa się z pięciu sprężarek, a załączenie kolejnych maszyn odbywa się sekwencyjnie. W takim wypadku, jeżeli nastąpiło wyłączenie jakiejś maszyny, to w jej miejsce zacznie pracować pierwsza dostępna, przy kolejnej okazji powróci maszyna podstawowa. O ile tutaj maszyna falownikowa pracuje cały czas dostosowując się do poboru, to nie udało się uchwycić korelacji pomiędzy przebiegiem ciśnienia a uruchamianymi maszynami. Analizując przebiegi czasowe, problemem nie jest zbyt niska częstotliwość próbkowania czy zbyt mała liczba zdarzeń. Stopień nieprzewidywalności załączanych maszyn prowadzi do okazjonalnych spadków ciśnienia. Tutaj zaleceniem byłoby wydłużenie biegu jałowego, by wyeliminować spadki ciśnienia, a następnie przeprowadzić uczenie na nowych zebranych danych.

Przed wdrożeniem proponuje się rozważenie następnych aspektów sprężarkowni:

- procent biegu luzem — wyznaczony dla całej stacji, informuje o tym, jak duży jest dostępny potencjał oszczędnościowy, przy kilkuprocentowym biegu jałowym

wdrożenie przynosi niewielkie korzyści, jednak przeliczenie godzin biegu jałowego na liczbę kWh traconych miesięcznie jednoznacznie informuje zakład o stratach,

- poprawność działania sterownika nadrzędnego — aby skutecznie zaobserwować zależności pomiędzy zapotrzebowaniem a reakcjami sterownika nadrzędnego, wymagane jest, by sterownik pracował w sposób prawidłowy, wszystkie maszyny wchodzące w skład układu reagowały na komendy, parametry sterownika były dobrane tak, by nie generowały niedeterministycznych zachowań, obiekt wytrenowany na nieprawidłowo działającym układzie będzie odwzorowywał jego wady,

- priorytety załączenia maszyn — czy układ pracuje w trybie wyrównywania godzin, czy ze względu na różne moce narzucona jest określona kolejność, czy występuje maszyna lub maszyny prowadzące, czy występuje maszyna, której wyłączenie jest zabronione (np. wyposażona w wymiennik ciepła), deterministyczne reguły załączania maszyn są łatwiejsze do modelowania,

- obecność dodatkowych algorytmów — czy w układzie obecny jest np. tryb pracy, który celowo uruchamia maszynę, by zachować ją w gotowości w przypadku niskich, zimowych temperatur, wyłączenie pochodzące z modułu SI wówczas przeszkodzi w dogrzewaniu,

- okres rozprężania maszyn — tutaj proponowanym dodatkowym założeniem jest, by horyzont predykcji był większy lub równy od okresu rozprężania, dzięki czemu maszyna jest gotowa do startu najszybciej jak to możliwe,

- długość pozostawania maszyn na biegu jałowym — jeżeli czas utrzymywania na biegu jałowym jest krótki, krótszy niż horyzont predykcji, to brakuje potencjału oszczędnościowego,

- liczba zdarzeń odciążenia — jaki jest rząd wielkości pojawiających się zdarzeń, czy uzasadnia wprowadzenie modułu SI,

- wierność odwzorowania procesu — czy dane zapisane dla trenowania oddają rzeczywisty przebieg w porównaniu z bieżącymi danymi procesowymi, ile próbek wypada na jedno zdarzenie odciążenia.

Zaproponowane kryteria są ze sobą powiązane, więc powinny być rozpatrywane całościowo.

6.4 Porównanie z literaturą i wskazanie wkładu własnego

Założeniem niniejszej pracy było przede wszystkim wprowadzenie systemu bazującego na badaniach, który będzie działał w czasie rzeczywistym dostarczając predykcyjne decyzje sterujące. Publikacje dotyczące sprężonego powietrza, w umiarkowanym stopniu poruszają tematykę samego działania sterowników nadrzędnych [25], zalecając jego użycie by zsynchronizować pracę maszyn. Pomijanym aspektem też często jest kwestia maszyn, których ograniczenia należy wziąć pod uwagę przy projektowaniu, by nie doprowadzić do ich wcześniejszego zużycia lub awarii. W przypadku uczenia głębokiego, czy podejścia statystycznego, dostępne publikacje w oparciu o liczniki sprężonego powietrza lub energii elektrycznej dokonują predykcji zużycia. Wówczas można uzyskać charakterystykę danego medium w horyzoncie godzinowym lub dobowym [23]. Takie podejście użyteczne przy planowaniu nie jest pomocne przy decyzjach podejmowanych w czasie rzeczywistym. Odchyłka błędu RMSE, która przy ewaluacji wydaje się niewielka w przypadku podejmowania decyzji może prowadzić do zdarzeń klasy FN.

W niniejszej rozprawie zaproponowano podejście, które bazuje na standardowych sygnałach obciążenia sprężarek w połączeniu z pomiarem ciśnienia. Następnie dane dostępne dla uczenia zostały zgrupowane w liczbie próbek odpowiadającej horyzontowi predykcji. Dzięki temu, że horyzont jest krótki, predykcja również trwa bardzo krótko i dostarcza szybko decyzji wspierającej system sterowania. W momencie odciążenia sieć otrzymuje aktualne okno i przewiduje kolejne stany (jak się okazało, w wybranych przypadkach w sposób niezwykle skuteczny). Dane liczników często dostępne są tylko po stronie zakładu przemysłowego i nie są widoczne dla systemu sterowania.

Często pojawiają się publikacje dotyczące predykcyjnego utrzymania ruchu [32], [52], [66], których zadaniem jest detekcja anomalii oraz przewidywania awarii na bazie dodatkowego, kosztownego opomiarowania, takiego jak pomiary drgań, czy dodatkowe pomiary temperatury, ciśnień czy prądu. Tutaj również wypada podkreślić, że sprężarki wyposażone w sterowniki mikroprocesorowe same w sobie mają szereg regułowych, często bardzo zaawansowanych mechanizmów zabezpieczeń bezwzględnie chroniących komponenty maszyn. W praktyce przemysłowej wykrycie awarii na bazie pomiarów wibracji dla sprężarek śrubowych jest niezwykle trudne. Niniejsza rozprawa szuka natomiast optymalnego momentu wyłączenia maszyny. Takie podejście wymaga odpowiedniego zgrupowania danych oraz wprowadzenia dedykowanych wskaźników oceny, które będą również śledzić koszty najbardziej niebezpiecznych decyzji — takie właśnie rozwiązanie zostało zaproponowane.

W tym miejscu wypada rozróżnić jakie elementy systemu zostały zaczerpnięte z literatury, a jakie stanowią autorskie rozwiązania niniejszej rozprawy.

Wykorzystanie rozwiązań pochodzących z literatury:

- użycie rekurencyjnych sieci LSTM oraz następnie przejście do sieci dwukierunkowych BiLSTM pojawiających się w kontekście szeregów czasowych,
- adaptacja mechanizmów uczenia wrażliwego na koszt (ang. *cost sensitive learning*) przy koncepcji tworzenia asymetrycznej funkcji straty,
- wykorzystanie optymalizacji bayesowskiej dla potrzeb doboru hiperparametrów,
- walidacja nauczonego modelu metodą OOT *out-of-time*,
- wykorzystanie koncepcji dryftu koncepcyjnego, odzwierciedlającego stopniowe niedopasowanie modelu do profilu produkcyjnego.

Zaproponowane autorskie rozwiązania:

- zgrupowanie danych w postaci krótkich sekwencji, dzięki czemu przy podejmowaniu decyzji, sieć nie musi otrzymywać danych np. z miesiąca poprzedzającego, co przyspiesza obliczenia i umożliwia działanie systemu w czasie rzeczywistym,
- zaproponowanie funkcji straty `mse_bce` wykorzystującej połączenie składowej regresyjnej oraz klasyfikacyjnej,
- zaproponowanie asymetrycznej funkcji straty `industrial`, w której koszt błędu krytycznego oraz gradient ciśnienia wchodzą do kryterium wyboru modelu produkcyjnego,
- system wskaźników oceny wyznaczanych dla pojedynczych zdarzeń decyzyjnych, zamiast popularnego podejścia porównującego próbki, z rozdzieleniem klas błędów na błędy krytyczne FN oraz błędy niekrytyczne utraconego potencjału oszczędnościowego FP,
- szereg wskaźników określających stopień bezpieczeństwa proponowanego wdrożenia,
- wskaźnik nadzoru operacyjnego, który weryfikuje wydane decyzje już po wdrożeniu,
- parametry czasowe zabezpieczające działanie układu dla systemów o wysokiej zmienności,
- wytyczne dotyczące oceny wierności odwzorowania procesu w danych uczących wraz z metodą weryfikacji.

Wkład pochodzący z wyników badań eksperymentalnych oraz wdrożenia:

- system przetestowany został na produkcji przez 806 dni na ośmiu wybranych obiektach, z opisem przypadków, gdy metoda była nieskuteczna,

- wyznaczone granice stosowalności metody — na bazie trzech popularnych konfiguracji sprężarkowni spotykanych w przemyśle obejmujących 12 rzeczywistych obiektów przemysłowych,
- zidentyfikowanie czterech mechanizmów degradacji, które zostały zauważone w trakcie eksploatacji,
- studium przypadku trudnego obiektu z wytycznymi dotyczącymi podjętych kroków prowadzących do wdrożenia.

Tutaj należy podkreślić, że opis przypadków, gdy metoda jest nieskuteczna, jest szczególnie ważny, ponieważ pozwala uniknąć wdrożeń, gdy metoda nie będzie działać.

6.5 Ograniczenia badań

W celu poprawnej interpretacji wyników badań należy podkreślić ograniczenia, z którymi wiążą się przeprowadzone badania. Poniższe punkty przedstawiają je w uszeregowany sposób.

1. Praca dotyczy tylko układów zbudowanych na sprężarkach śrubowych. Nie uwzględniono układów osuszania. Nie dysponowano wiedzą o instalacji odbiorczej takiej jak jej pojemność, liczba zbiorników, liczba odbiorców itd.
2. Sterownik nadrzędny sprężarkowni musi dysponować pomiarem ciśnienia odpowiadającego zapotrzebowaniu oraz mieć możliwość sterowania poszczególnymi maszynami w zakresie obciążenia.
3. Proponowana metodyka przeznaczona jest dla sprężarkowni zawierających od dwóch do czterech sprężarek, przy czym najlepsze rezultaty uzyskano dla trzech maszyn. Wyniki obiektu B4, złożonego z pięciu sprężarek, wskazują granicę stosowalności metody (sekcja 4.6.4).
4. Nie dysponowano wiedzą o cyklach produkcyjnych zakładów. Budowane modele opierały się na dostępnych danych procesowych.
5. Asymetryczna kara w funkcji straty realizowana jest na etapie selekcji modelu.
6. Nie dysponowano danymi liczników sprężonego powietrza i energii elektrycznej. W związku z tym nie można pokazać jednoznacznie liczby oszczędzonych kWh energii elektrycznej. Tutaj odniesieniem pozostaje procent biegu jałowego całej stacji sprężarek, jednak przy zmianie profilu produkcji, lub z innych przyczyn może występować sytuacja, gdy wskaźnik biegu jałowego wzrasta, ale bez modułu SI byłby jeszcze wyższy (obiekt D1). W przypadku dysponowania tymi danymi zaleca się wprowadzenie wskaźnika liczby zużytych kWh na określoną wielkość wyprodukowanego powietrza np. 1000 Nm^3.

