

Politechnika Śląska
Wydział Mechaniczny Technologiczny
Katedra Podstaw Konstrukcji Maszyn

Rozprawa doktorska

Modelowanie i optymalizacja procesu
cięcia elektroerozyjnego z zastosowaniem metod
sztucznej inteligencji

mgr inż. Łukasz Łomozik

Promotor:

Prof. dr hab. inż. Andrzej Sokołowski

Promotor pomocniczy:

dr inż. Grzegorz Dyrbuś

Opiekun pomocniczy:

dr hab. inż. Henryk Bąkowski

Gliwice, 2024

SPIS TREŚCI

Wykaz ważniejszych oznaczeń	6
1. Wprowadzenie	8
2. Cel i zakres badań.....	10
2.1. Cel pracy.....	10
2.2. Zakres pracy	11
2.3. Pojęcia podstawowe	12
2.4. Narzędzia programistyczne i środowisko obliczeniowe użyte w badaniach.....	14
3. Podstawy obróbki elektroerozyjnej	15
3.1. Historia i rozwój obróbki elektroerozyjnej.....	15
3.2. Cięcie elektroerozyjne drutem.....	18
3.2.1. Mechanizm usuwania materiału	18
3.2.2. Charakterystyka podstawowych parametrów procesu	19
3.2.3. Charakterystyka impulsu elektrycznego w procesie	21
3.2.4. Budowa obrabiarki WEDM.....	23
4. Przegląd badań doświadczalnych cięcia elektroerozyjnego	25
4.1. Badania wpływu parametrów procesu WEDM.....	25
4.1.1. Dokładność geometryczna.....	25
4.1.2. Chropowatość powierzchni obrabianej	28
4.1.3. Prędkość usuwania materiału	31
4.1.4. Stabilność obróbki	32
4.2. Wykorzystanie metod sztucznej inteligencji w procesie WEDM	33
4.2.1. Metody sztucznej inteligencji w modelowaniu procesu.....	34
4.2.2. Metody sztucznej inteligencji w optymalizacji procesu.....	38
4.3. Podsumowanie.....	41
5. Przegląd zastosowanych metod i algorytmów	43
5.1. Metody statystyczne	43
5.1.1. Analiza korelacji.....	43
5.1.2. Regresja liniowa	44

5.1.3. Regresja wielomianowa.....	45
5.1.4. Regresja logistyczna.....	45
5.2. Metody uczenia maszynowego.....	46
5.2.1. Drzewa decyzyjne	46
5.2.2. Las losowy.....	49
5.2.3. Wzmacnianie gradientowe	51
5.2.4. Maszyny wektorów nośnych.....	52
5.2.5. Sztuczne sieci neuronowe.....	55
5.3. Metody optymalizacji.....	57
5.3.1. Algorytm genetyczny	58
5.3.2. Algorytm ewolucji różnicowej.....	59
5.3.3. Algorytm optymalizacji rojem cząstek.....	60
5.3.4. Algorytm mrówkowy	61
6. Zasady i techniki symulacji komputerowych.....	63
6.1. Przygotowanie danych do modelowania.....	63
6.1.1. Postępowanie z niekompletnymi danymi.....	64
6.1.2. Transformacja danych	66
6.1.3. Balansowanie klas	68
6.2. Trenowanie i ewaluacja modeli.....	68
6.2.1. Podział danych na zbiór treningowy i testowy.....	69
6.2.2. Walidacja krzyżowa i optymalizacja hiperparametrów	69
6.3. Metody oceny modeli.....	73
6.3.1. Ocena modeli regresji.....	73
6.3.2. Ocena modeli klasyfikacyjnych	76
6.4. Interpretacja modeli.....	80
7. Opis badań eksperymentalnych.....	83
7.1. Materiały i aparatura badawcza.....	83
7.1.1. Obrabiarka elektroerozyjna	83
7.1.2. Materiał obrabiany.....	84

7.1.3. Wycinane próbki	85
7.1.4. Stanowisko do badań chropowatości powierzchni.....	85
7.1.5. Mikroskopy współrzędnościowe	86
7.2. Metody badań eksperymentalnych	87
7.2.1. Plan badań	87
7.2.2. Wybór zmiennych procesu	88
7.2.3. Plan eksperymentu.....	92
8. Analiza statystyczna wyników pomiarów	95
8.1. Ogólna charakterystyka zbioru danych	95
8.2. Analiza pomiarów pod względem obserwacji odstających	98
8.3. Wyniki analizy korelacji.....	99
9. Badania modelowe procesu WEDM	101
9.1. Model Ra	103
9.1.1. Dobór metody modelowania	103
9.1.2. Przygotowanie danych treningowych.....	104
9.1.3. Analiza jakości modelu	105
9.2. Model MRR.....	108
9.2.1. Dobór metody modelowania	108
9.2.2. Przygotowanie danych treningowych.....	110
9.2.3. Analiza jakości modelu	111
9.3. Model C	114
9.3.1. Dobór metody modelowania	114
9.3.2. Przygotowanie danych treningowych.....	115
9.3.3. Analiza jakości modelu	116
9.4. Model Kw	119
9.4.1. Dobór metody modelowania	119
9.4.2. Dobór metody przygotowania danych.....	120
9.4.3. Analiza jakości modelu	120
9.5. Analiza wpływu parametrów technologicznych na proces WEDM.....	122

9.5.1. Wpływ parametrów technologicznych na chropowatość powierzchni .	122
9.5.2. Wpływ parametrów technologicznych na prędkość usuwania materiału	125
9.5.3. Wpływ parametrów technologicznych na ciągłość procesu.....	128
9.5.4. Wpływ parametrów technologicznych na szerokość szczeliny erozyjnej	132
9.6. Podsumowanie badań modelowych	135
10. Optymalizacja procesu WEDM.....	136
10.1. Kryteria optymalizacji	136
10.2. Funkcja celu.....	137
10.3. Przestrzeń poszukiwań	139
10.4. Analiza porównawcza algorytmów optymalizacji	140
10.4.1. Kryteria porównania	140
10.4.2. Konfiguracja algorytmów	140
10.4.3. Porównanie wyników optymalizacji	142
10.5. Analiza wybranych wyników optymalizacji	144
10.6. Weryfikacja eksperymentalna wyników optymalizacji.....	148
10.7. Podsumowanie optymalizacji procesu WEDM.....	150
11. Opracowanie i wdrożenie oprogramowania optymalizującego	151
11.1. Architektura oprogramowania.....	151
11.2. Wykorzystana technologia	153
11.3. Interfejs użytkownika	154
11.4. Podsumowanie wdrożenia	156
12. Podsumowanie i wnioski	158
12.1. Wnioski końcowe	158
12.2. Plan przyszłych badań	159
Bibliografia.....	161

WYKAZ WAŻNIEJSZYCH OZNACZEŃ

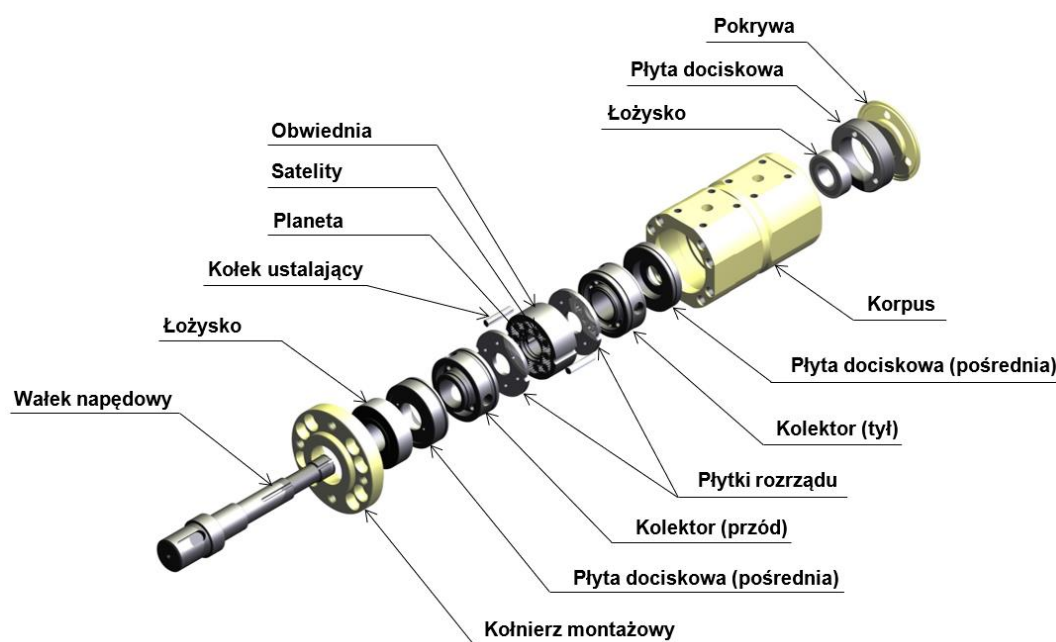
β	Wektor oszacowań współczynników regresji
Tl	Ilość liści w drzewie decyzyjnym
TPR	Wskaźnik wyników prawdziwie pozytywnych (True Positive Rate)
SPE	Specyficzność klasyfikatora
SEN	Czułość klasyfikatora
PRE	Precyzja klasyfikatora
$MAPE$	Średni procentowy błąd bezwzględny
JB	Wynik testu Jarque-Bera
FPR	Wskaźnik wyników prawdziwie negatywnych (True Negative Rate)
ACC	Dokładność klasyfikacji
Λ	Przestrzeń hiperparametrów
Δr	Dokładność kształtu w kierunku profilu
Δd	Dokładność kształtu w kierunku normalnym do posuwu
WT	Naciąg drutu
WF	Prędkość przewijania drutu
SR	Chropowatość powierzchni
S	Skośność rozkładu
Ra	Średnie arytmetyczne odchylenie od linii średniej
R^2	Współczynnik determinacji
MSE	Średni błąd kwadratowy
MRR	Prędkość usuwania materiału
MCC	Współczynnik korelacji Matthews'a
MAE	Średni błąd bezwzględny
M	Liczba drzew decyzyjnych
\mathcal{L}	Funkcja oceniająca jakość modelu
Kw	Szerokość szczeliny erozyjnej
K	Kurtoza rozkładu
FR	Natężenie przepływu dielektryka
DW	Wynik testu Durбина-Watsona
C	Ciągłość obróbki
ω_{ij}	Waga połączenia między j -tym neuronem wejściowym a i -tym neuronem ukrytym

ω'_{ij}	Waga połączenia między i -tym neuronem ukrytym a k -tym neuronem wyjściowym
λ^*	Optymalna konfiguracja hiperparametrów
y_i	Rzeczywista wartość zmiennej zależnej dla i -tej obserwacji
x_i	Wartość zmiennej niezależnej dla i -tej obserwacji
u_e	Napięcie wyładowania
t_{on}	Czas trwania impulsu wyładowania
t_{off}	Czas wyłączenia impulsu
r_{xy}	Współczynnik korelacji liniowej Pearsona
$f_{RF}(x)$	Przewidywane wartości przez las losowy dla obserwacji x
c_m	Stała drzewa decyzyjnego przewidująca wartość zmiennej zależnej w regionie
b'_k	Bias k -tego neuronu wyjściowego
b_i	Bias i -tego neuronu ukrytego
Z_i	Wyjście z i -tego neuronu w warstwie ukrytej
$L^{(t)}$	Funkcja straty w algorytmie XGBoost
I_G	Współczynnik Giniego,
$C_{RF}(x)$	Przewidywana klasa przez las losowy dla obserwacji x
$\psi_{CV}^{(val)}$	Uśredniona jakość modelu w walidacji krzyżowej
M_λ	Model z określoną konfiguracją hiperparametrów
\hat{y}_i	Przewidywana wartość zmiennej zależnej dla i -tej obserwacji
\hat{u}_i	Napięcie otwartej przerwy
\hat{I}_e	Prąd szczytowy
$\hat{C}_m(x)$	Przewidywana klasa dla m -tego drzewa
$(X, Y)^{(train)}$	Zbiór danych treningowych
$(X, Y)^{(test)}$	Zbiór danych testowych

1. WPROWADZENIE

Współczesny przemysł nieustannie dąży do zwiększenia wydajności oraz dokładności procesów technologicznych, co wynika z rosnących wymagań dotyczących jakości wytwarzanych komponentów oraz optymalizacji kosztów produkcji. W dobie zaawansowanych technologii obróbki, jakość produktów i efektywność produkcji stają się kluczowymi czynnikami determinującymi konkurencyjność przedsiębiorstw na rynku [1]. Szczególne znaczenie mają procesy produkcyjne o wysokim stopniu złożoności, które wymagają precyzyjnego nadzoru i ciągłej optymalizacji, aby sprostać rygorystycznym normom jakościowym i ekonomicznym.