7. By metoda działała, dane treningowe muszą odwzorowywać normalny stan produkcji — przestoje technologiczne, okres wakacyjny powodujący spowolnienie produkcji, awarie maszyn, zbyt mała liczba zdarzeń, może prowadzić do tego, że sieć nauczy się nieprawidłowych wzorców. Dane muszą zostać zweryfikowane i ocenione przed uczeniem.
8. 12 obiektów badawczych, osiem wdrożonych stanowi reprezentantów krajowego przemysłu, jednak ta liczba jest zbyt mała, by móc zaproponowane wnioski uogólnić na szerszą grupę. Zapotrzebowanie na sprężone powietrze często znacząco różni się od siebie, więc tutaj proponuje się podejście indywidualne.
9. Liczba zdarzeń decyzyjnych wpływa na odwzorowanie charakterystyki przez wskaźnik S_{FN}, zdecydowano się na podawanie liczb bezwzględnych zdarzeń, by była widoczna skala, ale wskaźnik wyznaczony na niewielkiej liczbie zdarzeń powoduje, że pojedyncze zdarzenie będzie zmieniać go o kilka punktów procentowych.
10. Wskaźnik S_{FN} odnosi się do okien, gdzie dana maszyna wróciła do stanu obciążenia. S_{FP} natomiast odnosi się do utraconych możliwości oszczędności FP dla wszystkich okien decyzyjnych. Przedmiotem rozprawy było przede wszystkim uniknięcie sytuacji niebezpiecznych FN a w drugiej kolejności sprawdzenie potencjału oszczędnościowego FP. W przypadku wskaźników również zdecydowano się na podanie wartości bezwzględnych by wiadoma była skala, której dotyczą wskaźniki.
11. Eksperymenty porównawcze wykonano na stałych hiperparametrach sieci, co umożliwiło porównanie wyników pomiędzy obiektami. Jednak oznacza to, że dana konfiguracja mogła osiągnąć lepsze wyniki przy ich innych wartościach.
12. Interpretowalność decyzji modelu jest ograniczona. Przy przeglądaniu wykresów porównawczych często niemożliwe było dojście do tego, skąd wzięła się adekwatna decyzja sieci neuronowej. W związku z tym wprowadzono jasne regułowe zabezpieczenia, które uspokajają personel utrzymania ruchu odpowiedzialny za proces.
13. Eksperymenty wdrożeniowe prowadzono reaktywnie. W przypadku wykrycia zachowania wymagającego interwencji dla jednego obiektu, reguła była rozszerzana na wszystkie pozostałe. Tak więc nie było to badanie kontrolowane, ale zdobywanie cennego doświadczenia wdrożeniowego.
14. Wskaźnik nadzoru operacyjnego R_{FN} obserwuje wyłącznie przypadki, gdy po wydanej decyzji wyłączenia, maszyna wróciła do stanu obciążenia, nie analizuje innych przypadków, które mogą wystąpić, przez co stanowi dolny próg liczby niewłaściwych decyzji, a pełna ocena wymagałaby kolejnych wskaźników.
15. Kryterium optymalizacji może również zostać wymuszone przez odbiorcę. Odbiorca dysponujący opomiarowaniem może żądać określonej kolejności maszyn, która będzie lepiej się sprawdzać w danych warunkach. Może również zdecydować o stratach energetycznych w zamian za dostawy stabilniejszego ciśnienia, jak to

miało miejsce dla obiektu D3.

16. Strona popytowa dysponuje również wiedzą o tym na ile stabilne są dostawy sprężonego powietrza. Poza obiektem D3 nie nastąpił kontakt z jego strony w celu poprawienia parametrów systemu.
17. Świadomie zrezygnowano z podejścia stochastycznego [79] na rzecz asymetrycznego w prognozowaniu szeregów czasowych omówionego w sekcji 2.4.3.

6.6 Znaczenie w kontekście założeń Przemysłu 4.0

Proponowane rozwiązanie wykorzystuje istniejące już dane procesowe w postaci pomiaru ciśnienia oraz stanów obciążenia maszyn (wraz z innymi niezbędnymi dla działania sygnałami dwustanowymi), bez rozbudowywania istniejącej infrastruktury. Wykorzystanie danych w taki sposób wpisuje się w koncepcję Przemysłu 4.0 [7], [8], gdzie wykorzystanie wymienionych danych procesowych wykracza poza zakres, jaki był przyjęty na etapie budowy instalacji. Być może dane potraktowane jako zasób umożliwiający rozwój istniejącego systemu w przyszłości zostaną rozbudowane o dodatkowe wskaźniki na bazie rozwijanych instalacji w ramach realizowania zaleceń audytów energetycznych.

Wdrożenia produkcyjne często realizowane są w instalacjach, które zostały wdrożone w poprzedniej epoce technologicznej i których wymiana na najnowocześniejsze rozwiązania nie jest uzasadniona, jeżeli nie przynosi wymiernego zwrotu poniesionych kosztów w określonym horyzoncie. Nie zawsze dostępne są zasoby, które umożliwiłyby integrację wszystkich warstw zakładu produkcyjnego, rozpoczynając od strony procesowej aż do warstwy zarządczej w ramach np. systemów ERP (ang. *Enterprise Resource Planning*). Zaproponowane w niniejszej rozprawie rozwiązanie nie jest inwazyjne i stanowi rozszerzenie istniejącego sterowania sprężarkowni. W przypadku awarii komunikacji z modułem SI, układ nadal pracuje według swoich reguł, przez co nie ma ryzyka utraty stałości dostaw sprężonego powietrza.

Kolejnym elementem wiążącym założenia Przemysłu 4.0 z zaproponowanym rozwiązaniem jest wykorzystanie archiwizowanych danych procesowych do oceny jakości podejmowanych decyzji. Baza danych zawierająca dane historyczne, przydatne przy diagnostyce działania układu i np. przyczyn awarii, została wykorzystana do trenowania modeli uczenia głębokiego oraz śledzenia wskaźników weryfikujących działanie modułu SI. Obieg przetwarzania sygnałów zamknął się zatem w ramach już posiadanej infrastruktury.

Ryzyko zachwiania dostaw sprężonego powietrza, określone błędem FN zostało odzwierciedlone na trzech spójnych poziomach:

- funkcji straty dla uczenia, gdzie niebezpieczne zdarzenie FN jest karane,

- wskaźników oceny modelu przed wdrożeniem bazujących na faktycznych decyzjach FN mogących wpłynąć na stabilność ciśnienia,
- mechanizmu weryfikacji po wdrożeniu sprawdzającym działanie układu na produkcji z naciskiem na wykrywanie zdarzeń FN.

Takie rozwiązanie jest łatwe w interpretacji dla personelu utrzymania ruchu i łączy ze sobą wszystkie etapy od badania do wdrożenia.

Wdrożenie również obaliło założenie, z którym przystępowano do treningu, że model powinien być uczony przy wyłączonym module SI, podczas normalnej pracy, tak by zauważyć korelację między poborem a pracującymi maszynami. Zaskakującą tutaj rolę odegrały zabezpieczające deterministyczne reguły. Ich podstawową funkcją było zabezpieczenie stabilności dostaw sprężonego powietrza, podczas gdy w praktyce odzwierciedliły dynamikę układu. Dzięki temu uniknięto sytuacji, gdy model uczyłby się na własnych nieprawidłowych decyzjach. W związku z tym warstwa zabezpieczeń wykroczyła poza swoją pierwotną rolę i stała się składową procesu.

Prawdopodobnie wraz z rozwojem sztucznej inteligencji, integracji wszystkich warstw systemów tworzących zakład produkcyjny, nowatorskiego wyszukiwania wzorców w zebranych danych w przyszłości, w ramach założeń Przemysłu 4.0, uda uzyskać się większą skuteczność niż zaproponowane w niniejszej rozprawie rozwiązanie. Rozprawa zdecydowanie wykazała, że mile widziany byłby dodatkowy czynnik, informujący np. o liczbie aktualnych odbiorców, a modyfikacja sekwencji uczących czy funkcji straty o tą własnością jest niewymagająca. Takie pełne współdziałanie wpisywałoby się w trend Przemysłu 4.0.

7 Podsumowanie

Rozdział stanowi zakończenie rozprawy. Zawiera podsumowanie przeprowadzonych prac, weryfikację tez badawczych, wskazanie oryginalnego wkładu autora, zalecenia praktyczne prowadzące do wdrożenia oraz kierunki dalszych badań.

7.1 Podsumowanie przeprowadzonych prac

Rynek oferuje szereg sterowników nadrzędnych z różnym podejściem do kwestii sterowania zespołem sprężarek. Czołowi producenci z sektora sprężonego powietrza proponują własne sterowniki, które najlepsze efekty osiągają przy współpracy z maszynami własnej produkcji. Firma Marani Sp. z o.o. zaproponowała sterowniki nadrzędne, które integrują w sobie różne typy maszyn. Drugim założeniem była konkurencyjność osiągnięta przez rozwiązania dedykowane konkretnym odbiorcom, która szeroko wykracza poza współpracę z systemami DCS. Najważniejszy był indywidualizm w podejściu do każdego procesu, dzięki czemu każdy sterownik w ofercie firmy jest inny i dostosowany pod potrzeby konkretnego klienta. Mimo to pozostało pewne założenie wspólne z innymi producentami, czyli reaktywne sterowanie bazujące na bieżącym zapotrzebowaniu. W reakcji na zmiany zapotrzebowania, sygnalizowane spadkiem ciśnienia, następuje załączenie sprężarki. Rozbudowane opomiarowanie obejmujące pomiar energii elektrycznej oraz wyprodukowanego powietrza wiąże się z dodatkowymi kosztami. Na bazie tych obserwacji, wraz z rozwojem metod uczenia głębokiego, zauważono możliwość rozwoju istniejących układów.

Problem badawczy został określony jako opracowanie metodyki predykcyjnego wsparcia układów sterowania nadrzędnego z założeniem wykorzystania standardowych dostępnych dla sterownika nadrzędnego sygnałów, bez rozbudowywania instalacji o dodatkowe opomiarowanie. Układ opiera się na aktualnym pomiarze ciśnienia względnego odzwierciedlającego stan zapotrzebowania oraz szereg sygnałów dwustanowych z prowadzącą rolą sygnału obciążenia, które są niezbędne dla realizacji sterowania nadrzędnego. Zidentyfikowano największe zagrożenie dla stałości dostaw sprężonego powietrza, jakim jest wyłączenie sprężarki, która była potrzebna do obciążenia, co może prowadzić do spadku ciśnienia i zachwiania stabilności produkcji. Drugim niebezpieczeństwem, jest sytuacja pominięcia okazji do wyłączenia, która prowadzi częściej do utraty potencjalnych oszczędności, a nie zagraża stabilności dostaw.

Przyjęte założenie problemu badawczego wpisuje się w koncepcję Przemysłu 4.0, w której dane procesowe gromadzone przez systemy SCADA stają się zasobem wykorzystywanym do sterowania pracą sprężarkowni i wzbogacenia jej o nowe funkcje. Proponowane rozwiązanie wprowadza elementy inteligencji do istniejących układów sterowania nadrzędnego bez ingerencji w infrastrukturę sprężarkowni (sekcja 6.6).

Proponowana metodyka obejmuje wszystkie etapy tworzenia systemu od fazy zbierania danych po wdrożenie łącznie z nadzorem produkcyjnym, czyli:

- właściwy dobór okresu próbkowania oddający dynamikę obiektu,
- dobór architektury dwukierunkowej sieci rekurencyjnej wraz z horyzontem predykcji,
- dwa warianty funkcji straty z uwzględnieniem kosztu procesowego i metodyką doboru jej parametrów,
- system wskaźników bezpieczeństwa dla faktycznych zdarzeń decyzyjnych,
- regułowa warstwa zabezpieczeń,
- system nadzoru operacyjnego.

Weryfikacja proponowanego rozwiązania została przeprowadzona dwuetapowo:

- etap badań — wyznaczono trzy popularne i spotykane w przemyśle grupy sprężarkowni, dla których na zebranych danych historycznych pochodzących z dwunastu sprężarkowni zbadano proponowaną metodykę,
- etap wdrożenia — wybrano cztery obiekty z poprzedniej grupy (z której jeden stanowił inspirację dla badań) oraz cztery kolejne obiekty, łącznie przepracowały na produkcji 806 dni podejmując 133 tysiące zweryfikowanych decyzji.

Podsumowaniem wdrożenia niech będzie fakt, że żadna z decyzji nie doprowadziła do spadku ciśnienia, a błędne decyzje utrzymały się na poziomie 0,085%. Gdyby nie przeprowadzenie wdrożenia, nie zaproponowano by warstwy zabezpieczającej, dodatkowych parametrów czasowych oraz efektywnego wskaźnika nadzoru operacyjnego. Te rozwiązania pojawiły się w odpowiedzi na obserwowane działanie systemu. W związku z tym proponuje się by badania symulacyjne były weryfikowane w rzeczywistych warunkach produkcyjnych, a nie pozostawały jedynie na poziomie symulacji.

7.2 Weryfikacja tez badawczych

Teza pierwsza

Predykcyjne wsparcie nadrzędnego procesu sterowania zespołem sprężarek z wykorzystaniem uczenia głębokiego obliguje do zindywidualizowanego podejścia, które wymaga

doboru następujących elementów: okresu próbkowania zbioru uczącego (odzwierciedlającego dynamikę obiektu), funkcji straty (wraz z ustaleniem jej parametrów oraz kosztów procesowych), horyzontu predykcji, poziomów reguł bezpieczeństwa (zabezpieczenie stabilności dostaw).