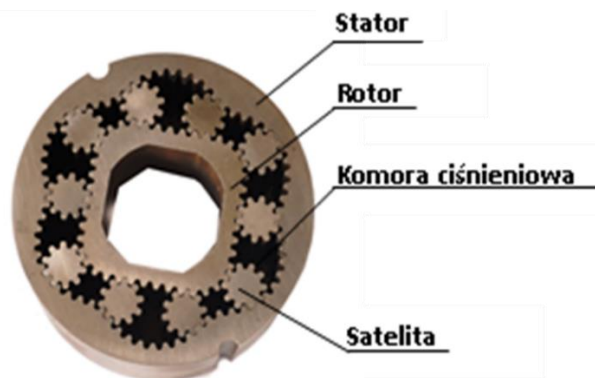
Przykładem przedsiębiorstwa, które zmagają się z tymi wyzwaniami, jest firma SMHydro, Sp. z o.o. z siedzibą w Katowicach (<https://smhydro.com.pl/>). Specjalizuje się ona w produkcji satelitowych silników hydraulicznych (Rys. 1.1), które znajdują zastosowanie w wymagających warunkach przemysłowych, ze szczególnym uwzględnieniem sektora górniczego, gdzie niezawodność i odporność na trudne warunki eksploatacyjne są priorytetowe. W takich środowiskach, jak kopalnie, trwałość i bezawaryjność silników są istotnymi kryteriami oceny ich wartości użytkowej.



Rys. 1.1 Satelitowy silnik hydrauliczny produkowany przez firmę SMHydro

Aby spełnić wymagania stawiane produkowanym silnikom, SMHydro wprowadziło mechanizm satelitowy (Rys. 1.2). Rozwiązanie to polega na umieszczeniu uzębionych satelitów pomiędzy statorem a rotorem silnika. Satelity te pełnią funkcję rozdzielacza

i uszczelnienia komór ciśnieniowych, a ich ewolwentowe zazębienie umożliwia toczenie się powierzchni uszczelniających bez tarcia poślizgowego, co znacząco poprawia trwałość silnika. Ponadto, dzięki temu rozwiązaniu możliwe jest stosowanie cieczy roboczych o niskiej smarowności i lepkości, takich jak niskoprocetowa emulsja wodno-olejowa, co zmniejsza wymagania dotyczące filtracji cieczy.



Rys. 1.2 Mechanizm satelitowy

Produkcja elementów mechanizmu satelitowego wiąże się jednak z koniecznością zastosowania wyjątkowo precyzyjnej obróbki. Dlatego firma zdecydowała się na wdrożenie technologii obróbki elektroerozyjnej drutem (ang. Wire Electrical Discharge Machining, WEDM), która umożliwia uzyskanie skomplikowanych kształtów z wysoką dokładnością. Proces WEDM, mimo swoich zalet, generuje jednak wysokie koszty, szczególnie gdy celem jest uzyskanie elementów o wysokiej jakości powierzchni. Koszty te wynikają głównie z konieczności wielokrotnego wykonywania serii cięć elektrodą drutową po tej samej ścieżce obróbki. Początkowe przejścia mają na celu zgrubne usunięcie materiału, natomiast kolejne są realizowane w celu stopniowego poprawiania jakości powierzchni i precyzji wymiarowej obrabianego elementu. W efekcie proces obróbki staje się dłuższy, co zwiększa koszty i zużycie części eksploatacyjnych obrabiarki.

W kontekście produkcji seryjnej, takie wyzwania technologiczne i ekonomiczne mogą istotnie wpływać na rentowność całego procesu. Jedną z możliwości poprawy efektywności jest optymalizacja procesu WEDM, aby zredukować koszty produkcji, jednocześnie zachowując wysoką jakość i precyzję wytwarzanych komponentów. W tym celu współczesna nauka oferuje narzędzia, takie jak metody sztucznej inteligencji, które mają potencjał do wspierania procesów optymalizacji w produkcji.

2. CEL I ZAKRES BADAŃ

2.1. Cel pracy

Celem ogólnym niniejszej pracy doktorskiej jest opracowanie i wdrożenie rozwiązań, które odpowiadają na potrzeby przedsiębiorstwa SMHydro w zakresie optymalizacji procesu cięcia elektroerozyjnego (WEDM). W kontekście rosnących wymagań jakościowych i ekonomicznych, istotne jest, aby proces ten był precyzyjny oraz opłacalny ekonomicznie. Aby osiągnąć te cele, zaproponowano zastosowanie metod sztucznej inteligencji do modelowania i optymalizacji procesu WEDM, co pozwoli na lepsze zarządzanie jego parametrami i efektywniejsze wykorzystanie zasobów.

Realizacja celu ogólnego pracy potencjalnie nastąpi poprzez osiągnięcie następujących celów szczegółowych:

- 1) Analiza wpływu parametrów technologicznych na wybrane wielkości charakteryzujące proces WEDM, takie jak prędkość usuwania materiału, szerokość szczeliny erozyjnej oraz chropowatość powierzchni obrabianej. Analiza ta dostarczy niezbędnych danych do dalszej optymalizacji procesu.
- 2) Opracowanie modeli procesu WEDM. Modele te będą zastosowane do przewidywania wyników obróbki na podstawie zdefiniowanych parametrów technologicznych, co umożliwi lepszą kontrolę nad procesem oraz jego optymalizację.
- 3) Opracowanie oprogramowania do dynamicznej optymalizacji procesu cięcia elektroerozyjnego, które umożliwi bieżące dostosowywanie parametrów technologicznych do specyficznych wymagań jakościowych obrabianych powierzchni oraz wymagań wydajnościowych, wykorzystując metody sztucznej inteligencji.
- 4) Wdrożenie opracowanego oprogramowania w przedsiębiorstwie, mające na celu praktyczne zastosowanie narzędzia do optymalizacji procesu cięcia elektroerozyjnego w warunkach produkcyjnych.

2.2. Zakres pracy

Treść pracy została podzielona na dwanaście rozdziałów. W rozdziale trzecim omówiono podstawy obróbki elektroerozyjnej, w tym historię i rozwój tej technologii, oraz charakterystykę procesu cięcia elektroerozyjnego drutem. Przedstawiono mechanizm usuwania materiału, charakterystykę parametrów procesu oraz budowę obrabiarki WEDM, co stanowi teoretyczne tło dla dalszych badań.

Rozdział czwarty zawiera przegląd badań doświadczalnych związanych z procesem WEDM. Skupiono się w nim na badaniach wpływu parametrów technologicznych na dokładność geometryczną, chropowatość powierzchni oraz prędkość usuwania materiału. Omówiono także stabilność obróbki oraz zastosowanie metod sztucznej inteligencji w modelowaniu i optymalizacji procesu. W rozdziale tym przedstawiono wnioski z przeglądu literatury, zidentyfikowane luki badawcze oraz potencjalne kierunki dalszego rozwoju badań nad technologią WEDM.

Rozdział piąty poświęcono przeglądowi metod i algorytmów, które stanowią podstawę modelowania i optymalizacji procesu WEDM w niniejszej pracy. Omówiono w nim zarówno podejścia statystyczne, jak i metody uczenia maszynowego, a także techniki optymalizacyjne inspirowane biologicznie, które mogą wspierać procesy związane z obróbką elektroerozyjną.

W rozdziale szóstym skoncentrowano się na zasadach i technikach symulacji komputerowych, niezbędnych dla efektywnego modelowania procesu WEDM. Zawiera on opis przygotowania danych do modelowania, trenowania modeli oraz metod oceny ich jakości.

Rozdział siódmy poświęcono opisowi badań eksperymentalnych, które stanowią integralną część procesu weryfikacji równolegle opracowywanych modeli procesu. W rozdziale tym scharakteryzowano materiał obrabiany oraz aparaturę badawczą, w tym obrabiarkę elektroerozyjną i mikroskopy współrzędnościowe. Omówiono także metody badań, plan eksperymentu oraz dobór zmiennych procesu.

Rozdział ósmy zawiera analizę statystyczną wyników pomiarów. Przeanalizowano w nim ogólną charakterystykę zbioru danych oraz wyniki analizy korelacji.

W rozdziale dziewiątym przedstawiono wyniki badań modelowych procesu WEDM, w tym modele wpływu parametrów technologicznych na chropowatość powierzchni, prędkość usuwania materiału, ciągłość obróbki oraz szerokość szczeliny erozyjnej. Rozdział ten zawiera także analizę wpływu parametrów technologicznych na ustalone zmienne zależne.

Rozdział dziesiąty koncentruje się na optymalizacji procesu WEDM. Omówiono w nim kryteria optymalizacji, funkcję celu oraz przestrzeń poszukiwań, a także

przeprowadzono analizę porównawczą algorytmów optymalizacyjnych. Przedstawiono również wyniki optymalizacji i ich weryfikację eksperymentalną.

W rozdziale jedenastym opisano opracowanie i wdrożenie oprogramowania do optymalizacji procesu WEDM. Przedstawiono architekturę oprogramowania, wykorzystaną technologię oraz interfejs użytkownika.

Pracę kończy rozdział dwunasty, w którym przedstawiono podsumowanie wyników, wnioski z przeprowadzonych badań oraz zarysowano kierunki dalszych badań.

2.3. Pojęcia podstawowe

W zagadnieniach związanych z uczeniem maszynowym i sztuczną inteligencją, zarówno w literaturze polskiej, jak i anglojęzycznej, często pojawia się problem z odpowiednio ścisłym definiowaniem pojęć. Na przykład, terminy takie jak „model” i „algorytm” są często używane zamiennie i mają różne znaczenia w zależności od kontekstu.

W związku z tym, w niniejszym podrozdziale zostaną przedstawione definicje wybranych pojęć oraz skrótów myślowych (wyróżnione pogrubioną czcionką) zastosowanych w pracy.

- **Model**

Pojęcie „model” w pracy występuje zgodnie z definicją przedstawioną w [2]. Zgodnie z tą definicją model to formalna reprezentacja funkcji matematycznej f , która mapuje zbiór wejściowych zmiennych niezależnych na wartość zmiennej zależnej.

A zatem np. występujące w pracy wyrażenie „model regresji liniowej” należy rozumieć jako formalną reprezentację funkcji regresji liniowej.