Teza potwierdzona. Przy proponowanej metodyki nie wykazano możliwości wprowadzenia uniwersalnej funkcji straty, która sprawdzi się dla każdego rozpatrywanego obiektu. Każda z zaproponowanych składowych musi być dobierana indywidualnie. W przypadku niespełnienia tego założenia należy liczyć się z problemami.

W rozprawie zawarto studium przypadku dla obiektu D3, dla którego domyślny **okres próbkowania** parametrów nie oddawał wiernie dynamiki obiektu (przy braku pominięcia jakichkolwiek istotnych tranzycji odciążenia). Wskaźnik S_{FN} opisujący model nauczony na domyślnych 20 s danych był na, wydawać by się mogło, doskonałym poziomie 3,5% (jeden z najlepszych wyników w zbiorze). Przy czym na produkcji w ciągu pierwszych dziesięciu minut podjął dwie ryzykowne decyzje, które mogły prowadzić do znacznego spadku ciśnienia. Zwiększenie częstotliwości próbkowania dostarczyło model, który przepracował na produkcji 35 dni do decyzji klienta o zmianie profilu produkcji.

Funkcje straty zostały zbadane w studium ablacji. Dla obiektu A3 wskaźnik S_{FN} poprawił się z poziomu 50,0% do 35,7%, dla obiektu B1 wzrósł aż do 94,0%, dla obiektu C3 poziom został dokładnie ten sam, a dla obiektu D3, poza studium ablacji, umożliwił wdrożenie. Przytoczone studium potwierdza założenie, że nie istnieje uniwersalna konfiguracja funkcji straty optymalna dla wszystkich obiektów, a o wdrożeniu będzie decydować osiągany poziom wskaźników bezpieczeństwa.

Horyzont predykcji jest kwestią indywidualną zależącą silnie od rozpatrywanego obiektu, z czego można zaobserwować następujące przypadki wraz ze wzrostem długości okna obserwuje się następujące zachowanie wskaźnika S_{FN}:

- spadek (obiekt A1, B4),
- brak zmian (obiekt A3, C3),
- wzrost (obiekt A4, C4),
- inna tendencja (wzrostowa lub malejąca) bez wyraźnie utrzymującego się trendu dla każdego z horyzontów.

Nasuwa się więc wniosek, że decydującym w tym przypadku będzie nie horyzont, lecz charakterystyka obiektu.

Reguły bezpieczeństwa zostały wprowadzone w wyniku wdrożenia, ich poziom definiowany jest indywidualnie dla każdego zakładu. Wdrożenie wykazało też potrzebę wprowadzania dodatkowych parametrów czasowych w celu odpowiedniej obserwowalności procesu oraz złagodzenia skutków stanów nieustalonych.

Teza druga

Przydatność opracowanego modelu predykcyjnego wymaga wskaźników oceny bazujących na rzeczywistych, pojedynczych zdarzeniach decyzyjnych z ukierunkowaniem na koszt błędu procesowego.

Teza potwierdzona. Teza zostanie omówiona na trzech przykładach ją potwierdzających.

Przy prowadzeniu pierwszych badań oraz w dostępnych publikacjach wyznaczane wskaźniki bazowały na wiernym odwzorowaniu poszczególnych próbek trendu. Taka informacja nie odróżnia od siebie modelu, który będzie popełniał błędy krytyczne od modelu działającego prawidłowo. W związku z tym zdecydowano się na wprowadzenie wskaźników bezpieczeństwa. Warianty jednokierunkowe w studium ablacji zdecydowały się sygnalizować bezwarunkowy powrót obciążenia (zerowy wskaźnik S_{FN}), przez co model nigdy nie rekomendował wyłączenia maszyny.

Wskaźniki rozdzielono na dwie grupy: wskaźnik S_{FN} odzwierciedlający ryzyko zachwiania dostaw sprężonego powietrza, oraz wskaźnik S_{FP}, który informuje o utracie potencjału oszczędnościowego. Rozpatrzenie obu wskaźników daje informacje niezbędne dla adekwatnej oceny obiektu.

Kolejną funkcją wskaźników jest ocena dopuszczenia obiektu do wdrożenia. Po rozpatrzeniu wskaźników wraz ze sprawdzeniem trendów predykcyjnych nie dopuszczono do wdrożenia obiektu B4 (zagregowany S_{FN} na poziomie 96,7%, zagregowany wskaźnik S na poziomie 83,2%). Sama wstępna inspekcja zachowania układu wskazywała na możliwe wsparcie modułem SI, ale w praktyce okazało się, że ryzyko jest zbyt duże.

Wdrożenie wykazało też konieczność wprowadzenia wskaźnika nadzoru operacyjnego R_{FN}, którego zadaniem jest wykrywanie degradacji wprowadzonego modelu. Założenie działania wskaźnika sprowadza się do weryfikacji następczej wydanych rekomendacji. Natomiast wyznaczane wskaźniki bezpieczeństwa z danych zebranych nie używały dodatkowych reguł zabezpieczających, nie brały też pod uwagę dostępnej rezerwy ruchowej. W praktyce okazało się, że wskaźnik S_{FN} dla obiektu D1 na poziomie 100% spotkał się z jedną błędną decyzją na wydane ponad dwadzieścia dwa tysiące rekomendacji wyłączenia.

Teza trzecia

Zastosowanie dwukierunkowej sieci rekurencyjnej BiLSTM zapewnia lepszą zdolność modelowania procesu niż wykorzystanie rekurencyjnych sieci jednokierunkowych LSTM oraz GRU.

Teza potwierdzona. Studium ablacji wykazało, że sieć dwukierunkowa przy zachowaniu tych samych parametrów jak inne sieci, osiąga wyższą skuteczność. Wskaźnik S (odwzorowanie prawidłowych decyzji w zbiorze) osiąga:

- dla obiektu A3 — 99,2% BiLSTM wobec 6,0% LSTM,
- dla obiektu B1 — 87,5% BiLSTM wobec 51,2% LSTM,
- dla obiektu C3 — 99,8% BiLSTM wobec 32,4% LSTM.

Studium potwierdziło również zbliżony charakter sieci LSTM oraz GRU co jest zgodne z literaturą. W rozpatrywanym przypadku sieci te przyjęły zasadę prawie ciągłego przewidywania obciążenia. Sieć dwukierunkowa, rozpatrując zdarzenia w obu kierunkach, uzyskuje szerszy kontekst. Jest zatem w stanie odróżnić sytuację potrzeby pozostania pod obciążeniem od bezpiecznego wyłączenia. Kierunkowość przetwarzania zatem jest kluczowa.

Uzyskany wynik jest związany ściśle z zaproponowaną konfiguracją sieci neuronowej oraz dedykowaną funkcją straty i w wypadku zaproponowanej metodyki jest najlepszym wyborem, co nie wyklucza użycia sieci jednokierunkowych w innych niż zaproponowana konfiguracjach.

Teza czwarta

Zaproponowana metodyka zakłada wykorzystanie już istniejącej infrastruktury systemu sterowania nadrzędnego bazującej na sterowniku PLC, a zakres jej stosowalności jest możliwy do wyznaczenia przed wdrożeniem. Po wdrożeniu wymagany jest ciągły nadzór operacyjny dla utrzymania efektów.

Teza potwierdzona. Współpraca z istniejącymi sterownikami PLC sprowadziła się do ustalenia obszaru wymiany, sposobu współpracy oraz przesłania rekomendacji. Nie było potrzeby rozbudowy istniejącego systemu pomiarowego. Komunikacja pomiędzy warstwą inteligentnego wsparcia decyzji sterujących a istniejącym sterownikiem nadrzędnym odbywa się przy pomocy standardowego protokołu przemysłowego. W przypadku zerwania komunikacji sterownik nadrzędny działa nadal na bazie własnych reguł. Zakres stosowalności metody umożliwiają opisane wcześniej wytyczne. Poza procesem kwalifikacji obiektu do wsparcia modułem SI (ocena wstępna i odrzucenie obiektu B4), przedstawiono następujące kryteria dotyczące oceny: podstawowa ocena prawidłowości działania układu sterowania nadrzędnego, udział biegu jałowego w roboczogodzinach całej stacji, wierność odwzorowania procesu, liczba zdarzeń decyzyjnych, obecność maszyn o regularnych cyklach pracy, relacja biegu jałowego i horyzontu predykcji, priorytety załączenia maszyn oraz obecność dodatkowych algorytmów sterownika nadrzędnego, czas rozprężania maszyn.

System nadzoru wdrożonych rozwiązań jest niezbędny w związku z zaobserwowanymi czterema mechanizmami degradacji sieci: czynniki nieprzewidywalne, w tym awarie maszyn lub powiązanej infrastruktury, dryft koncepcyjny wynikający ze zmiany profilu produkcji, zanikanie możliwości korzyści energetycznych oraz brak potencjału

optymalizacyjnego obiektu. Trend zmian wskaźnika R_{FN} w czasie pozwala na ustalenie prawdopodobnego mechanizmu degradacji, będąc systemem wczesnego ostrzegania, zachęcającego do indywidualnego rozpatrzenia danego przypadku.

Stosowalność metody związana jest z określonymi warunkami: spośród dwunastu badanych obiektów, trzy nie osiągnęły wyników wystarczających dla wdrożenia, z ośmiu wdrożonych obiektów dwa zostały wyłączone w trakcie (w związku z czynnikami nieprzewidywalnymi). Wyniki te określają zakres stosowalności metody.

7.3 Oryginalny wkład autora

Rozwiązania autorskie oraz zaczerpnięte z literatury zostały przedstawione w sekcji 6.4. Poniżej przedstawiono oryginalny wkład niniejszej rozprawy.

System oceny bazujący na zdarzeniach decyzyjnych. Chcąc odzwierciedlić rzeczywiste działanie systemu sprężonego powietrza, zdecydowano się na ocenę opartą o okna decyzyjne i wystąpienia tranzycji, co odzwierciedla rzeczywiste wystąpienie potrzeby wsparcia decyzjami predykcyjnymi. Rozróżniono również klasy błędów, które mogą prowadzić do zachwiania stabilności dostaw sprężonego powietrza lub do ominięcia potencjału oszczędności energii elektrycznej. Zaproponowany zestaw wskaźników bazujący na decyzjach umożliwia ocenę bezpieczeństwa przed wdrożeniem, która wykracza poza klasyczne podejście rozpatrywania każdej próbki indywidualnie.

Funkcje straty dostosowane do procesu z uwzględnieniem kosztu procesowego. Zaproponowane funkcje straty odzwierciedlają proces przez wprowadzenie składowych z określoną wagą, spośród których każda ma pełnić określoną funkcję. Za odzwierciedlenie przebiegu ciśnienia odpowiada składowa regresyjna, za rozkład pracy maszyn — składowa klasyfikacyjna, natomiast asymetryczna kara za błędy krytyczne oraz składowa trendu ciśnienia wchodzą do kryterium wyboru modelu produkcyjnego. Zaproponowano również metodykę doboru wag dla rozpatrywanego obiektu.

Wskaźnik wdrożeniowy nadzoru operacyjnego z mechanizmem weryfikacji następczej. Opracowany mechanizm zapisuje każde wydane polecenie i sprawdza je po upływie ustalonego okna w celu weryfikacji, czy nie została aktywowana reguła zabezpieczająca, która je anulowała. Takie rozwiązanie pozwala na bieżący monitoring jakości decyzji predykcyjnych w warunkach produkcyjnych. Ta sama definicja błędu krytycznego odnosi się do funkcji straty (kara), wskaźników oceny przed wdrożeniem oraz wskaźnika oceny na produkcji.

Sprawdzenie wierności odwzorowania procesu. Przedstawiono przykład, kiedy wskaźniki bezpieczeństwa wyznaczone na danych, które nie odwzorowywały wiernie procesu, nie przełożyły się na rzeczywiste wdrożenie, tylko opisywały zgodność wytrenowanego modelu z zebranymi danymi. Wprowadzono zatem wymóg, zgodnie z którym, stan maszyn musi być powiązany z trendem charakterystyki ciśnienia oraz stanami

obciążenia maszyn. Przy ocenie należy również wziąć pod uwagę liczbę próbek wypadającą na pojedyncze zdarzenie.

Parametry czasowe związane z realizacją decyzji. Dzięki wdrożeniu zidentyfikowano dwa niezależne parametry czasowe niezbędne dla prawidłowego działania układu. Pierwszym jest opóźnienie wywołania predykcji, związane z czasem, jaki jest potrzebny na wskazanie trendu nadchodzącej charakterystyki. Drugim jest parametr czasowy chroniący przed stanem nieustalonym układu w wypadku odciążeń sekwencyjnych — opóźnienie realizacji decyzji.

Lista punktów dopuszczająca obiekt do wdrożenia. Przygotowano listę punktów, która pozwala ocenić sens zastosowania wsparcia predykcyjnego na danym obiekcie. Wzięto pod uwagę czynniki techniczne, jak i ekonomiczne potencjalnych oszczędności energii.

Doświadczalny zakres stosowalności metody. Przebadano osiem obiektów w łącznym okresie 806 dni działania na produkcji. Na bazie zdobytego doświadczenia opisano przypadki, gdy metoda działała skutecznie, jak również cztery zauważone mechanizmy degradacji prowadzące do utraty efektywności (podając również sposób ich rozpoznania).

Test metody na produkcji. Wdrożenie objęło osiem obiektów, cztery spośród badanych grup oraz cztery nowe obiekty. Pokazano przypadki redukcji udziału biegu jałowego, jak i takie, gdzie metoda nie przyniosła korzyści. Podanie przykładów niepowodzeń pozwala na uniknięcie wdrożeń w przypadkach, gdy metoda będzie nieskuteczna.

7.4 Zalecenia praktyczne

Poniżej opisano zalecenia wdrożeniowe dotyczące przeprowadzonych badań oraz doświadczeń związanych z uruchomieniem predykcyjnego wsparcia sterowania sprężarkownią bazujących na sprężarkach śrubowych.

Przed wdrożeniem — sprawdzenie sterowania nadrzędnego. Układ sterowania nadrzędnego musi działać w najlepszy możliwy sposób w ramach dostępnych możliwości, tak by reagować na bieżące zapotrzebowanie. Nauczenie modelu podczas nieprawidłowego działania odwzoruje jego nieprawidłowe zachowanie po wdrożeniu. Czynnikami destabilizującymi mogą być problemy z konfiguracją maszyn lub np. przeprowadzenie trenowania z maszyną wyłączoną z systemu i jej powrót do działania po wdrożeniu (mając model nauczony, że wspomniana maszyna nigdy się nie obciąża).

Sprawdzenie danych treningowych. Należy porównać zarchiwizowane dane z niższą częstotliwością zapisu z przebiegami dostępnymi na sterowniku nadrzędnym lub jeżeli jest taka możliwość techniczna, z zapisami z wyższą częstotliwością. Zapisane przebiegi muszą odzwierciedlać stan obciążenia maszyny w powiązaniu z ciśnieniem

przy zachowaniu liczby próbek utrzymujących obserwację. Nieprawidłowa konfiguracja może oznaczać np. zapis zmiany obciążenia przy poziomie ciśnienia omijającego górną granicę (stan zapisany taki sam próg ciśnienia stan obciążenia, kolejny zapis zbliżone ciśnienie, stan odciążenie).

Sprawdzenie potencjału oszczędności energii elektrycznej. Im wyższy udział biegu jałowego w roboczogodzinach całej stacji, tym większy potencjał oszczędności energii uzyskanych przy proponowanej metodzie. Przy kilkuprocentowym poziomie biegu jałowego stacji wdrożenie nie ma ekonomicznego uzasadnienia, a zaproponowane wskaźniki bezpieczeństwa będą utrzymywać się na wysokich poziomach (z racji niewielkiej liczby zdarzeń).

Sprawdzenie rzeczywistej wydajności maszyn. Zalecane jest wykonanie pomiarów ilości sprężonego powietrza dostarczanego przez pojedyncze sprężarki wraz z pomiarem pobieranej energii elektrycznej. Celem pomiarów jest sprawdzenie, czy wydajność sprężarek jest prawidłowa, czy nie dochodzi do sytuacji, że wydajność maszyny znacznie odbiega od deklarowanej w związku z niezidentyfikowanymi problemami. To działanie ma na celu wyeliminowanie przypadków, gdy załączenia danej maszyny ma niewielki wpływ na cały układ sterowania.

Startowa funkcja straty. Punkt startowy to domyślna funkcja straty z domyślnymi hiperparametrami, które były przyjęte w publikacjach z wykorzystaniem architektury BiLSTM. Zaleca się sprawdzenie przestrzeni hiperparametrów przy użyciu siatki w celu wybrania jej najlepszego wariantu. W przypadku niewystarczającej adekwatności należy użyć zaawansowanej funkcji straty.

Ustalenie horyzontu predykcji. Proponowany horyzont powinien być większy lub równy od czasu rozprężania maszyn wchodzących w skład sprężarkowni. Empirycznie należy sprawdzić warianty od dwóch do pięciu minut i dokonać analizy wskaźników bezpieczeństwa, krytycznych błędów predykcji, błędów utraconych oszczędności ze wzięciem pod uwagę liczby zdarzeń.

Dobór parametrów czasowych. Pierwszy z nich to opóźnienie wywołania predykcji, które musi dać sieci czas na właściwą interpretację kierunku charakterystyki (zaleca się przyjęcie domyślnej wartości niezależnie od częstotliwości próbkowania). Drugi parametr to opóźnienie kolejnych decyzji predykcyjnych w przypadku sekwencji odciążeń co zwiększa bezpieczeństwo układu i powoduje ograniczenie skutków stanu nieustalonego.

Dobór poziomów zabezpieczeń. Rekomendacja wyłączenia sprężarki profilaktycznie musi zostać obwarowana zabezpieczającym poziomem ciśnienia poniżej którego, nie nastąpi wysłanie decyzji sterującej. Rozwiązanie nie ingeruje w logikę sterowania związaną ze sterownikiem nadrzędnym, a zabezpiecza proces w przypadku nagłych spadków ciśnienia. Kolejnym zabezpieczeniem jest liczba załączeń na godzinę każdej

maszyny, której nie wolno bezwzględnie przekroczyć. Oczywiście przy wysyłaniu komend należy wziąć również sygnały dwustanowe informujące o stanie poszczególnych maszyn związane ze specyfiką instalacji klienckiej (np. bezwzględny zakaz wyłączenia, którejś z maszyn).

Monitoring wdrożonego rozwiązania — nadzór operacyjny. Ciągły monitoring wskaźnika prowadzącego nadzór operacyjny wysyłanych decyzji R_{FN} jest niezbędny dla wykrycia degradacji proponowanego rozwiązania. Narastający wskaźnik wskazuje na zmianę profilu produkcji, nagły skok sugeruje awarię infrastruktury, wysoki, utrzymujący się stały poziom oznacza brak możliwości optymalizacji.

Dotrenowywanie modelu podczas produkcji. Zalecane jest dotrenowywanie modelu na liczbie danych odpowiadającej pierwotnemu zbiorowi treningowemu. Dane muszą zostać jednak wcześniej sprawdzone, czy na obiekcie nie występują anomalie (np. zmiana profilu produkcji związana z przerwą wakacyjną). Niezbędna jest zatem ocena liczebności zdarzeń w danych treningowych.

Ustalenie kryterium optymalizacji. Większe oszczędności mogą być związane z większą ilością błędnych decyzji, które mogą prowadzić do spadku ciśnienia do poziomu, który być może nie będzie powodował zatrzymania zakładu, ale może negatywnie wpływać na określone grupy urządzeń. W związku z tym najbezpieczniej ustalić z odbiorcą poziom ciśnienia, poniżej jakiego bezwzględnie nie można dopuścić do spadku. Jego wartość może również ulec zmianie w trakcie eksploatacji, gdy większe straty energii będą dopuszczalne, kosztem stabilniejszych dostaw sprężonego powietrza.

7.5 Kierunki dalszych badań

Przystępując do rozprawy, jasne było, że zróżnicowanie pomiędzy zakładami produkcyjnymi będzie znaczne, przez co konieczne będzie określenie granic stosowalności metody. Inny charakter każdego z obiektów prowadzi do kolejnych pomysłów wartych zbadania dotyczących indywidualnego podejścia. W trakcie badań pojawiły się nowe kierunki rozwoju, które warto byłoby eksplorować. Poniżej wymieniono kilka kierunków, które chciano by zbadać w przyszłości.

Identyfikacja cyklu pracy zakładu i dobranie modelu. Zauważono, że część obiektów pracuje w dwóch cyklach produkcyjnych. Pojawiają się okresy stabilnej pracy z niewielką liczbą tranzycji oraz okresy z ich dużą liczbą. Czasami charakterystyka jest powtarzalna, czasami nie [89]. Model wytrenowany na obu dynamikach nabiera cech pośrednich. Prawdopodobnie rozgraniczenie tych przypadków mogłoby prowadzić do adekwatniejszych wyników i mogło doprowadzić do wdrożenia obiektów, które nie zostały zakwalifikowane do wdrożenia.

Dalsze badania mechanizmów dotrenowania. Dotrenowanie wymaga adekwatnego zbioru uczącego. Nie ustalono jaki okres byłby optymalny, zauważono, że

nieprawidłowy horyzont prowadzi do degradacji modelu. Konieczne jest zbadanie przełożenia długości okna dotrenowania w wielomiesięcznym horyzoncie [23], [88].

Wprowadzenie dodatkowych danych do wektora wejściowego. Sprawdzenie po uzyskaniu danych z liczników przepływu oraz pomiarów energii elektrycznej czy te informacje będą tylko powielać pomiar ciśnienia, czy będą też przekładać się na lepsze predykcje. Uzyskanie dostępu do pomiaru przepływu umożliwiłoby estymację rozkładu zapotrzebowania, a co za tym idzie ocenę przydatności sterowania stochastycznego [79], którego zastosowanie w niniejszej pracy wykluczono z powodu braku pomiaru.

Przeniesienie asymetrii kosztów do kroku gradientowego. W zaproponowanej metodzie kara za błąd krytyczny nakładana jest na decyzję po przekształceniu progowym, przez co oddziałuje na wybór modelu, nie zaś na przebieg uczenia. Alternatywą jest miara oparta na pewności predykcji [74], przenosząca asymetrię bezpośrednio do kroku gradientowego. Należałoby zbadać czy przyniosłoby to niższe wartości S_{FN}, czy przez wzmocnienie sygnału pochodzącego z rzadkich zdarzeń, prowadziłoby do degeneracji obserwowanej w studium ablacji dla architektur jednokierunkowych. Rozstrzygnięcie wymaga porównania obu wariantów na tym samym zbiorze, przy kontrolowanym kryterium zatrzymania.

Automatyzacja doboru okresu próbkowania. Sprawdzenie, czy na bazie zdobytego doświadczenia możliwe byłoby wprowadzenie mechanizmu oceniającego odwzorowanie rzeczywistej charakterystyki pracy. W rozprawie w związku z ograniczeniami technologicznymi skupiono się na narzuconej z góry częstotliwości próbkowania, natomiast przykład obiektu D3 wykazał potrzebę zwiększania częstotliwości próbkowania. Być może rozwiązania przetwarzania brzegowego mogłyby tutaj dostarczyć więcej informacji.

Opracowanie inteligentnego mechanizmu nadzorczego i zaawansowanych wskaźników oceny. Ponieważ dla operatora ludzkiego ocena prawidłowości działania układu jest intuicyjna, chciano by przełożyć to rozwiązanie na model, który weryfikowałby również charakterystykę, poziomy do jakich spadło ciśnienie i podejmował być może działania prewencyjne.

Rozbudowa wskaźników nadzoru operacyjnego. Planowane jest wprowadzenie wskaźnika monitorującego negatywne skutki procesowe decyzji. Jego zadaniem byłaby obserwacja czasu trwania spadków ciśnienia poniżej określonego poziomu ciśnienia (znajdującej się poniżej dolnej granicy ciśnienia roboczego) po wysłaniu decyzji sterującej [88]. Należałoby wówczas odróżnić dane związane z chwilowymi zmianami zapotrzebowania od spowodowanych wyłączeniem. Obecny wskaźnik R_{FN} skupiony jest na wykrywaniu zabezpieczeń bez oceny skutków procesowych.

Wyznaczenie współczynnika kWh/1000 Nm^3 dla zoptymalizowanego obiektu. Wybranie obiektu, który dysponowałby pełnym opomiarowaniem w zakresie energii

elektrycznej i sprężonego powietrza, a następnie porównanie osiągniętych oszczędności w ramach publikacji wyników.