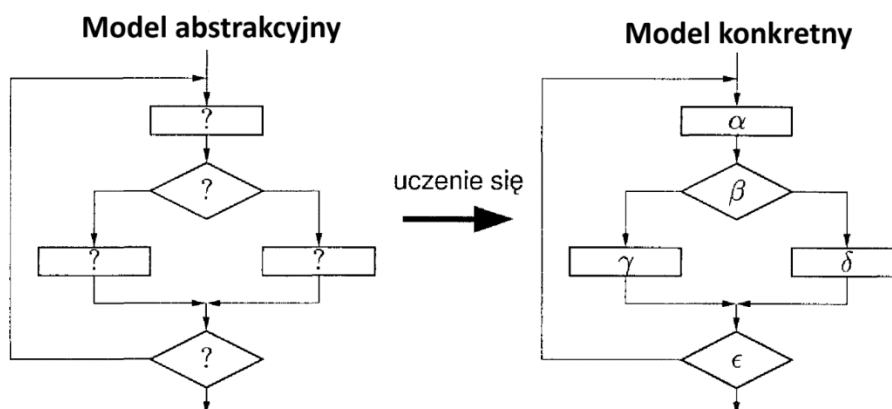
- **Jakość modelu**

Jakość modelu to wielowymiarowe pojęcie, które obejmuje różne aspekty oceny modelu w zależności od jego zastosowania i wymagań analitycznych [3].

W pracy jakość modelu określa zbiór wskaźników (metryk) jakości, które wyrażają przykładowo dokładność przewidywania wartości zmiennej zależnej, zdolność do rozróżniania klas lub powtarzalność wartości generowanych przez model (zob. 6.3). W szczególności, rozpatrywana jest także zdolność generalizacji, rozumiana jako zdolność uogólnienia wyuczonej umiejętności na inne, podobne zadania, nieprezentowane w trakcie uczenia [4].

- **Trenowanie modelu**

Modele podlegają procesowi „trenowania” lub zamiennie „uczenia się”. Trenowanie modelu oznacza wykorzystanie dostępnych danych do optymalizacji parametrów modelu w odniesieniu do funkcji straty, która określa, jak dobrze model przewiduje wartości zmiennej zależnej [2]. W ujęciu [5] model przed trenowaniem jest modelem abstrakcyjnym, natomiast po trenowaniu staje się modelem konkretnym, np. modelem przewidującym chropowatość powierzchni. Proces treningu jako konkretyzacja modelu został przedstawiony na Rys. 2.1 [5].



Rys. 2.1 Uczenie się jako konkretyzacja modelu [5]

W pracy jako synonim „modelu konkretnego” stosuje się pojęcie „model wytrenowany”.

- **Przewidywanie modelu**

Pojęcie „przewidywanie modelu” odnosi się w pracy do wartości generowanej na wyjściu modelu na podstawie danych wejściowych. Wartość ta jest interpretowana jako prognoza „stanu” lub charakterystyki danej zmiennej, np. przewidywana chropowatość powierzchni. Użycie terminu „przewidywana” jest uzasadnione tym, że model, na podstawie dostarczonych danych wejściowych, estymuje wartość zmiennej zależnej, a nie podaje jej rzeczywistego wyniku. W tym sensie model nie zna prawdziwej wartości zmiennej, a jedynie „przewiduje” ją na podstawie danych wejściowych i reguł ustalonych podczas procesu uczenia. Wynik ten stanowi szacunek oparty na wzorcach odkrytych w danych treningowych, co odróżnia go od bezpośredniej obserwacji. Wartość ta może przyjmować różne formy, w zależności od rodzaju modelu: od wartości ciągłej w przypadku regresji po przypisanie do określonej klasy w przypadku klasyfikacji.

- **Algorytm uczenia**

W procesie trenowania modelu wykorzystywane jest pojęcie „algorytm uczenia”. Algorytm ten rozumiany jest jako konkretna procedura lub zestaw reguł stosowanych w procesie trenowania modelu, np. algorytm bazujący na spadku gradientu.

2.4. Narzędzia programistyczne i środowisko obliczeniowe użyte w badaniach

Zdecydowaną większość prac zrealizowano przy użyciu języka programowania Python. Język ten, w połączeniu z dedykowanymi bibliotekami, stanowił kluczowe narzędzie do modelowania, optymalizacji, analizy danych oraz wdrożenia oprogramowania, wspierając wszystkie etapy realizacji niniejszej pracy. Dodatkowo, do programowania interfejsu użytkownika wykorzystano język JavaScript, w szczególności framework React Native.

Wykaz najważniejszych bibliotek i frameworków oraz ich zastosowania przedstawiono w Tabeli 2.1.

Tabela 2.1 Wykaz ważniejszych bibliotek i frameworków oraz ich zastosowanie w pracy

Biblioteka/Framework	Zastosowanie w pracy
Pandas	Przetwarzanie i analiza danych, operacje na tabelach
Random	Generowanie liczb losowych, losowanie próbek
NumPy	Operacje matematyczne
Matplotlib, Seaborn	Tworzenie wykresów, diagramów, wizualizacje danych
Statsmodels	Analiza statystyczna, testy statystyczne
Scikit-learn	Implementacja modeli klasyfikacji i regresji
Imblearn	Przetwarzanie niezbalansowanych danych
TensorFlow	Implementacja sieci neuronowych
Deap, Scipy	Implementacja algorytmów ewolucyjnych, optymalizacja
Doepy	Projektowanie planów eksperymentów
Django	Opracowanie backendu oprogramowania
React Native	Opracowanie aplikacji mobilnej

3. PODSTAWY OBRÓBKİ ELEKTROEROZYJNEJ

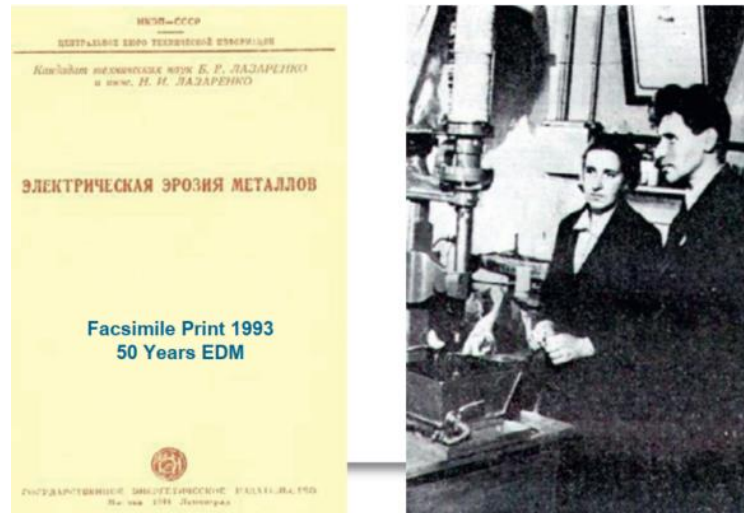
Obróbka elektroerozyjna (ang. Electrical Discharge Machining, EDM) jest technologią obróbki materiałów, która wykorzystuje zjawisko erozji elektrycznej zachodzącej pomiędzy elektrodą narzędziową a obrabianym przedmiotem w środowisku dielektrycznym. Proces ten umożliwia kształtowanie materiałów trudno skrawalnych poprzez kontrolowane wyładowania elektryczne, które powodują miejscowe topienie i odparowywanie materiału.

W niniejszym rozdziale przedstawiono wybiórczy przegląd podstawowych aspektów obróbki elektroerozyjnej, z naciskiem na specjalistyczną odmianę tej technologii, jaką jest cięcie elektroerozyjne drutem (WEDM). Zaprezentowano syntetyczny przegląd historii oraz podstaw teoretycznych związanych z tą technologią.

Na początku rozdziału (p. 3.1) przedstawiono zwięzły przegląd historii i rozwoju obróbki elektroerozyjnej, koncentrując się na kluczowych momentach oraz innowacjach, które wpłynęły na jej współczesną formę. W dalszej części (p. 3.2) zaprezentowano skrócony opis teoretycznych i technologicznych podstaw cięcia elektroerozyjnego drutem, ze szczególnym uwzględnieniem mechanizmu usuwania materiału (p. 3.2.1), który obejmuje wybrane fizyczne i chemiczne zjawiska zachodzące w trakcie procesu. Kolejno omówiono kluczowe parametry procesu WEDM (p. 3.2.2), takie jak napięcie, prąd i czas trwania impulsu oraz charakterystykę impulsu elektrycznego (p. 3.2.3). Ostatnia część rozdziału (p. 3.2.4) zawiera skrótowy opis budowy obrabiarki do cięcia elektroerozyjnego.

3.1. Historia i rozwój obróbki elektroerozyjnej

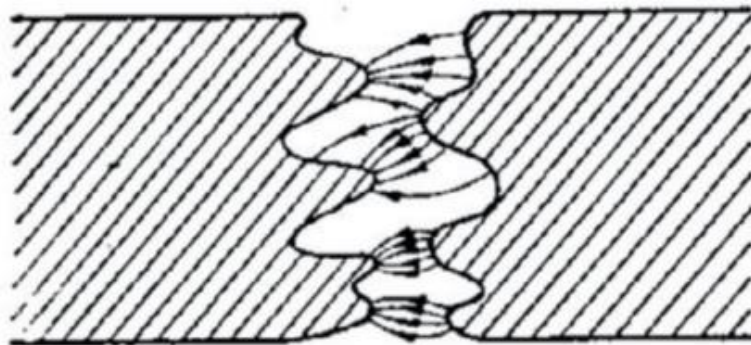
Obróbka elektroerozyjna to technologia, która rozwijała się przez dekady, napędzana zarówno przez potrzeby rynkowe, jak i postęp naukowy. Początki tej metody sięgają XVIII wieku, kiedy Benjamin Priestley w 1751 roku zauważył efekty erozji metalu spowodowane wyładowaniami elektrycznymi [6]. Jednakże, przełomowy rozwój technologii elektroerozyjnej rozpoczął się dopiero w latach 40. XX wieku, dzięki pracom rosyjskich naukowców, B.R. Lazarenko i L.I. Lazarenko [7] (Rys. 3.1).



Rys. 3.1 Wydana w 1943 roku praca B.R. Lazarenko : „Inversion of the erosion of metals and measured against contact devastation” [7]

Opracowali oni pierwsze urządzenie do elektroerozyjnej obróbki materiałów, wykorzystując układ zasilania typu rezystancyjno-pojemnościowego. Ich innowacyjne podejście pozwoliło na precyzyjne kontrolowanie procesu erozji materiału [6].

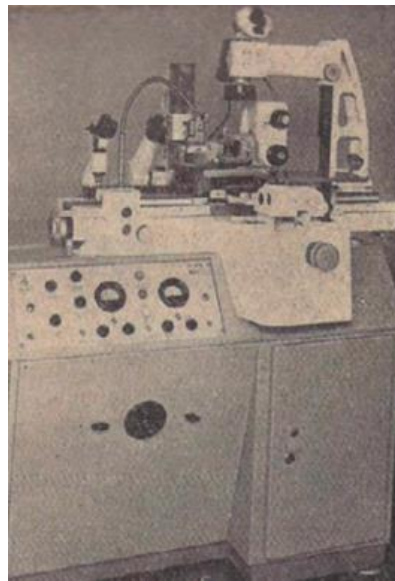
Późne lata czterdzieste, po katastrofalnej wojnie, nie pozwoliły na szybki rozwój maszyn, jednak badania nad obróbką elektroerozyjną rozpoczęły się na całym świecie. B.N. Sołotych, jako asystent Lazarenko, przeprowadził pierwszą systematyczną analizę procesu elektroerozyjnego [8], opracowując przy tym teorię inicjacji wyładowania elektrycznego. Teoria ta zakłada, że wyładowanie elektryczne inicjowane jest przez zimną emisję elektronów, wywołaną wysokim natężeniem pola elektrycznego na mikroskopijnych wierzchołkach powierzchni elektrod, które są najbliżej przeciwelektrody (Rys. 3.2 [8]).