Transfer wiedzy między obiektami. Przy początkowych badaniach, w publikacji [18] zwrócono uwagę na jedną z sieci, która miała uniwersalne własności i dawała akceptowalne wyniki (wg ówczesnych metryk) pomiędzy różnymi obiektami. Badania wówczas przeprowadzono tylko w zakresie predykcji ciśnienia bez przewidywania stanów maszyn. Chciano by sprawdzić, czy w zależności od danej grupy obiektów możliwe byłoby zbudowanie uniwersalnego modelu predykcyjnego [57]. Następnym celem byłoby wykorzystanie modeli pomiędzy obiektami o zbliżonej charakterystyce.

Automatyzacja mechanizmu tłumaczącego podjęte decyzje. Zapis dokładnego stanu predykcji w momencie jej podjęcia w postaci wykresów z ukazaniem sytuacji poprzedzającej oraz następczej oraz zestawieniem predykcji z rzeczywistością ułatwiłaby operatorowi procesu diagnostykę oraz umożliwiła proste diagnozowanie nieprawidłowości.

Wyjście poza sieci rekurencyjne. Sprawdzenie na bazie najnowszych publikacji innego podejścia do przedstawionego problemu z wykorzystaniem takich modeli jak transformer [65] dostosowanych do istniejącego problemu.

Sprawdzenie wskaźników oceny poza sektorem sprężonego powietrza. Wskaźniki oceny wyznaczane dla zdarzeń decyzyjnych prawdopodobnie mogłyby być użyteczne w innych rodzajach przemysłowych układów, które brałyby pod uwagę asymetryczny koszt. W ramach działalności badawczej prowadzonej przy projektach badawczo-rozwojowych proponowana metodyka znajdzie szersze zastosowanie.

Bibliografia

[1] Eurostat, *Final energy consumption in industry – detailed statistics*, Statistics Explained. term. wiz. 27 kw. 2026. adr.: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Final_energy_consumption_in_industry_-_detailed_statistics.

[2] *Dyrektywa Parlamentu Europejskiego i Rady (UE) 2023/1791 z dnia 13 września 2023 r. w sprawie efektywności energetycznej oraz zmieniająca rozporządzenie (UE) 2023/955 (wersja przekształcona) (Tekst mający znaczenie dla EOG)*, 2023-09-13, 2023-09. term. wiz. 27 kw. 2026.

[3] „Zalecenie Komisji (UE) 2024/2002 z dnia 24 lipca 2024 r. ustanawiające wytyczne dotyczące interpretacji art. 11 dyrektywy Parlamentu Europejskiego i Rady (UE) 2023/1791 w odniesieniu do systemów zarządzania energią i audytów energetycznych (notyfikowana jako dokument nr C(2024) 5155)," European Commission, spraw. tech., 2024-07-24, 2024-07. term. wiz. 1 maj. 2026.

[4] International Organization for Standardization, *ISO 50001:2018 energy management systems – requirements with guidance for use*, 2018.

[5] International Organization for Standardization, *ISO 50002-1:2025 energy audits – part 1: General requirements with guidance for use*, 2025.

[6] Polski Komitet Normalizacyjny, *PN-EN 16247-1:2012 Audyty energetyczne – Część 1: Wymagania ogólne*, 2012.

[7] K. Schwab, *The Fourth Industrial Revolution*. Geneva: World Economic Forum, 2016, ISBN: 978-1-944835-00-2.

[8] A. Gilchrist, *Industry 4.0: The Industrial Internet of Things*. Berkeley, CA: Apress, 2016, ISBN: 978-1-4842-2046-7. DOI: 10.1007/978-1-4842-2047-4.

[9] D. Gryboś i J. S. Leszczyński, „A Review of Energy Overconsumption Reduction Methods in the Utilization Stage in Compressed Air Systems," *Energies*, t. 17, nr. 6, s. 1495, sty. 2024, ISSN: 1996-1073. DOI: 10.3390/en17061495. term. wiz. 1 maj. 2026.

[10] British Compressed Air Society, „Reducing energy consumption from compressed air usage," British Compressed Air Society, Best Practice Guide BPG 106, 2021.

[11] F. Krokowski, „Energieeffiziente Druckluftinstallationen," Viega GmbH & Co. KG, Attendorn, Whitepaper, 2023.

[12] Umweltbundesamt, „Potenzialstudie Energie-/Kosteneinsparung in der Fluidtechnik," Umweltbundesamt, Dessau-Roßlau, Studie, 2021.

[13] U.S. Department of Energy, *Improving Compressed Air System Performance: A Sourcebook for Industry*, 2016.

[14] J. J. Cabello Eras, A. Sagastume Gutiérrez, V. Sousa Santos i M. J. Cabello Ulloa, „Energy management of compressed air systems. Assessing the production and use of compressed air in industry," *Energy*, t. 213, s. 118 662, grud. 2020, ISSN: 03605442. DOI: 10.1016/j.energy.2020.118662. term. wiz. 27 kw. 2026.

[15] K. Kasprzyk i A. Gałuszka, „Energy Savings Possibilities Gained from Neural Network Application in Compressed Air Supervisory Control Systems," w *2023 27th International Conference on Methods and Models in Automation and Robotics (MMAR)*, Międzyzdroje, Poland: IEEE, sierp. 2023, s. 279–285, ISBN: 979-8-3503-1107-5. DOI: 10.1109/MMAR58394.2023.10242517. term. wiz. 23 lut. 2026.

[16] P. Król, P. Małka, A. Gallina i K. Kasprzyk, „Modeling Wastewater Treatment Plants to Study Energy Optimal Control," *Energies*, t. 18, nr. 22, s. 6011, sty. 2025, ISSN: 1996-1073. DOI: 10.3390/en18226011. term. wiz. 11 maj. 2026.

[17] *Misja, wizja, wartości* -. term. wiz. 28 sierp. 2026. adr.: https://www.marani.pl/firma/misja-wizja-wartosci/.

[18] K. Kasprzyk i A. Gałuszka, „Prediction of compressed air demand depending on the type of production with the use of neural networks," *Advances in Science and Technology Research Journal*, t. 17, nr. 3, s. 154–159, 2023, ISSN: 2080-4075.

[19] A. Gałuszka i K. Kasprzyk, „LSTM Networks in Prediction of the Demand for Compressed Air Depending on the Type of Production," w *Abstracts & Proceedings Book of the International Conference on Science and Technology (ICONST 2022)*, (Budva, Montenegro), Isparta, Turkey: ICONST, lip. 2022, s. 20, ISBN: 978-605-70965-6-2.

[20] International Organization for Standardization, *ISO 8573-1:2010 compressed air – part 1: Contaminants and purity classes*, 2010.

[21] Marani Sp. z o.o. „Sprężarki śrubowe z falownikiem – oferta, ceny i dobór," term. wiz. 4 wrz. 2026. adr.: https://www.marani.pl/produkty/sprezarki-srubowe/.

[22] Tamturbo Ltd., *Sprężarki powietrza Tamturbo® Touch-Free. Broszura produktowa*, https://tamturbo.com/wp-content/uploads/tamturbo_brochure_210x280_PL.pdf, Bezolejowe turbosprężarki z aktywnymi łożyskami magnetycznymi, 2017. term. wiz. 23 wrz. 2026.

[23] V. Bezold, N. Weckerlein, F. Schnell i A. Sauer, „Data-Efficient Forecasting for Compressed Air Systems: Retraining Instead of More Data," w *2025 6th International Conference on Clean and Green Energy Engineering (CGEE)*, Zagreb, Croatia: IEEE, sierp. 2025, s. 95–99, ISBN: 979-8-3315-2398-5. DOI: 10.1109/CGEE65971.2025.11327906. term. wiz. 18 maj. 2026.

[24] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu i Y. Zhang, „Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network," *IEEE Transactions on Smart Grid*, t. 10, nr. 1, s. 841–851, 2019. DOI: 10.1109/TSG.2017.2753802. term. wiz. 27 lip. 2026.

[25] D. Salomone-González, „Advanced Strategies for the Efficient Optimization and Control of Industrial Compressed Air Systems," *Results in Engineering*, t. 26, s. 105 429, czer. 2025, ISSN: 2590-1230. DOI: 10.1016/j.rineng.2025.105429. term. wiz. 22 maj. 2026.

[26] M. H. Albadi i E. F. El-Saadany, „Demand Response in Electricity Markets: An Overview." DOI: 10.1109/PES.2007.385728. term. wiz. 4 wrz. 2026.

[27] International Organization for Standardization, *ISO 11011:2013 compressed air – energy efficiency – assessment*, 2013.

[28] M. Mastrowski, J. Smolka, D. Butrymowicz, Ł. Antczak, J. Gagan i K. Śmierciew, „Experimental Validation of the Theoretical Ejector Model in a Low-Grade Waste Heat Refrigeration System Using R1233zdE as a Working Fluid," *Applied Thermal Engineering*, t. 236, s. 121 716, sty. 2024, ISSN: 1359-4311. DOI: 10.1016/j.applthermaleng.2023.121716. term. wiz. 6 wrz. 2026.

[29] B. Semlitsch i M. Mihăescu, „Flow phenomena leading to surge in a centrifugal compressor," maj 2016, ISSN: 0360-5442. DOI: 10.17863/CAM.48184. term. wiz. 6 wrz. 2026.

[30] C. A. C. Wengerkievicz, C. F. Mazzola, N. Sadowski, N. J. Batistela, E. Silva i C. J. Deschamps, „Efficiency Determination of Compressor Embedded Induction Motors," *Journal of Microwaves, Optoelectronics and Electromagnetic Applications*, t. 20, s. 658–674, 2021, ISSN: 2179-1074. DOI: 10.1590/2179-10742021v20i3254761. term. wiz. 6 wrz. 2026.

[31] Y. Liu, Y. Li, Z. Chen, C. Wang i Z. Xing, „Numerical Research on Refrigerant Injection for Motor Cooling of Screw Compressor for Heat Pump with Large Temperature Lift," *Applied Thermal Engineering*, t. 253, s. 123 837, wrz. 2024, ISSN: 1359-4311. DOI: 10.1016/j.applthermaleng.2024.123837. term. wiz. 6 wrz. 2026.

[32] O. Serradilla, E. Zugasti i U. Zurutuza, *Deep Learning Models for Predictive Maintenance: A Survey, Comparison, Challenges and Prospect*, paź. 2020. DOI: 10.48550/arXiv.2010.03207. arXiv: 2010.03207 [cs.LG]. term. wiz. 11 czer. 2026.

[33] L. A. Bryan i E. A. Bryan, *Programmable Controllers: Theory and Implementation*, 2 wyd. Atlanta, GA: Industrial Text Company, 1997, ISBN: 978-0-944107-32-4.

[34] W. Bolton, *Programmable Logic Controllers*, 6 wyd. Oxford: Newnes, 2015, ISBN: 978-0-12-802929-9.

[35] E. A. Johnson, „Touch Display - A Novel Input/Output Device for Computers," *Electronics Letters*, t. 1, nr. 8, s. 219–220, 1965. DOI: 10.1049/el:19650200.

[36] F. Beck i B. Stumpe, „Two Devices for Operator Interaction in the Central Control of the New CERN Accelerator," CERN, Geneva, CERN Report CERN-73-06, 1973. DOI: 10.5170/CERN-1973-006.

[37] G. S. Hurst, „Electrical Sensor of Plane Coordinates," US3662105A, maj 1972.

[38] G. S. Hurst i J. E. Parks, „Transparent Sensing Device," US3911215A, paź. 1975.

[39] S. A. Boyer, *SCADA: Supervisory Control and Data Acquisition*, 4 wyd. Research Triangle Park, NC: International Society of Automation (ISA), 2009, ISBN: 978-1-936007-09-7.

[40] G. Clarke i D. Reynders, *Practical Modern SCADA Protocols: DNP3, 60870.5 and Related Systems*. Oxford: Newnes, 2004, ISBN: 978-0-7506-5799-0.

[41] T. J. Williams, „The Purdue Enterprise Reference Architecture," *Computers in Industry*, t. 24, nr. 2-3, s. 141–158, 1994. DOI: 10.1016/0166-3615(94)90017-5.

[42] *Humanoid Robot Market: $38.4B by 2035 and Why Early Movers Win*. term. wiz. 22 lip. 2026. adr.: https://www.kaisoresearch.com/blog/humanoid-robot-market-industry-analysis.

[43] J. Rifkin, *The Third Industrial Revolution: How Lateral Power Is Transforming Energy, the Economy, and the World*. New York: Palgrave Macmillan, 2011, ISBN: 978-0-230-34197-5.

[44] M. Hermann, T. Pentek i B. Otto, „Design Principles for Industrie 4.0 Scenarios," w *2016 49th Hawaii International Conference on System Sciences (HICSS)*, sty. 2016, s. 3928–3937. DOI: 10.1109/HICSS.2016.488. term. wiz. 6 wrz. 2026.

[45] R. A. Rojas i E. Rauch, „From a literature review to a conceptual framework of enablers for smart manufacturing control," *The International Journal of Advanced Manufacturing Technology*, t. 104, nr. 1, s. 517–533, wrz. 2019, ISSN: 1433-3015. DOI: 10.1007/s00170-019-03854-4. term. wiz. 6 wrz. 2026.

[46] J. L. Elman, „Finding Structure in Time," *Cognitive Science*, t. 14, nr. 2, s. 179–211, 1990, ISSN: 1551-6709. DOI: 10.1207/s15516709cog1402_1. term. wiz. 6 wrz. 2026.

[47] S. Hochreiter i J. Schmidhuber, „Long Short-Term Memory," *Neural Computation*, t. 9, nr. 8, s. 1735–1780, list. 1997, ISSN: 0899-7667. DOI: 10.1162/neco.1997.9.8.1735. term. wiz. 6 sierp. 2026.

[48] S. Hochreiter, „The Vanishing Gradient Problem During Learning Recurrent Neural Nets and Problem Solutions," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, t. 06, nr. 02, s. 107–116, kw. 1998, ISSN: 0218-4885. DOI: 10.1142/S0218488598000094. term. wiz. 25 lip. 2026.

[49] Y. Bengio, P. Simard i P. Frasconi, „Learning long-term dependencies with gradient descent is difficult," *IEEE Transactions on Neural Networks*, t. 5, nr. 2, s. 157–166, 1994. DOI: 10.1109/72.279181. term. wiz. 25 lip. 2026.

[50] K. Cho i in., „Learning Phrase Representations Using RNN Encoder–Decoder for Statistical Machine Translation," w *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, A. Moschitti, B. Pang i W. Daelemans, red., Doha, Qatar: Association for Computational Linguistics, paź. 2014, s. 1724–1734. DOI: 10.3115/v1/D14-1179. term. wiz. 26 lip. 2026.

[51] J. Chung, C. Gulcehre, K. Cho i Y. Bengio, *Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling*, grud. 2014. DOI: 10.48550/arXiv.1412.3555. arXiv: 1412.3555 [cs.NE]. term. wiz. 26 lip. 2026.

[52] U. Khan, D. S. Cheng, F. Setti, F. Fummi, M. Cristani i L. Capogrosso, „A Comprehensive Survey on Deep Learning-based Predictive Maintenance," *ACM Transactions on Embedded Computing Systems*, t. 25, nr. 2, 25:1–25:43, 2026, ISSN: 1539-9087. DOI: 10.1145/3732287. term. wiz. 21 wrz. 2026.

[53] K. Kasprzyk i A. Gałuszka, „Frequency Driven Screw Compressors Unload State Prediction in Modern Air Control System," w *23rd Industrial Simulation Conference 2025*, Skövde, Szwecja: EUROSIS-ETI, 3–2025, s. 61–65.

[54] M. Schuster i K. Paliwal, „Bidirectional Recurrent Neural Networks," *IEEE Transactions on Signal Processing*, t. 45, nr. 11, s. 2673–2681, list. 1997, ISSN: 1941-0476. DOI: 10.1109/78.650093. term. wiz. 8 czer. 2026.

[55] G. I. Webb, R. Hyde, H. Cao, H. L. Nguyen i F. Petitjean, „Characterizing concept drift," *Data Mining and Knowledge Discovery*, t. 30, nr. 4, s. 964–994, lip. 2016, ISSN: 1573-756X. DOI: 10.1007/s10618-015-0448-4. term. wiz. 27 lip. 2026.

[56] G. I. Webb, L. K. Lee, F. Petitjean i B. Goethals, *Understanding Concept Drift*, kw. 2017. DOI: 10.48550/arXiv.1704.00362. arXiv: 1704.00362 [cs.LG]. term. wiz. 27 lip. 2026.

[57] S. J. Pan i Q. Yang, „A Survey on Transfer Learning," *IEEE Transactions on Knowledge and Data Engineering*, t. 22, nr. 10, s. 1345–1359, 2010. DOI: 10.1109/TKDE.2009.191. term. wiz. 27 lip. 2026.

[58] C. Finn, P. Abbeel i S. Levine, *Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks*, lip. 2017. DOI: 10.48550/arXiv.1703.03400. arXiv: 1703.03400 [cs.LG]. term. wiz. 27 lip. 2026.

[59] J. Yosinski, J. Clune, Y. Bengio i H. Lipson, *How Transferable Are Features in Deep Neural Networks?* List. 2014. DOI: 10.48550/arXiv.1411.1792. arXiv: 1411.1792 [cs.LG]. term. wiz. 27 lip. 2026.

[60] A. Salih i in., „A Perspective on Explainable Artificial Intelligence Methods: SHAP and LIME," *Advanced Intelligent Systems*, t. 7, nr. 1, s. 2 400 304, sty. 2025, ISSN: 2640-4567, 2640-4567. DOI: 10.1002/aisy.202400304. arXiv: 2305.02012 [stat.ML]. term. wiz. 27 lip. 2026.

[61] S. Lundberg i S.-I. Lee, *A Unified Approach to Interpreting Model Predictions*, list. 2017. DOI: 10.48550/arXiv.1705.07874. arXiv: 1705.07874 [cs.AI]. term. wiz. 27 lip. 2026.

[62] M. T. Ribeiro, S. Singh i C. Guestrin, *"Why Should I Trust You?": Explaining the Predictions of Any Classifier*, sierp. 2016. DOI: 10.48550/arXiv.1602.04938. arXiv: 1602.04938 [cs.LG]. term. wiz. 27 lip. 2026.

[63] Y. Wang i in., „An electricity load forecasting model based on multilayer dilated LSTM network and attention mechanism," *Frontiers in Energy Research*, t. 11, czer. 2023, ISSN: 2296-598X. DOI: 10.3389/fenrg.2023.1116465. term. wiz. 11 czer. 2026.

[64] C. Pavlatos, E. Makris, G. Fotis, V. Vita i V. Mladenov, „Enhancing Electrical Load Prediction Using a Bidirectional LSTM Neural Network," *Electronics*, t. 12, nr. 22, s. 4652, sty. 2023, ISSN: 2079-9292. DOI: 10.3390/electronics12224652. term. wiz. 27 lip. 2026.

[65] K. Wang, J. Zhang, X. Li i Y. Zhang, „Long-Term Power Load Forecasting Using LSTM-Informer with Ensemble Learning," *Electronics*, t. 12, nr. 10, s. 2175, sty. 2023, ISSN: 2079-9292. DOI: 10.3390/electronics12102175. term. wiz. 27 lip. 2026.

[66] A. H. Alaoui, A. Meddaoui i M. Hain, *AI-Driven Predictive Maintenance for Industry 4.0: A Systematic Review of Models, Methods, and Challenges*, wrz. 2025. DOI: 10.21203/rs.3.rs-7618987/v1. term. wiz. 21 wrz. 2026.

[67] W. Li i T. Li, „Comparison of deep learning models for predictive maintenance in industrial manufacturing systems using sensor data," *Scientific Reports*, t. 15, nr. 1, s. 23 545, lip. 2025, ISSN: 2045-2322. DOI: 10.1038/s41598-025-08515-z. term. wiz. 11 czer. 2026.

[68] K. Kasprzyk i A. Gałuszka, „Frequency Driven Screw Compressors State Prediction in Modern Air Control System," w *Modelling and Simulation 2024. The 2024 European Simulation and Modelling Conference*, EUROSIS-ETI, 2024, s. 252–258.

[69] K. Kasprzyk i Ł. Antczak, „Predykcja Stanów Sprężarek Śrubowych z Użyciem Sieci LSTM w Celu Wsparcia Nadrzędnych Systemów Sterowania Nowoczesnych Przemysłowych Układów Sprężonego Powietrza," w *Automatyzacja Procesów Dyskretnych. Teoria i Zastosowania*, 2024, s. 133–144, ISBN: 978-83-7880-972-2.

[70] H. Alizadegan, B. Rashidi Malki, A. Radmehr, H. Karimi i M. A. Ilani, „Comparative study of long short-term memory (LSTM), bidirectional LSTM, and traditional machine learning approaches for energy consumption prediction," *Energy Exploration & Exploitation*, t. 43, nr. 1, s. 281–301, sty. 2025, ISSN: 0144-5987. DOI: 10.1177/01445987241269496. term. wiz. 11 czer. 2026.

[71] K. Kasprzyk, A. Gałuszka i Ł. Antczak, „Redukcja Zużycia Energii Elektrycznej w Przemysłowych Układach Sprężonego Powietrza z Wykorzystaniem Uczenia Głębokiego — Doświadczenia z Wdrożeń," w *Automatyzacja Procesów Dyskretnych. Teoria i Zastosowania*, 2026, s. 159–169.

[72] C. Elkan, „The Foundations of Cost-Sensitive Learning," w *Proceedings of the 17th International Joint Conference on Artificial Intelligence (IJCAI)*, Seattle, WA: Morgan Kaufmann, 2001, s. 973–978.

[73] H. He i E. A. Garcia, „Learning from Imbalanced Data," *IEEE Transactions on Knowledge and Data Engineering*, t. 21, nr. 9, s. 1263–1284, 2009. DOI: 10.1109/TKDE.2008.239. term. wiz. 29 lip. 2026.

[74] T.-Y. Lin, P. Goyal, R. Girshick, K. He i P. Dollár, *Focal Loss for Dense Object Detection*, sierp. 2017. DOI: 10.48550/arXiv.1708.02002. arXiv: 1708.02002 [cs.CV]. term. wiz. 8 czer. 2026.

[75] P. Christoffersen i F. X. Diebold, *How Relevant is Volatility Forecasting for Financial Risk Management?* SSRN Scholarly Paper, Rochester, NY, paź. 1998. Social Science Research Network: 1298285. term. wiz. 17 wrz. 2026.

[76] K. J. Arrow, T. Harris i J. Marschak, „Optimal Inventory Policy," *Econometrica*, t. 19, nr. 3, s. 250–272, 1951, ISSN: 0012-9682. DOI: 10.2307/1906813. JSTOR: 1906813. term. wiz. 17 wrz. 2026.

[77] R. Koenker i G. Bassett, „Regression Quantiles," *Econometrica*, t. 46, nr. 1, s. 33–50, 1978, ISSN: 0012-9682. DOI: 10.2307/1913643. JSTOR: 1913643. term. wiz. 18 wrz. 2026.

[78] A. Zellner, „Bayesian Estimation and Prediction Using Asymmetric Loss Functions," *Journal of the American Statistical Association*, t. 81, nr. 394, s. 446–451, czer. 1986, ISSN: 0162-1459. DOI: 10.1080/01621459.1986.10478289. term. wiz. 19 wrz. 2026.

[79] A. Mesbah, „Stochastic Model Predictive Control: An Overview and Perspectives for Future Research," *IEEE Control Systems*, t. 36, nr. 6, s. 30–44, grud. 2016, ISSN: 1066-033X, 1941-000X. DOI: 10.1109/MCS.2016.2602087. term. wiz. 19 wrz. 2026.

[80] J. Davis i M. Goadrich, „The Relationship between Precision-Recall and ROC Curves," w *Proceedings of the 23rd International Conference on Machine Learning*, ser. ICML '06, New York, NY, USA: Association for Computing Machinery, 2006, s. 233–240, ISBN: 978-1-59593-383-6. DOI: 10.1145/1143844.1143874. term. wiz. 20 wrz. 2026.

[81] T. Saito i M. Rehmsmeier, „The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets," *PLOS ONE*, t. 10, nr. 3, e0118432, mar. 2015, ISSN: 1932-6203. DOI: 10.1371/journal.pone.0118432. term. wiz. 20 wrz. 2026.

[82] N. Tatbul, T. J. Lee, S. Zdonik, M. Alam i J. Gottschlich, *Precision and Recall for Time Series*, sty. 2019. DOI: 10.48550/arXiv.1803.03639. arXiv: 1803.03639 [cs.LG]. term. wiz. 20 wrz. 2026.

[83] A. Lavin i S. Ahmad, „Evaluating Real-Time Anomaly Detection Algorithms – The Numenta Anomaly Benchmark," w *2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA)*, grud. 2015, s. 38–44. DOI: 10.1109/ICMLA.2015.141. term. wiz. 20 wrz. 2026.