Rys. 3.2 Zapłon zainicjowany emisją zimnych elektronów [8]

W latach 50. i 60. XX wieku zarówno w Związku Radzieckim, jak i krajach Zachodu, prowadzono intensywne badania nad udoskonaleniem procesu EDM, a także na

zastosowaniu wyników tych prac w praktyce warsztatowej. W Instytucie Badawczym ENIMS w Moskwie, we współpracy z Instytutem Projektowania Obrabiarek i Grupą Rozwoju Generatorów (ZNIL-Elektrom), opracowano maszyny, technologię, akcesoria i materiały elektrodowe. Dyrektor grupy EDM, prof. Livshiz, opublikował również w 1957 r. przegląd różnych wariantów obróbki elektroerozyjnej, a zwłaszcza wczesnych osiągnięć i zastosowań EDM [6]. W przeglądzie tym m.in. wyróżniono obróbkę elektroerozyjną wykorzystującą miedziany drut jako elektrodę roboczą (WEDM). Pierwsze maszyny wykorzystujące technologię WEDM były zbudowane na bazie precyzyjnych maszyn pomiarowych niemieckiego producenta Leitz. Przykład takiej maszyny przedstawiono na Rys. 3.3 [6].



Rys. 3.3 Jedna z pierwszych maszyn EDM wykorzystująca drut jako elektrodę roboczą (1964) [6]

W 1954 roku doświadczeni konstruktorzy obrabiarek z Europy Zachodniej wznowili swoją obecność na rynku, biorąc udział w I Europejskiej Wystawie Obrabiarek w Mediolanie. Marc Bruma z Francji wygłosił wykład na temat obróbki elektroerozyjnej, a firmy takie jak Sparcatron, AGIE i Charmilles zaprezentowały nowo opracowane maszyny. Przedstawione przez nich osiągnięcia wzbudziły szerokie zainteresowanie. W rezultacie zainicjowano intensywne badania nad tą technologią w krajach Zachodu, co doprowadziło m.in. do powołania Komitetu Naukowo-Technicznego ds. Obróbki Elektroerozyjnej (STC-E) [6].

W latach 60. trzy kluczowe osiągnięcia naukowe dały istotny impuls rozwojowi EDM. W 1959 roku Jean Hoerni wynalazł technologię tranzystora planarnego, którą Robert Noyce wykorzystał do opracowania pierwszego układu scalonego. Wynalazki te stworzyły fundament dla rozwoju technologii układów scalonych, umożliwiając opracowanie

statycznych generatorów impulsów dla EDM. W 1965 roku firma Digital Equipment Corporation (DEC) zrealizowała, na bazie układów scalonych, pierwszy komputer procesowy, co otworzyło drogę do rozwoju kompaktowych maszyn, serwokontrolerów oraz urządzeń programowalnych. W 1971 roku firma Intel zaprezentowała pierwszy mikroprocesor, który stał się podstawą do opracowania sterowania numerycznego i metod programowania sekwencji pracy maszyn [6]. Firmy AGIE i Charmilles ze Szwajcarii jako pierwsze wyposażyły swoje maszyny w generatory impulsów statycznych, co przyniosło znaczny postęp w postaci większej szybkości obróbki, mniejszego zużycia i lepszej precyzji, wynikającej ze zmniejszenia napięcia obwodu otwartego (70-100V) i zmniejszonej szerokości szczeliny [6].

Lata siedemdziesiąte przyniosły wiele udoskonaleń obróbki EDM. Rozwinął się rynek wsparcia dla precyzyjnych uchwytów, bezpiecznych płynów dielektrycznych, skutecznych urządzeń filtrujących, dostaw wstępnie obrobionych elektrod itp. [4].

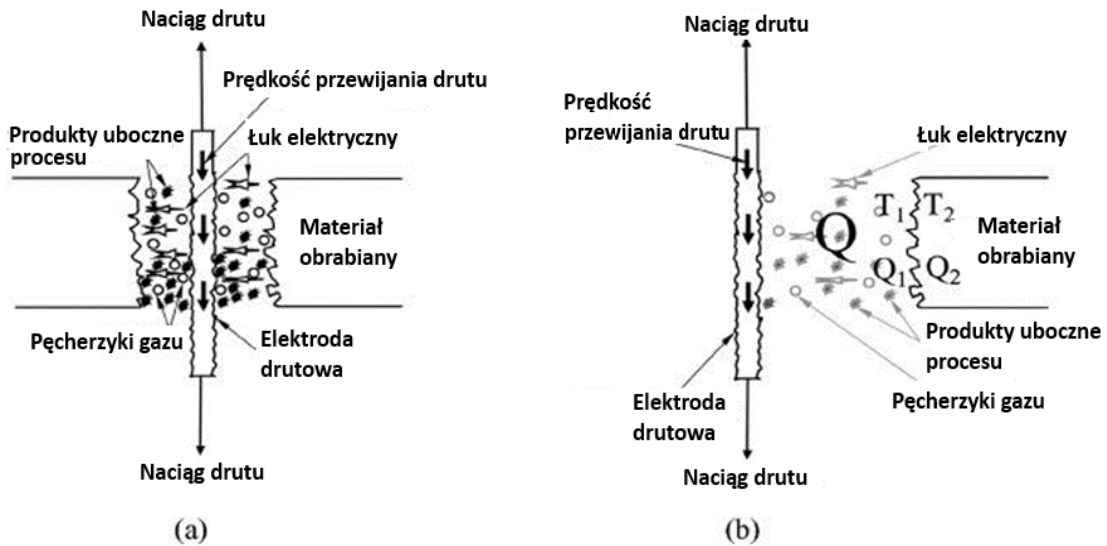
Lata 80. przyniosły rewolucję w technologii obróbki elektroerozyjnej dzięki wprowadzeniu komputerowego sterowania numerycznego (CNC). Wprowadzenie technologii CNC do EDM umożliwiło automatyzację wielu etapów procesu, od umieszczania elektrody po uzyskanie końcowej, precyzyjnie obrobionej powierzchni. Dzięki temu możliwe stało się masowe wytwarzanie skomplikowanych i precyzyjnych komponentów o wysokiej jakości [4].

3.2. Cięcie elektroerozyjne drutem

3.2.1. Mechanizm usuwania materiału

Mechanizm usuwania materiału w metodzie WEDM w sposób schematyczny przedstawia Rys. 3.4 [9]. Proces ten jest zbliżony do konwencjonalnego procesu EDM i polega na efekcie erozji wywołanej przez wyładowania elektryczne. Materiał usuwany jest z materiału obrabianego za pomocą wyładowań elektrycznych, które występują pomiędzy przedmiotem obrabianym a drutem, stale podawanym do strefy obróbki [10]. Przepływ energii elektrycznej prowadzi do utworzenia kanału plazmowego pomiędzy katodą a anodą, w wyniku czego energia elektryczna ulega konwersji na energię cieplną (Q) [11], osiągając temperatury w zakresie od 8000 do 12000 °C, a w niektórych przypadkach nawet do 20000 °C [12]. Energia cieplna dzieli się na dwie składowe: Q_1 , która jest absorbowana przez dielektryk odpowiedzialny za chłodzenie strefy obróbki, oraz Q_2 , która jest przewodzona w głąb materiału. Temperatura na powierzchni obrabianego materiału (T_1) jest znacznie wyższa niż w jego głębszych warstwach (T_2), co prowadzi do powstania strefy wpływu ciepła (ang. Heat Affected Zone, HAZ) i zmian strukturalnych w materiale. Wysoka temperatura powoduje intensywne ogrzewanie i topienie materiału na powierzchni obu elektrod. W momencie, gdy pulsacyjne zasilanie prądem stałym

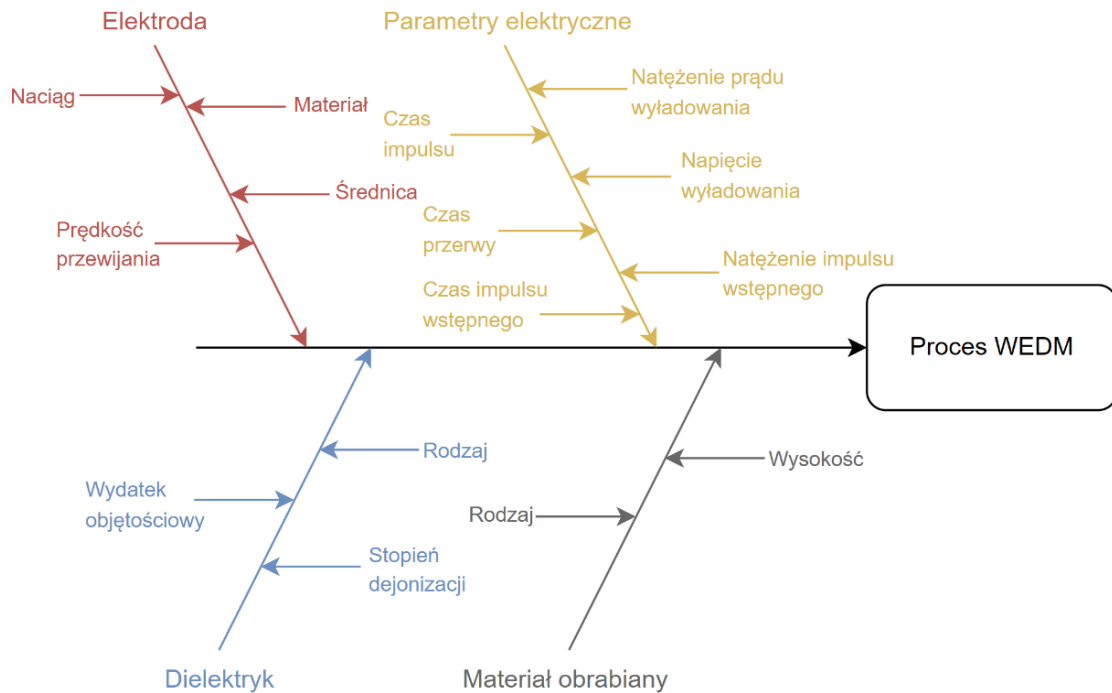
o częstotliwości od 20000 do 30000 Hz [12] zostaje wyłączony, kanał plazmowy ulega przerwaniu, co prowadzi do gwałtownego obniżenia temperatury. Wówczas cyrkulujący płyn dielektryczny przedostaje się do miejsca, gdzie wcześniej znajdował się kanał plazmowy, wypłukując stopione cząstki z powierzchni elektrod i usuwając je w postaci mikroskopijnych zanieczyszczeń [13].



Rys. 3.4 Mechanizm usuwania materiału podczas obróbki WEDM: (a) procesy mechaniczne i elektryczne oraz (b) procesy termiczne [9].

3.2.2. Charakterystyka podstawowych parametrów procesu

Podstawowe parametry i warunki obróbki występujące w procesie WEDM przedstawiano poniżej na diagramie przyczynowo skutkowym (Rys. 3.5 [14]).



Rys. 3.5 Podstawowe parametry i warunki obróbki występujące w procesie WEDM [14]

Parametry procesu WEDM można podzielić na elektryczne oraz nieelektryczne. Do najważniejszych parametrów elektrycznych należą:

- **Czas trwania impulsu** (ang. pulse on time, t_{on}) [μs] - określa czas, przez jaki prąd płynie przez materiał podczas jednego impulsu.
- **Czas wyłączenia impulsu** (ang. pulse off time, t_{off}) [μs] - określa przerwę między kolejnymi impulsami.
- **Prąd szczytowy** (ang. peak current, \hat{I}_e) [A] - maksymalne natężenie prądu elektrycznego podczas impulsu [15].
- **Napięcie otwartej przerwy** (ang. open-gap voltage, \hat{u}_i) [V] - odnosi się do kontrolowanego progu różnicy potencjałów, który jest zaprogramowany w generatorze impulsów. Napięcie to jest dostarczane do obwodu elektrycznego w momencie włączenia zasilania i służy do inicjacji wyładowania elektrycznego, gdy elektroda zbliża się do obrabianego materiału [15].
- **Napięcie wyładowania** (ang. discharge voltage, u_e) [V] - odnosi się do różnicy potencjałów mierzonej w szczelinie między elektrodą a obrabianym materiałem. Szczelina ta, nazywana szczeliną iskrową (ang. spark gap) jest kontrolowaną odległością, w której dochodzi do powstania kanału plazmowego. Typowa wartość szczeliny iskrowej wynosi od 10 do 100 μm . Aby zapewnić efektywną obróbkę i płukanie, odległość ta musi być ciągle kontrolowana podczas procesu [15].