[84] C.-W. Hsu, C.-C. Chang i C.-J. Lin, „A Practical Guide to Support Vector Classification," National Taiwan University, Department of Computer Science, spraw. tech., 2003.

[85] J. Bergstra i Y. Bengio, „Random Search for Hyper-Parameter Optimization," *Journal of Machine Learning Research*, t. 13, nr. 10, s. 281–305, 2012.

[86] J. Bergstra, R. Bardenet, Y. Bengio i B. Kégl, „Algorithms for Hyper-Parameter Optimization," w *Advances in Neural Information Processing Systems 24*, Curran Associates, 2011, s. 2546–2554.

[87] L. Prechelt, „Early Stopping — But When?" W *Neural Networks: Tricks of the Trade: Second Edition*, G. Montavon, G. B. Orr i K.-R. Müller, red., Berlin, Heidelberg: Springer, 2012, s. 53–67, ISBN: 978-3-642-35289-8. DOI: 10.1007/978-3-642-35289-8_5. term. wiz. 20 wrz. 2026.

[88] J. Lu, A. Liu, F. Dong, F. Gu, J. Gama i G. Zhang, „Learning under Concept Drift: A Review," *IEEE Transactions on Knowledge and Data Engineering*, t. 31, nr. 12, s. 2346–2363, grud. 2019, ISSN: 1558-2191. DOI: 10.1109/TKDE.2018.2876857. term. wiz. 20 wrz. 2026.

[89] J. Gama, I. Žliobaitė, A. Bifet, M. Pechenizkiy i A. Bouchachia, „A Survey on Concept Drift Adaptation," *ACM Computing Surveys (CSUR)*, t. 46, nr. 4, 44:1–44:37, 2014, ISSN: 0360-0300. DOI: 10.1145/2523813. term. wiz. 20 wrz. 2026.

[90] International Organization for Standardization, *ISO 1217:2009 displacement compressors – acceptance tests*, 2009.

[91] M. Karaboyacı, K. Taşdelen, A. Beram, H. Kandemir, E. Kala i S. Özdemir, „International Conferences on Science and Technology. Engineering Science and Technology. ICONST EST 2022, September 7-9, 2022, Budva, Montenegro," w *International Conferences on Science and Technology. Engineering Science and Technology*, Association of Kutbilge Academicians, 2022, ISBN: 978-605-70965-6-2. term. wiz. 6 wrz. 2026.

[92] R. Chen, L. Yang i C. Hafner, „Nonparametric Multistep-Ahead Prediction in Time Series Analysis," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, t. 66, nr. 3, s. 669–686, sierp. 2004, ISSN: 1369-7412. DOI: 10.1111/j.1467-9868.2004.04664.x. term. wiz. 2 sierp. 2026.

[93] P. T. von Hippel, „Mean, Median, and Skew: Correcting a Textbook Rule," *Journal of Statistics Education*, t. 13, nr. 2, sty. 2005, ISSN: null. DOI: 10.1080/10691898.2005.11910556. term. wiz. 9 wrz. 2026.

[94] A. Ghasemi i S. Zahediasl, „Normality Tests for Statistical Analysis: A Guide for Non-Statisticians," *International Journal of Endocrinology and Metabolism*, t. 10, nr. 2, s. 486–489, 2012, ISSN: 1726-913X. DOI: 10.5812/ijem.3505. term. wiz. 8 wrz. 2026.

[95] J. Snoek, H. Larochelle i R. P. Adams, *Practical Bayesian Optimization of Machine Learning Algorithms*, sierp. 2012. DOI: 10.48550/arXiv.1206.2944. arXiv: 1206.2944 [stat.ML]. term. wiz. 20 wrz. 2026.

[96] F. A. Gers, J. Schmidhuber i F. Cummins, „Learning to Forget: Continual Prediction with LSTM," *Neural Computation*, t. 12, nr. 10, s. 2451–2471, paź. 2000, ISSN: 0899-7667. DOI: 10.1162/089976600300015015. term. wiz. 8 czer. 2026.

[97] J. Kim, H. Kim, H. Kim, D. Lee i S. Yoon, „A comprehensive survey of deep learning for time series forecasting: Architectural diversity and open challenges," *Artificial Intelligence Review*, t. 58, nr. 7, s. 216, kw. 2025, ISSN: 1573-7462. DOI: 10.1007/s10462-025-11223-9. term. wiz. 6 wrz. 2026.

A Opis implementacji kluczowych modułów

Kod źródłowy systemu stanowi własność intelektualną firmy Marani Sp. z o.o. i nie jest publicznie dostępny. W niniejszym załączniku przedstawiono opisy algorytmiczne kluczowych komponentów w formie pseudokodu, umożliwiające odtworzenie zaproponowanych rozwiązań.

A.1 Funkcje straty

A.1.1 Wariant bazowy

Wariant bazowy stanowi ważoną sumę dwóch składowych: błędu średniokwadratowego dla przewidywanego ciśnienia oraz binarnej entropii krzyżowej dla stanów sprężarek. Pierwszy element wektora wyjściowego odpowiada ciśnieniu, pozostałe odpowiadają stanom poszczególnych maszyn. Indeksowanie tablic jest zgodne z konwencją języka Python np. zapis $[j{:}k]$ oznacza elementy od j do $k-1$, pominięcie granicy oznacza początek lub koniec osi, -1 to ostatni element, a $\ldots$ wszystkie pozostałe osie.

Algorithm 1 Bazowa funkcja straty (`mse_bce`)

Require: y_{true}, y_{pred} o wymiarze $(B \times h \times (N{+}1))$, wagi w_p, w_b

$p_{true} \leftarrow y_{true}[\ldots, 0], \quad p_{pred} \leftarrow y_{pred}[\ldots, 0]$

$c_{true} \leftarrow y_{true}[\ldots, 1{:}], \quad c_{pred} \leftarrow y_{pred}[\ldots, 1{:}]$

$\mathcal{L}_{MSE} \leftarrow \mathrm{mean}\big((p_{true} - p_{pred})^2\big)$

$\mathcal{L}_{BCE} \leftarrow \mathrm{mean}\big(\mathrm{BCE}(c_{true},\, c_{pred})\big)$

return $w_p \cdot \mathcal{L}_{MSE} + w_b \cdot \mathcal{L}_{BCE}$

A.1.2 Wariant asymetryczny

Wariant rozszerzony uzupełnia powyższą konstrukcję o dwie składowe uwzględniające specyfikę sterowania sprężarkami. Składowa trzecia wykrywa zdarzenia powrotu obciążenia w horyzoncie predykcji i nakłada asymetryczną karę walidacyjną, w której kara za pominięcie powrotu obciążenia (FN) jest wielokrotnie wyższa niż za niepotrzebne utrzymanie maszyny na biegu jałowym (FP). Składowa czwarta wprowadza

dodatkową karę w sytuacji, gdy pominięcie powrotu obciążenia zbiega się ze spadkiem ciśnienia w sieci (taki błąd może nieść największe konsekwencje procesowe). Zapis $\mathbf{1}[A]$ oznacza funkcję wskaźnikową zdefiniowaną jako:

$$\mathbf{1}[A] = \begin{cases} 1, & \text{gdy warunek } A \text{ jest spełniony,} \\ 0, & \text{w przeciwnym razie.} \end{cases}$$

Algorithm 2 Asymetryczna funkcja straty (`industrial`)

Require: y_{true}, y_{pred}, wagi w_p, w_b, w_r, w_t, mnożniki α (FN), β (FP)

Składowa 1 — błąd predykcji ciśnienia

$\mathcal{L}_p \leftarrow \text{mean}\big((p_{true} - p_{pred})^2\big)$

Składowa 2 — podstawowa jakość klasyfikacji

$\mathcal{L}_b \leftarrow \text{mean}\big(\text{BCE}(c_{true},\, c_{pred})\big)$

Składowa 3 — wykrywanie powrotu obciążenia

$s_t \leftarrow c_{true}[:,\, :{-}1,\, :], \quad s_{t+1} \leftarrow c_{true}[:,\, 1:,\, :]$

$R \leftarrow \mathbf{1}[\, s_t < 0{,}5 \;\wedge\; s_{t+1} > 0{,}5\,]$ ▷ zdarzenie $0 \to 1$

$\hat{c}_{t+1} \leftarrow \mathbf{1}[\, y_{pred}[:,\, 1:,\, 1:] > 0{,}5\,]$

$FN \leftarrow R \cdot (1 - \hat{c}_{t+1})$ ▷ powrót wystąpił, model go nie przewidział

$FP \leftarrow (1 - R) \cdot \hat{c}_{t+1}$ ▷ model przewidział powrót, który nie nastąpił

$\mathcal{L}_r \leftarrow \alpha \cdot \text{mean}(FN) + \beta \cdot \text{mean}(FP)$

Składowa 4 — uwzględnienie trendu ciśnienia

$\Delta p \leftarrow p_{true}[:,\, 1:] - p_{true}[:,\, :{-}1]$

$D \leftarrow \mathbf{1}[\, \Delta p < 0\,]$ ▷ ciśnienie opada

$\mathcal{L}_t \leftarrow \text{mean}\big(D \cdot R \cdot (1 - \hat{c}_{t+1})\big)$

Suma ważona

return $w_p\mathcal{L}_p + w_b\mathcal{L}_b + w_r\mathcal{L}_r + w_t\mathcal{L}_t$

Domyślne wartości wag wynoszą $w_p = 0{,}15$, $w_b = 0{,}10$, $w_r = 0{,}78$ oraz $w_t = 0{,}15$, przy mnożnikach $\alpha = 12{,}0$ i $\beta = 1{,}0$. Rozkład ten odzwierciedla priorytet przypisany wykrywaniu powrotu obciążenia; szczegółowy dobór wag dla poszczególnych obiektów opisano w rozdziale 3.

A.2 Architektura sieci

Model przyjmuje sekwencję L próbek i zwraca prognozę o tej samej długości $h = L$ kroków w przód, obejmującą ciśnienie (wyjście o aktywacji liniowej) oraz stany sprężarek (wyjście o aktywacji sigmoidalnej). Warstwy rekurencyjne przetwarzają sekwencję dwukierunkowo, co ma decydujący wpływ na jakość predykcji (wytłumaczono w rozdziale 4).

Algorithm 3 Budowa modelu predykcyjnego

Require: lista jednostek $[u_1, \ldots, u_k]$, współczynnik d (dropout), liczba sprężarek N, długość sekwencji $L = h$

$x \leftarrow \mathrm{Input}(L \times (N{+}1))$
for $i = 1$ **to** k **do**
    $x \leftarrow \mathrm{Bidirectional}\big(\mathrm{LSTM}(u_i/2,\ \mathit{return_sequences} = \mathrm{prawda})\big)(x)$
    $x \leftarrow \mathrm{Dropout}(d)(x)$
end for
$y \leftarrow \mathrm{TimeDistributed}\big(\mathrm{Dense}(N{+}1)\big)(x)$ ▷ wyjście dla każdego kroku
$y[\ldots, 0] \leftarrow y[\ldots, 0]$ ▷ ciśnienie, aktywacja liniowa
$y[\ldots, 1{:}] \leftarrow \sigma\big(y[\ldots, 1{:}]\big)$ ▷ stany sprężarek, aktywacja sigmoidalna
return model o wejściu x i wyjściu y o wymiarze $(h \times (N{+}1))$

A.3 Wyznaczanie metryk bezpieczeństwa

Ocena modelu oparta jest o pojedyncze zdarzenia decyzyjne, nie zaś o pojedyncze próbki. Dla każdej wykrytej tranzycji odciążenia wyznaczane jest okno decyzyjne, w obrębie którego stany rzeczywiste i predykowane agregowane są do postaci binarnej. Klasyfikacja zdarzenia wynika z pary tych wskaźników, zgodnie z opisem przedstawionym w rozdziale 3.

Algorithm 4 Klasyfikacja zdarzeń w oknach decyzyjnych

Require: szereg stanów rzeczywistych c, szereg predykcji $\hat{c}$, granice okna w_{start}, w_{end}

$TP \leftarrow 0, \quad TN \leftarrow 0, \quad FP \leftarrow 0, \quad FN \leftarrow 0$
$T \leftarrow$ zbiór chwil, w których wystąpiła tranzycja $1 \rightarrow 0$
for all $t \in T$ **do**
    $W \leftarrow [\, t + w_{start},\ t + w_{end} \,]$
    $g \leftarrow \mathbf{1}\big[\exists\, \tau \in W :\ c(\tau) = 1\big]$ ▷ rzeczywisty powrót obciążenia
    $\hat{g} \leftarrow \mathbf{1}\big[\exists\, \tau \in W :\ \hat{c}(\tau) > 0{,}5\big]$ ▷ powrót przewidziany
    if $g = 1$ **and** $\hat{g} = 1$ **then**
        $TP \leftarrow TP + 1$
    else if $g = 0$ **and** $\hat{g} = 0$ **then**
        $TN \leftarrow TN + 1$
    else if $g = 0$ **and** $\hat{g} = 1$ **then**
        $FP \leftarrow FP + 1$
    else
        $FN \leftarrow FN + 1$ ▷ błąd krytyczny
    end if
end for
$S \leftarrow (TP + TN)/(TP + TN + FP + FN)$
$S_{FN} \leftarrow FN/(TP + FN)$
$S_{FP} \leftarrow FP/(TP + TN + FP + FN)$
return S, S_{FN}, S_{FP}

A.4 Weryfikacja następcza decyzji w eksploatacji

Mechanizm nadzoru operacyjnego rejestruje każdą wydaną rekomendację wyłączenia i po upływie okna weryfikacji porównuje ją z rzeczywistym zachowaniem układu. Wynikiem jest wskaźnik R_{FN}, opisany w rozdziale 5. Poniższy pseudokod przedstawia logikę pojedynczego cyklu sterowania.

Algorithm 5 Cykl sterowania z weryfikacją następczą

Require: model f_θ, kolejka oczekujących decyzji Q, liczniki N_D, FN_{op}

loop ▷ cykl o okresie T_{check}
  $\mathbf{x} \leftarrow$ odczyt bieżącego stanu układu (ciśnienie, stany maszyn)
  $\mathbf{X} \leftarrow$ ostatnie L odczytów z bufora, łącznie z $\mathbf{x}$
  — *weryfikacja decyzji wydanych wcześniej* —
  for all $(i, t_d) \in Q$ takich, że upłynęło okno weryfikacji **do**
    if sprężarka i powróciła pod obciążenie w oknie po t_d **then**
      $FN_{op} \leftarrow FN_{op} + 1$; zapisz zdarzenie w dzienniku
    else
      $N_D \leftarrow N_D + 1$
    end if
    usuń (i, t_d) z Q
  end for
  — *generowanie nowej rekomendacji* —
  $\hat{\mathbf{Y}} \leftarrow f_\theta(\mathbf{X})$ ▷ predykcja na horyzoncie h
  for all sprężarek i w stanie biegu jałowego **do**
    $\hat{g}_i \leftarrow \mathbf{1}\big[\exists\, \tau \in W :\ \hat{\mathbf{Y}}_i(\tau) > 0{,}5\big]$
    if $\hat{g}_i = 0$ **then** ▷ model nie przewiduje powrotu obciążenia
      if spełnione warunki deterministyczne **then** ▷ limit załączeń, progi ciśnienia, stan układu
        if upłynął czas wstrzymania decyzji **then**
          wyślij rekomendację wyłączenia sprężarki i do sterownika
          dodaj (i, t) do kolejki Q
        end if
      else
        zarejestruj blokadę rekomendacji
      end if
    end if
  end for
  $R_{FN} \leftarrow FN_{op}/(FN_{op} + N_D)$
end loop

Rekomendacja przekazywana jest do sterownika dopiero po spełnieniu warunków deterministycznych. Ich niespełnienie skutkuje zarejestrowaniem blokady, przy czym, jak wykazano w rozdziale 5, liczba blokad bywa znacząca. Wstrzymanie realizacji decyzji o ustalony czas stanowi dodatkowe zabezpieczenie, niezależne od okresu próbkowania, chroniące przed podejmowaniem decyzji na podstawie stanu nieustalonego.

B Konfiguracja systemu

Załącznik przedstawia strukturę plików konfiguracyjnych wykorzystywanych w zaproponowanym rozwiązaniu. Ze względu na poufność danych obiektów przemysłowych przedstawiono wyłącznie szablony pozbawione wartości identyfikujących instalację. Wartości liczbowe odpowiadają rzeczywistym konfiguracjom, natomiast adresy sieciowe, nazwy baz danych oraz identyfikatory projektów zostały zastąpione oznaczeniami zastępczymi.

W systemie funkcjonują dwa odrębne pliki konfiguracyjne pełniące różne role. Pierwszy z nich opisuje parametry procesu uczenia oraz architekturę modelu i towarzyszy zapisanemu modelowi, umożliwiając jego odtworzenie. Drugi definiuje sposób integracji warstwy predykcyjnej z konkretną instalacją tzn. zawiera mapowanie sygnałów, parametry komunikacji ze sterownikiem oraz warunki bezpieczeństwa. Rozdzielenie to pozwala na wdrożenie tego samego modelu w różnych konfiguracjach sprzętowych oraz na modyfikację parametrów pracy systemu bez ingerencji w model.

B.1 Konfiguracja procesu uczenia

Plik konfiguracji uczenia definiuje obiekt, parametry przygotowania danych, funkcję straty, architekturę sieci oraz parametry procesu optymalizacji. Poniżej przedstawiono strukturę odpowiadającą modelowi wykorzystującemu asymetryczną funkcję straty wraz z hiperparametrami dobranymi metodą optymalizacji bayesowskiej.

```yaml
object:
  name: OBIEKT_2min          # identyfikator obiektu i horyzontu
  group: X                   # grupa konfiguracyjna (A, B, C)
  n_compressors: 2           # liczba sprężarek w układzie
  csv_path: ...              # ścieżka do zbioru danych

data:
  lookback: 24               # dłgość okna obserwacji [próbki]
  forecast_horizon: 24       # horyzont predykcji [próbki]
  train_split: 0.75          # podział zbiór uczący i walidacyjny
  Pd: 0.0                    # limit filtracji ciśnienia Pd [bar]
  Pg: 10.0                   # limit filtracji ciśnienia Pg [bar]
  train_mode: sliding        # metoda tworzenia sekwencji

loss:
```

```yaml
  loss_type: industrial       # wybrana funkcja straty
  wsp_linear: 0.6             # waga składowej ciśnienia
  wsp_sigmoid: 0.4            # waga składowej klasyfikacyjnej
  wsp_reload: 0.72            # waga wykrywania powrotu obciążenia
  fn_multiplier: 20.0         # mnożnik kary za FN

model:
  architecture: bilstm        # wybrana architektura
  params:
    units: [96, 32]           # liczba jednostek warstw rekur.
    dropout: 0.260848

training:                     # parametry uczenia
  epochs: 150
  batch_size: 8192
  learning_rate: 0.010248
  optimizer: adam
  seed: 42
```

Listing B.1: Konfiguracja procesu uczenia — struktura pliku

Gwoli przypomnienia: przyjęto założenie gdzie `lookback = forecast_horizon` czyli model obserwuje tyle samo próbek wstecz, ile przewiduje w przód. Przy okresie próbkowania wynoszącym 5 s wartość 24 próbek odpowiada oknu dwuminutowemu. Wartości `units`, `dropout` oraz `learning_rate` odbiegają od konfiguracji domyślnej, ponieważ pochodzą z optymalizacji bayesowskiej przeprowadzonej dla tego obiektu.

B.2 Konfiguracja wdrożeniowa

Plik konfiguracji wdrożeniowej opisuje integrację z konkretną instalacją. Poniżej przedstawiono jego strukturę w podziale na sekcje funkcjonalne, z pominięciem fragmentów zawierających dane identyfikujące obiekt.

B.2.1 Komunikacja ze sterownikiem

Komunikacja z warstwą sterowania realizowana jest przez dedykowany blok danych sterownika. Parametr `tryb_zimowy` obsługuje sytuację, w której sterownik nadrzędny wykrywa przypadki wychłodzenia maszyn i uruchamia je w celu dogrzania, aby zapewnić im gotowość rozruchową (w takim wypadku ich wyłączenie jest blokowane niezależnie od predykcji).

```yaml
plc:
  ip: 0.0.0.0                 # adres sterownika
  rack: 0
  slot: 1
```

```
  db_number: 2000              # blok danych wymiany PLC moduł SI
  timeout_ms: 5000

  tryb_zimowy:
    enabled: true              # opcja dogrzewania sprężarek
```

Listing B.2: Parametry komunikacji

B.2.2 Mapowanie sygnałów

Sekcja mapowania odzwierciedla konkretną instalację przemysłową oraz jej reprezentację w bazie danych. Wybrane kolumny tabel mogą różnić się pomiędzy obiektami w zależności od konfiguracji sterownika oraz liczby wprowadzonych maszyn. Sygnały `awaria` oraz `brak_kom` wykorzystywane są do filtracji danych opisanej w rozdziale 3. Próbki zarejestrowane w stanie awarii lub braku komunikacji nie są uwzględniane ani w procesie uczenia, ani przy podejmowaniu decyzji.

```
baza:
  enabled: true           # system aktywny
  database: BAZA_OBIEKTU
  tabela_decyzje: obiekt_decyzje
  tabela_alarmy: obiekt_alarmy

  mapowanie:
    timestamp: ts          # data i czas
    cisnienie: v1          # pomiar ciśnienia
    wylaczenie: v2         # system sterowania nieaktywny

    tagi_sprezarki:
      - aktywne: v10       # sprężarka pracuje w trybie zdalnym
        awaria: v20        # wystąpienie awarii
        brak_kom: v30      # brak komunikacji z maszyną
      - aktywne: v11
        awaria: v21
        brak_kom: v31
```

Listing B.3: Mapowanie sygnalow procesowych

B.2.3 Parametry sprężarek

Poniżej przedstawiono parametry związane z poszczególnymi sprężarkami.

```
sprezarki:
  - idx: 0                     # indeks urządzenia
    nazwa: SP1                 # identyfikator literalny
    predykcja_enabled: true    # predykcja aktywna
```

```
    czas_rozprezania_s: 120  # czas rozprezania po zatrzymaniu

  - idx: 1
    nazwa: SP2
    predykcja_enabled: true
    czas_rozprezania_s: 30
```

LISTING B.4: Definicja maszyn w ukladzie

Parametr `czas_rozprezania_s` określany jest indywidualnie dla każdej maszyny i wynika z jej konstrukcji. Sprężarka o dłuższym czasie rozprężania wymaga większego marginesu bezpieczeństwa, ponieważ jej ponowne załączenie nie jest możliwe w krótkim czasie po wyłączeniu. Wartości te mogą różnić się między maszynami w obrębie jednego układu.

B.2.4 Parametry sterowania i bezpieczeństwa

Sekcja ta zawiera parametry omówione w rozdziale 5. Parametr `max_zalaczen_na_godzine` realizuje ograniczenie liczby wyłączeń maszyny, chroniąc ją przed zbyt dużą liczbą rozruchów. Parametr `cooldown_obserwacji_s` odpowiada za wstrzymanie realizacji decyzji do czasu ustalenia się stanu układu. Rozdzielenie `sample_interval_s` oraz `check_interval_s` pozwala na częstsze sprawdzanie stanu układu.

```
sterowanie:
  sample_interval_s: 20      # okres probkowania danych
  check_interval_s: 5        # okres cyklu sprawdzania S7
  max_zalaczen_na_godzine: 10
  cooldown_obserwacji_s: 20  # wstrzymanie realizacji decyzji
  safe_window_probek: 6      # domyslne okno decyzyjne
  safe_tryb: any             #
  safe_threshold: 0.5
  weryfikacja_po_probkach: 6 # horyzont weryfikacji następczej
```

LISTING B.5: Warunki bezpieczenstwa i parametry decyzji

B.2.5 Dotrenowanie i monitorowanie

Mechanizm dotrenowania może zostać wyłączony niezależnie dla każdego obiektu (jak wykazano w rozdziale 5), dla części instalacji zdecydowano się na jego pominięcie. Sekcja `monitoring` definiuje progi, po których przekroczeniu system generuje ostrzeżenie o pogorszeniu jakości predykcji.

```
dotrenowanie:
  enabled: true
  interval_h: 24             # okres uruchamiania
  okno_danych_h: 24          # zakres danych do dotrenowania
```

```
  epochs: 20
  batch_size: 2048

monitoring:
  min_unload_match: 0.70      # próg ostrzegawczy jakosci predykcji
  max_safe_fn: 0.30           # próg ostrzegawczy ryzyka
```

Listing B.6: Parametry dotrenowania i nadzoru