Czas trwania impulsu, napięcie wyładowania oraz prąd szczytowy definiują energię wyładowania, którą można ująć w sposób formalny. W tym celu w pracy [16] wprowadzono następującą formułę:

$$W_e = \int u_e(t) \cdot i(t) dt \approx \bar{u}_e \cdot \bar{i}_e \cdot t_{on}, \quad (3.1)$$

gdzie:

- W_e – energia wyładowania,
- $u_e(t)$ – napięcie w funkcji czasu,
- $i(t)$ – natężenie prądu w funkcji czasu,
- \bar{u}_e – średnie napięcie wyładowania,
- \bar{i}_e – średnie natężenie prądu wyładowania.

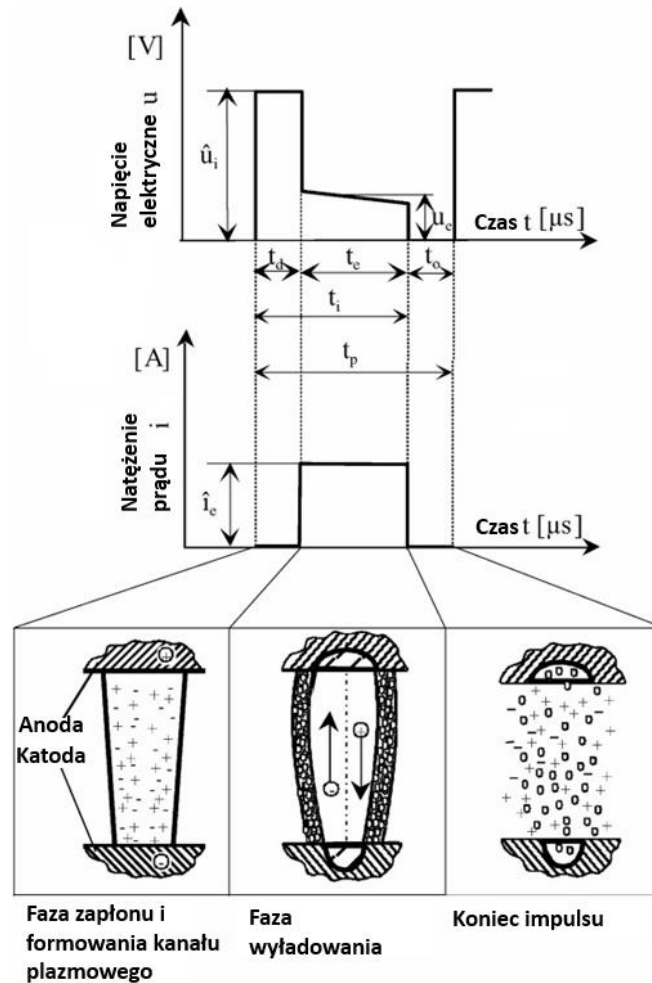
Powyższa formuła pokazuje, że energia wyładowania jest bezpośrednio zależna od wartości średniego napięcia, średniego prądu oraz czasu trwania wyładowania. Szczegóły charakterystyki impulsu elektrycznego przedstawiono w p.3.2.3.

Do najważniejszych parametrów nieelektrycznych należą:

- **Prędkość przewijania drutu** (ang. wire feedrate, WF) [m/min] - określa szybkość, z jaką drut jest przesuwany przez przestrzeń obróbczą.
- **Natężenie przepływu dielektryka** (ang. dielectric flow rate, FR) [l/min] - określa ilość dielektryka dostarczanego do przestrzeni obróbki na minutę.
- **Naciąg drutu** (ang. wire tension, WT) [N] - odnosi się do siły, z jaką drut jest napinany podczas procesu obróbki.

3.2.3. Charakterystyka impulsu elektrycznego w procesie

Fazy impulsu elektrycznego w procesie WEDM mogą być przedstawione za pomocą schematu na Rys. 3.6 [10].



Rys. 3.6 Fazy pojedynczego wyładowania EDM [10]

Zasadniczo cały proces może być opisany za pomocą trzech faz [15]:

- 1) **Faza zapłonu** (ang. ignition time) jest etapem inicjacji wyładowania elektrycznego między elektrodą drutową a obrabianym elementem. Początkowo, napięcie szybko rośnie, osiągając wartość szczytową \hat{u}_i , podczas gdy natężenie pozostaje na bardzo niskim poziomie lub jest nieobecne. W miarę jak napięcie rośnie, pole elektromagnetyczne zaczyna przyciągać i polaryzować jony, co prowadzi do ich migracji z katody do anody. Temperatura w szczelinie wzrasta, a rezystywność dielektryka maleje, aż do momentu, gdy kanał plazmowy jest w pełni uformowany. W tym momencie szczelina osiąga minimalną wartość rezystywności, co oznacza zakończenie fazy zapłonu i przejście do fazy wyładowania. Czas trwania tej fazy jest oznaczony jako t_d .
- 2) **Faza wyładowania** (ang. discharge phase) rozpoczyna się natychmiast po zakończeniu fazy zapłonu, gdy kanał plazmowy jest już uformowany. W tym momencie natężenie prądu elektrycznego szybko rośnie do swojej wartości

szczytowej \hat{I}_e i utrzymuje się na tym poziomie przez cały czas trwania impulsu t_{on} . Napięcie w tej fazie może nieznacznie spadać z powodu zmniejszenia rezystywności kanału plazmowego. Wysoka energia i ciśnienie w kanale plazmowym prowadzą do topienia i odparowania materiałów elektrody drutowej oraz obrabianego elementu. Jednocześnie następuje gwałtowne rozszerzenie bańki gazowej wokół kanału plazmowego. Faza ta jest krytyczna dla procesu obróbki, ze względu na właściwą erozję materiału w tej fazie

- 3) **Faza końca impulsu** (ang. end of pulse) następuje po zakończeniu fazy wyładowania. W tej fazie zasilanie elektryczne jest przerywane, co powoduje gwałtowne zmniejszenie natężenia prądu oraz napięcia elektrycznego do zera. Ta faza jest znana jako czas przerwy między impulsami t_{off} i trwa do momentu rozpoczęcia kolejnego impulsu. Podczas tej przerwy dielektryk przepływa przez szczelinę, usuwając odpady i zanieczyszczenia oraz chłodząc obrabiany materiał. Czas przerwy pozwala na stabilizację procesu i przygotowanie powierzchni na kolejny cykl wyładowania, zapewniając ciągłość i efektywność obróbki. Czas trwania tej fazy jest oznaczony jako t_o .

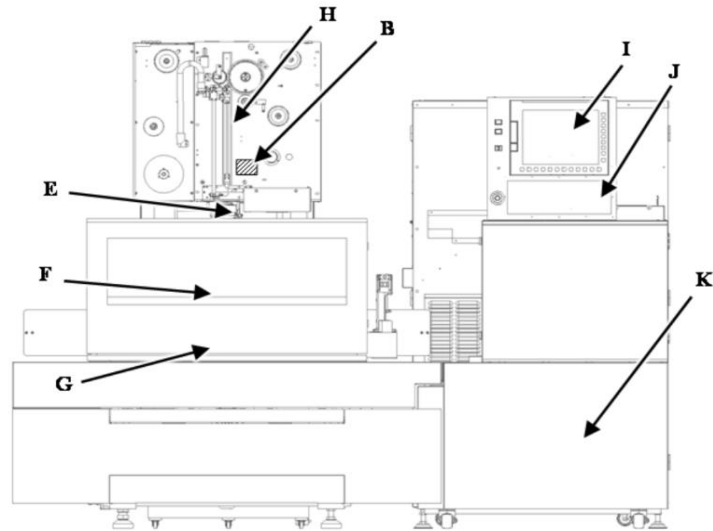
3.2.4. Budowa obrabiarki WEDM

Budowę obrabiarki WEDM przedstawiono na Rys. 3.7 [17] oraz Rys. 3.8 [18]. Maszyna ta różni się budową w zależności od producenta, można jednak wyróżnić kilka najważniejszych i wspólnych komponentów:

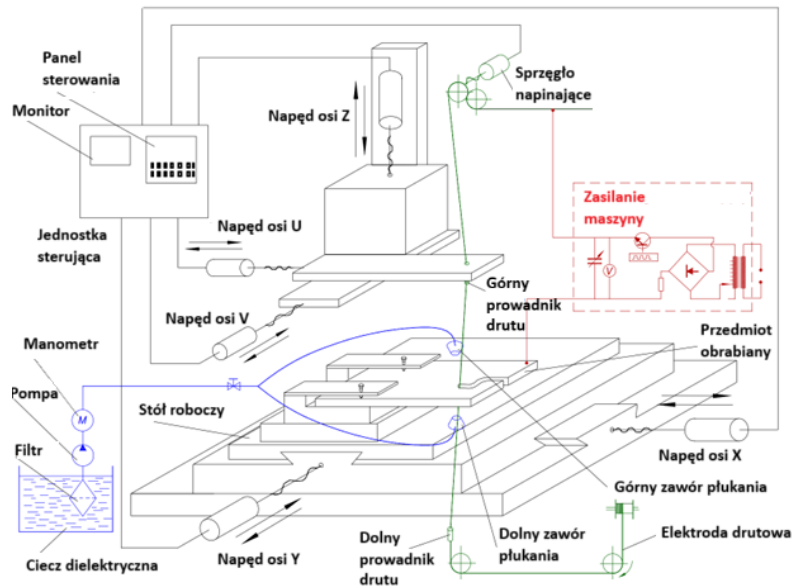
- System podawania drutu - zapewnia ciągły ruch elektrody drutowej wzdłuż zaprogramowanej trajektorii. Drut jest nawijany na szpulę, a następnie przeprowadzany przez zestaw prowadnic i rolek, które kontrolują jego napięcie i pozycję. Na Rys. 3.7 oznaczony jako „E” (głowica) i „H” (prowadnica drutu).
- Stół roboczy - platforma, na której mocowany jest obrabiany materiał. Stół jest zazwyczaj ruchomy w osiach X i Y, co pozwala na przemieszczanie obrabianego materiału względem drutu. Na Rys. 3.7 stół roboczy oznaczony jest jako „F” i „G”.
- System sterowania - nowoczesne obrabiarki WEDM są wyposażone w systemy sterowania CNC (ang. Computer Numerical Control), które pozwalają na programowanie i kontrolowanie procesu obróbki. Na Rys. 3.7 oznaczony jako „I” oraz „J”.
- Układ zasilania i generowania impulsu elektrycznego - dostarcza odpowiednie impulsy prądu elektrycznego, które generują wyładowania elektryczne między drutem a obrabianym materiałem. System ten musi być odpowiednio

kontrolowany, aby zapewnić odpowiednie parametry elektryczne, co jest kluczowe dla efektywności procesu erozji.

- Układ chłodzenia i dielektryka - obszar obróbki jest stale chłodzony i płukany dielektrykiem, który usuwa produkty erozji i zapobiega przegrzewaniu się materiału. Dielektryk jest pompowany przez system filtracji, który zapewnia jego czystość. Na Rys. 3.7 zbiornik dielektryka oznaczono jako „K”.



Rys. 3.7 Widok maszyny WEDM [17]



Rys. 3.8 Schematyczne przedstawienie maszyny WEDM [18]

4. PRZEGLĄD BADAŃ DOŚWIADCZALNYCH CIĘCIA ELEKTROEROZYJNEGO

Rozdział ten został poświęcony przeglądowi badań doświadczalnych z zakresu cięcia elektroerozyjnego. Zaprezentowano wyniki najnowszych badań, które ukazują wpływ różnych parametrów procesowych na jakość i wydajność obróbki, a także omówiono zastosowanie metod sztucznej inteligencji w modelowaniu i optymalizacji procesu WEDM.

W pierwszej sekcji rozdziału (p.4.1) omówiono badania dotyczące wpływu parametrów procesu WEDM na jakość i wydajność cięcia. Analiza obejmuje aspekty takie jak dokładność geometryczna (p.4.1.1), chropowatość powierzchni obrabianej (p.4.1.2), prędkość usuwania materiału (p.4.1.3), a także stabilność obróbki (p.4.1.4).

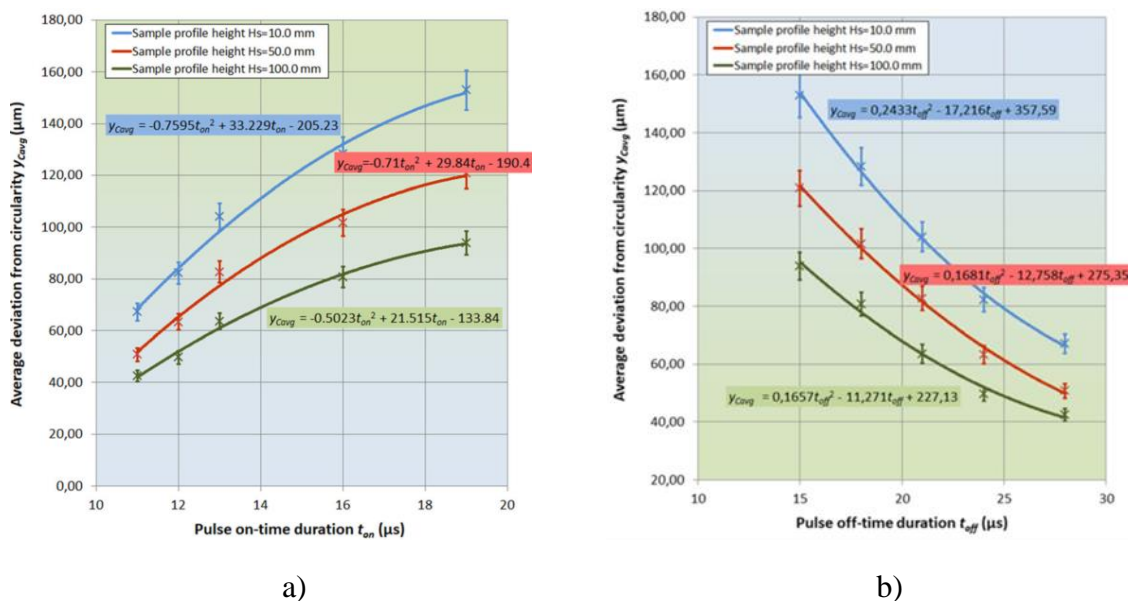
Druga część rozdziału (p.4.2) skupia się na wykorzystaniu metod sztucznej inteligencji w procesie WEDM. Sekcja ta została podzielona na dwa podrozdziały: pierwszy omawia metody sztucznej inteligencji w modelowaniu procesu WEDM (p.4.2.1), prezentując różne podejścia i ich skuteczność, natomiast drugi koncentruje się na optymalizacji procesu (p.4.2.2).

Na zakończenie rozdziału (p.4.3) zamieszczono podsumowanie, które syntetyzuje najważniejsze wnioski z przeprowadzonych badań, a także wskazuje potencjalne kierunki dalszych prac oraz możliwości rozwoju technologii WEDM.

4.1. Badania wpływu parametrów procesu WEDM

4.1.1. Dokładność geometryczna

Badania [19] udowodniły znaczący wpływ elektrycznych parametrów na dokładność geometryczną obrabianej powierzchni. Przykładowy rezultat badań przedstawiono poniżej.



Rys. 4.1 Zależność średnich odchyłek geometrycznych zaokrąglenia od t_{on} (a) i t_{off} (b) w poszczególnych profilach wysokości próbki eksperymentalnej [19].

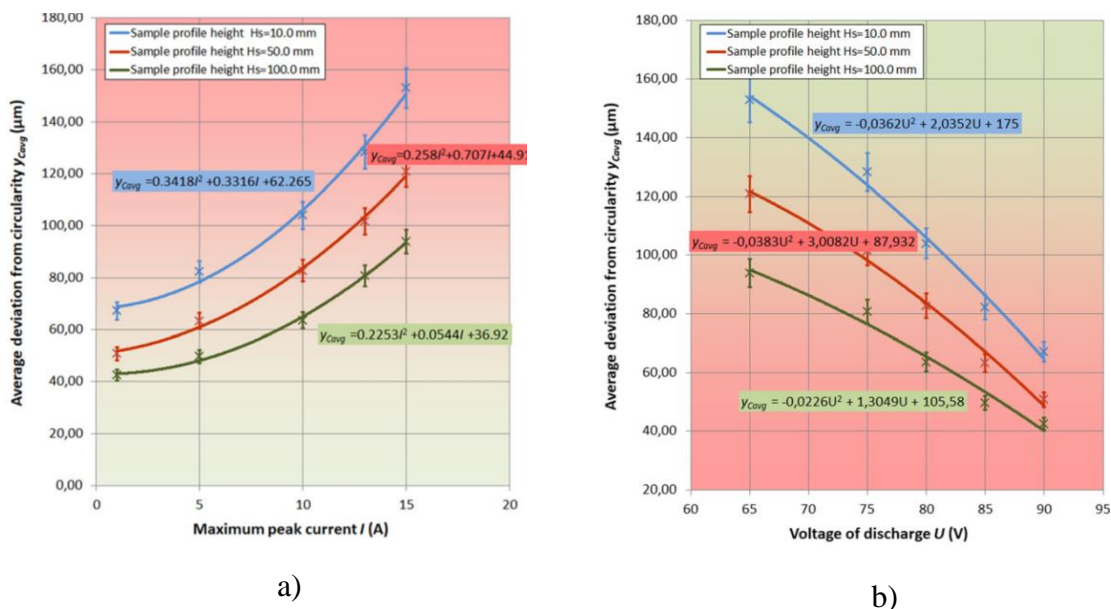
Wykresy Rys. 4.1a i Rys. 4.1b [19] przedstawiają zależność średnich odchyłek okrągłości γ_{avg} wycinanych próbek o różnej wysokości od odpowiednio czasu trwania impulsu (t_{on}) oraz czasu przerwy (t_{off}). Można łatwo zauważyć, że wzrost czasu trwania impulsu zwiększa średnie odchyłki okrągłości, szczególnie w próbkach o mniejszej wysokości. Wynikać to może z faktu, że wzrost czasu trwania impulsu dostarcza więcej energii do obszaru obróbki, co prowadzi do większej erozji materiału i zwiększa odchyłki okrągłości. Próbki o większej wysokości mają lepszą zdolność do odprowadzania ciepła, co zmniejsza deformacje termiczne i odchyłki okrągłości [19].

Natomiast wydłużenie czasu przerwy między impulsami zmniejsza średnie odchyłki okrągłości, co jest najbardziej zauważalne również w krótszych próbkach. Występuje tutaj zjawisko podobne do wpływu parametru t_{on} , jednak efekt jest odwrotny. Wzrostu wartości parametru t_{off} umożliwia lepsze chłodzenie i usuwanie produktów erozji, co zmniejsza deformacje termiczne i poprawia jakość powierzchni, redukując odchylenia od okrągłości.

Rys. 4.2a i Rys. 4.2b [13] przedstawiają zależność średnich odchyłek okrągłości od odpowiednio maksymalnego prądu szczytowego (\hat{I}_e) oraz napięcia wyładowania (u_e). Z wykresów wynika, że wraz ze wzrostem maksymalnego prądu szczytowego, średnie odchyłki okrągłości zwiększają się dla wszystkich trzech wysokości próbek. Zachodzi tutaj analogiczne zjawisko jak w przypadku wpływu t_{on} .

Odwrotną zależność można zaobserwować w przypadku badania wpływu napięcia wyładowania – wraz ze wzrostem u_e średnie odchyłki okrągłości maleją. Efekt ten wynika

prawdopodobnie z faktu, że zwiększenie napięcia poprawia stabilność wyładowań, co prowadzi do bardziej równomiernego usuwania materiału, co z kolei minimalizuje odchyłki okrągłości.

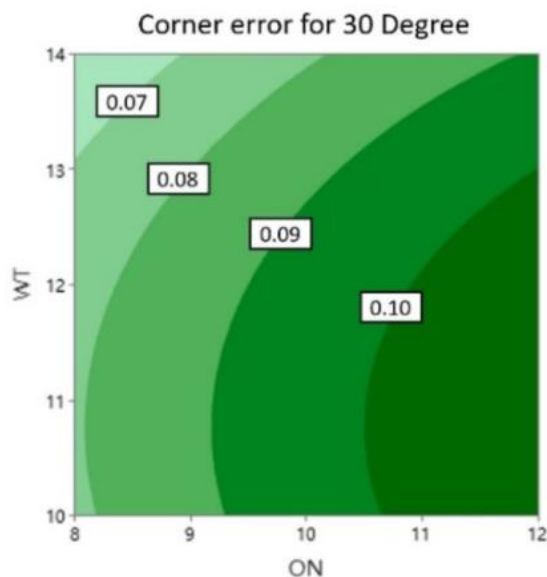


Rys. 4.2 Zależność średnich odchyłek geometrycznych zaokrąglenia od \hat{I}_e (a) i u_e (b) w poszczególnych profilach wysokości próbki eksperymentalnej [19].

Powyższe wyniki znajdują częściowe potwierdzenie w innych badaniach. Na przykład w pracy [20] badano m.in. wpływ parametrów elektrycznych na dokładność kształtu w kierunku profilu (Δr) i dokładność kształtu w kierunku normalnym do posuwu (Δd). Autorzy pracy [20] wykazali, że wyższy prąd szczytowy \hat{I}_e powoduje znaczący wzrost odchyłek Δr , co pokrywa się z wynikami uzyskanymi w pracy [19].

Zaobserwowano jednak istotną interakcję parametrów \hat{I}_e i t_{off} , prowadzącą do wzrostu odchyłek Δd przy niskich wartościach t_{off} i wysokich wartościach \hat{I}_e , natomiast wzrost parametrów \hat{I}_e i u_e prowadzi do minimalizacji odchyłek Δd .

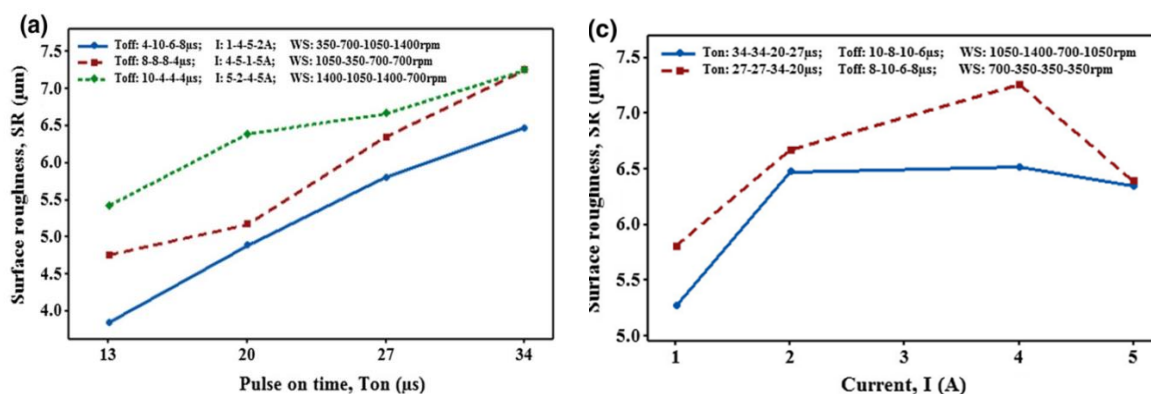
Spśród parametrów nieelektrycznych, które wywierają wpływ na dokładność geometryczną obrabianych powierzchni najczęściej wymieniany jest naciąg drutu oraz natężenie przepływu dielektryka. Na przykład wyniki badań [21] wskazują, że naciąg drutu przyczynia się do poprawy dokładności wycinanych narożników.

Rys. 4.3 Kontur 2D błędu narożnika dla 30° w zależności od t_{on} i WT [21]

Jak widać na powyższym rysunku wzrost naciągu drutu (od 10 do 14 N) powoduje zmniejszenie błędu narożnika. Wyższy naciąg drutu stabilizuje drut, co minimalizuje jego vibracje i przemieszczenia, prowadząc do większej precyzji w obróbce narożników.

4.1.2. Chropowatość powierzchni obrabianej

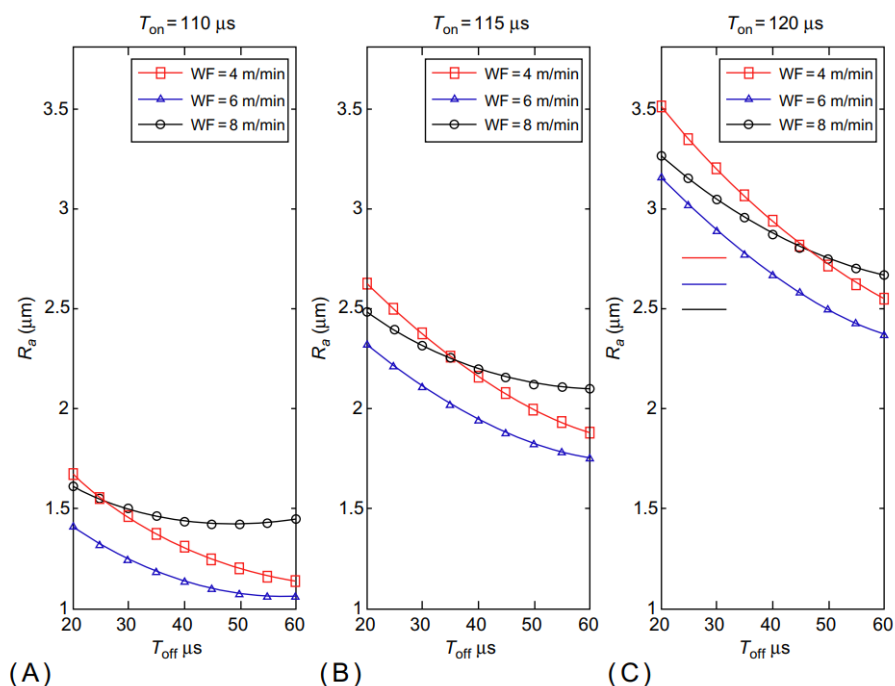
Inne prace badawcze, np. [13], [22] – [24] wskazują na znaczny wpływ parametrów obróbki na chropowatość powierzchni obrabianej. Przykładowy wynik badań w tym zakresie przedstawiono na Rys. 4.4 [24].

Rys. 4.4 Zmienność parametru chropowatości Ra w zależności od (a) t_{on} i (b) \hat{I}_e [24]

Jak przedstawiono na Rys. 4.4a, chropowatość powierzchni wzrasta wraz ze zwiększeniem t_{on} w całym obszarze eksperymentalnym. Jest to spowodowane większą ilością energii dostarczanej do miejsca obróbki, co prowadzi do intensywniejszego topienia i odparowywania materiału. To z kolei skutkuje powstawaniem głębszych kraterów

erozyjnych na powierzchni obrabianej, co bezpośrednio przekłada się na wyższą chropowatość powierzchni. Obserwacja ta znajduje swoje potwierdzenie w wielu innych pracach badawczych, np. [25],[26]

W przypadku analizy wpływu prądu szczytowego (\hat{I}_e) (Rys. 4.4c), chropowatość powierzchni wzrasta wraz ze zwiększaniem natężenia prądu, jednak według pracy [24] wzrost ten szybko zanika, a wpływ prądu szczytowego na chropowatość powierzchni stabilizuje się w zakresie 2-4A. Po przekroczeniu 4A zauważalna jest niewielka poprawa chropowatości. Początkowy wzrost chropowatości powierzchni wynika z faktu wzrostu dostarczonej ilości energii, co zwiększa prawdopodobieństwo tworzenia się kraterów erozyjnych na obrabianej powierzchni. Późniejsza stabilizacja i spadek chropowatości według [24] wynika z interakcji z innymi parametrami obróbki: t_{on} , t_{off} i WS. Wpływ tych parametrów na chropowatość powierzchni jest na tyle istotny, że w pewnym zakresie może niwelować wpływ wzrostu prądu szczytowego. Szczególnie w przypadku maksymalnej rozpatrywanej wartości t_{off} i WS następuje wyraźny spadek chropowatości. Analizę wpływu parametrów t_{on} , WS i t_{off} na chropowatość powierzchni przeprowadzano np. w pracy [27] (Rys. 4.5).

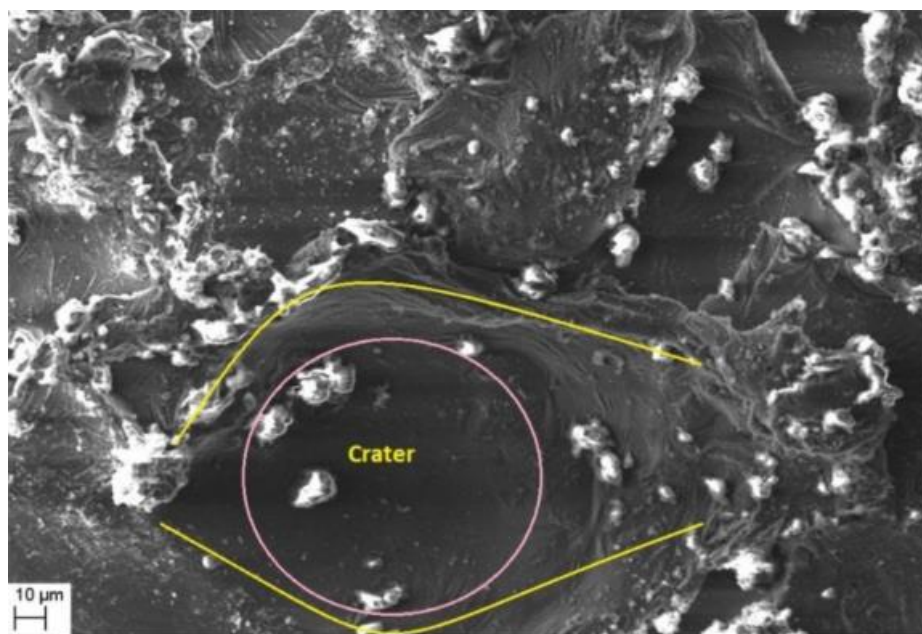


Rys. 4.5 Wpływ parametrów t_{off} , t_{on} i WF na chropowatość powierzchni [27]

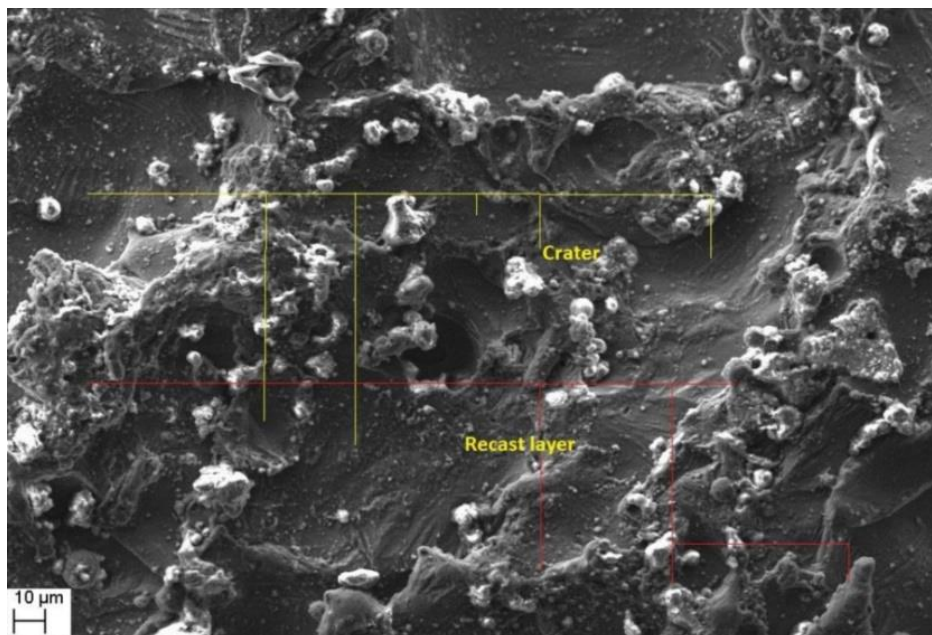
Jak widać na Rys. 4.5 wydłużenie czasu przerwy powoduje spadek chropowatości powierzchni. Dzieje się tak z powodu wydłużenia czasu odprowadzania ciepła, co ogranicza powstawanie defektów termicznych na powierzchni obrabianej [27]. Można

również zauważyć, że prędkość przewijania drutu (WF) ma znaczący wpływ na chropowatość powierzchni. Najlepsza jakość powierzchni jest osiągana przy prędkości przewijania drutu wynoszącej 6 m/min, podczas gdy niższa i wyższa prędkość prowadzi do zwiększenia chropowatości. Niska prędkość przewijania drutu prowadzi do ciągłego generowania wyładowań, co obniża prędkość cięcia. W efekcie powierzchnia staje się bardziej szorstka z powodu nierównomiernego rozmieszczenia oraz nachodzenia na siebie kraterów erozyjnych. Z kolei zbyt wysoka prędkość przewijania (WF) może skutkować nierównomiernym zużyciem drutu, co prowadzi do powstawania niejednorodnych kraterów na powierzchni obrobionej, pogarszając jej jakość.

Poniżej zaprezentowano wybrane zdjęcia z badań morfologicznych obrabianych powierzchni przeprowadzonych przy użyciu skaningowej mikroskopii elektronowej (SEM) (Rys. 4.6 [28], Rys. 4.7 [28]). Oba zdjęcia przedstawiają kratery erozyjne (ang. Crater), ponadto na Rys. 4.7 zaznaczono przetopioną warstwę materiału obrabianego (ang. Recast layer). Tego typu defekty powierzchniowe bezpośrednio wpływają na chropowatość powierzchni.



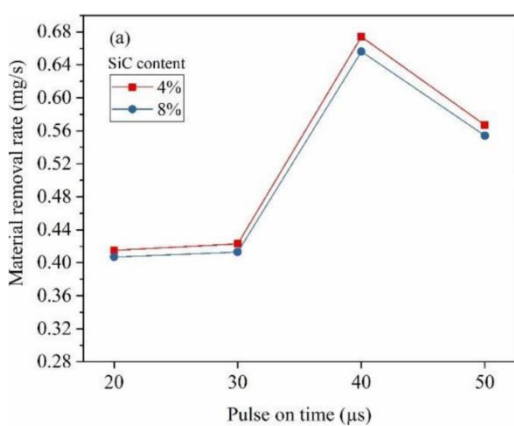
Rys. 4.6 Formacja krateru erozyjnego [28]



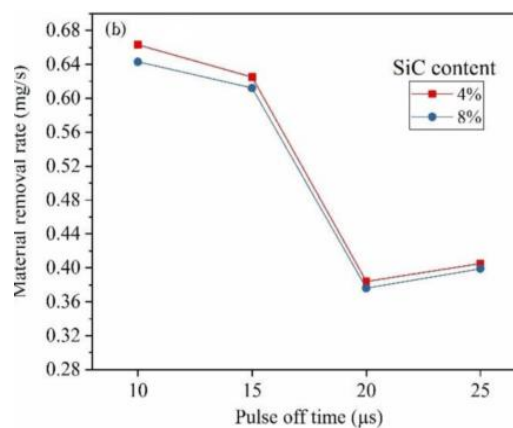
Rys. 4.7 Widok przetopionej warstwy materiału i kraterów erozyjnych [28]

4.1.3. Prędkość usuwania materiału

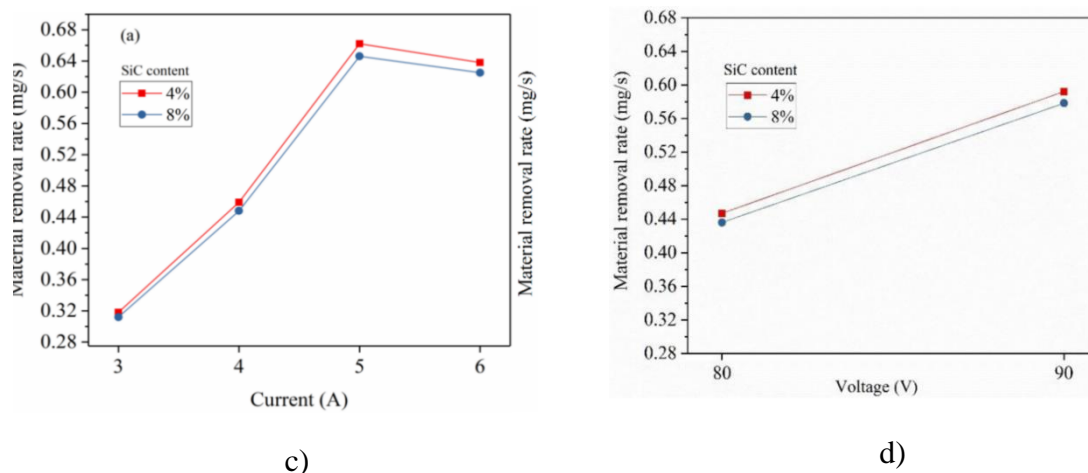
Parametry elektryczne mają także istotne znaczenie ze względu na wpływ na wydajność obróbki. Wpływ na prędkość usuwania materiału (ang. Material Removal Rate, MRR) został zilustrowany na Rys. 4.8 [28].



a)



b)



Rys. 4.8 (a) Wpływ t_{on} na MRR. (b) Wpływ t_{off} na MRR, (c) Wpływ \hat{I}_e na MRR, (d) Wpływ u_e na MRR [28]

Jak można łatwo zauważyć, na Rys. 4.8a, im dłuższy czas trwania impulsu (t_{on}) tym większa prędkość usuwania materiału. Zjawisko to powodowane jest zwiększeniem ilości energii dostarczanej do strefy cięcia, co prowadzi do intensywniejszego procesu erozji. Jednakże, po osiągnięciu pewnej wartości granicznej (w rozpatrywanym przypadku jest to 40 μ s) prędkość usuwania materiału może zacząć spadać. Zjawisko to może wynikać z nadmiernego nagrzewania materiału oraz tworzenia się większych defektów termicznych, które utrudniają efektywne usuwanie stopionego materiału [28].

Na Rys. 4.8b widoczne jest, jak skrócenie czasu t_{off} prowadzi do zwiększenia prędkości usuwania materiału. Wynika to z faktu, że krótsze przerwy między impulsami pozwalają na wyższą częstotliwość generowania wyładowań elektrycznych. Większa liczba wyładowań na jednostkę czasu oznacza intensywniejszy proces erozji, co skutkuje wyższym MRR.

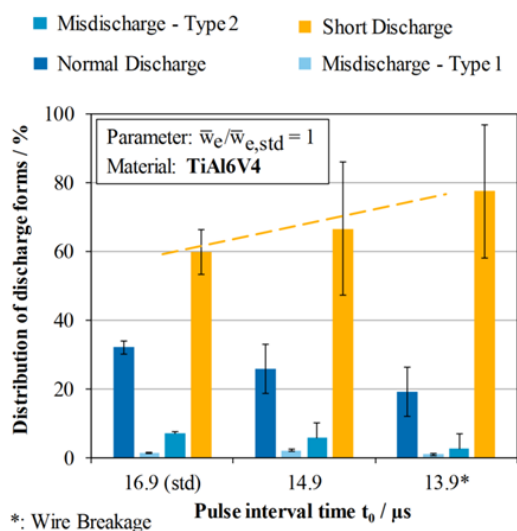
W przypadku \hat{I}_e , na Rys. 4.8c widać, że zwiększanie wartości prądu od 3A do 5A prowadzi do znaczącego wzrostu MRR z powodu dostarczania większej ilości energii, co intensyfikuje proces erozji. Najwyższy MRR osiągnany jest przy natężeniu 5A. Dalsze zwiększanie natężenia prądu do 6A powoduje niewielki spadek MRR, prawdopodobnie z powodu nadmiernego nagrzewania strefy obróbki i tworzenia się defektów termicznych.

W przypadku u_e , na Rys. 4.8d widać, że wzrost wartości napięcia od 80V do 90V prowadzi do liniowego wzrostu MRR, ponieważ większe napięcie również zwiększa energię wyładowania.

4.1.4. Stabilność obróbki

W pracy [29] poruszono temat analizy parametrów procesu w celu identyfikacji niestabilnych warunków obróbki. Artykuł wyraźnie wskazuje, że parametry t_{on} i t_{off} mają

kluczowy wpływ na stabilność procesu WEDM. Podkreślono, że niestabilne warunki procesu charakteryzują się zwiększoną liczbą nieprawidłowych wyładowań elektrycznych, takich jak krótkie wyładowania i niepełne wyładowania, które mogą prowadzić do przerwania drutu. Krótszy czas t_{off} może nie zapewniać wystarczającego czasu na efektywne odprowadzenie produktów erozji i ciepła z obszaru roboczego, co zwiększa ryzyko zwarć i przerw drutu. Z kolei parametr t_{on} zwiększa energię wyładowania, co zwiększa ryzyko przegrzania i uszkodzeń drutu.



Rys. 4.9 Dystrybucja form wyładowań elektrycznych w zależności od t_{off} [29]

Na Rys. 4.9 [29] przedstawiono wpływ czasu przerwy między impulsami na stabilność procesu obróbki elektroerozyjnej dla materiału TiAl6V4. Przy standardowym czasie t_{off} wynoszącym 16.9 μs dominują wyładowania normalne, co wskazuje na stabilny proces. Skrócenie t_{off} do 14.9 μs prowadzi do wzrostu liczby wyładowań krótkich, co wskazuje na spadek stabilności procesu. Dalsze skrócenie t_{off} do 13.9 μs powoduje dominację wyładowań krótkich i znaczną niestabilność procesu, co prowadzi do przerwania drutu [29].

4.2. Wykorzystanie metod sztucznej inteligencji w procesie WEDM

Metody sztucznej inteligencji stanowią istotny element badań nad obróbką elektroerozyjną. Wykorzystywane są one głównie do dwóch celów: modelowania procesu obróbki oraz jego optymalizacji. Oba te zagadnienia mają na celu poprawę jakości, wydajności procesu WEDM.

4.2.1. Metody sztucznej inteligencji w modelowaniu procesu

Znaczna część badań dotyczących modelowania procesu cięcia elektroerozyjnego koncentruje się na opracowywaniu matematycznego modelu tego procesu, często przy użyciu Metodologii Powierzchni Odpowiedzi (ang. Response Surface Methodology, RSM) lub metody Taguchi. RSM to zestaw metod statystycznych i matematycznych stosowanych do modelowania i analizy problemów, w których odpowiedź systemu lub procesu zależy od wielu zmiennych. Procedura ta obejmuje planowanie eksperymentu (ang. Design of Experiments, DOE), modelowanie procesu za pomocą regresji wielomianowej oraz techniki optymalizacyjne. Podstawowym celem RSM jest skonstruowanie modelu, który precyzyjnie opisuje związek między zmiennymi wejściowymi a zmienną wyjściową [30].

Jednakże, wraz z postępem technologicznym oraz zwiększeniem mocy obliczeniowej komputerów, coraz częściej stosowane są metody sztucznej inteligencji do opracowywania modeli procesu cięcia elektroerozyjnego (WEDM). Zastosowanie tych metod wynika z potrzeby zwiększenia dokładności i elastyczności modeli.

W pracy [31] rozpatrywano trzy różne podejścia do modelowania procesu WEDM: podejście oparte na sztucznych sieciach neuronowych (ang. Artificial Neural Network, ANN), maszynach wektorów nośnych (ang. Support Vector Machine, SVM) oraz RSM. Porównanie dokonano na podstawie analizy dokładności przewidywanych wartości przez model, w stosunku do wartości empirycznych chropowatości powierzchni. Analizę przeprowadzono z wykorzystaniem trzech metryk statystycznych: współczynnika determinacji (R-value), średniego absolutnego błędu procentowego (ang. Mean absolute percentage error, MAPE) oraz maksymalnego błędu (Δ_{max}). Wyniki przedstawiono w Tabeli 4.1.

Tabela 4.1 Skuteczność modeli ANN, SVM i RSM [20]

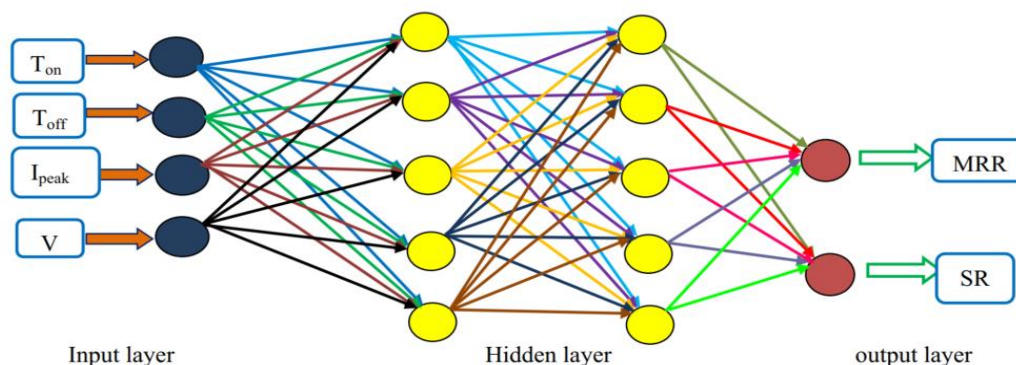
S. No.	Model	R-value	MAPE	Δ_{max}
1	ANN	0.9998	0.2541%	2.291%
2	SVM	0.9999	0.0347%	0.1427%
3	RSM	0.9873	3.577%	11.089%

Na podstawie przedstawionych wyników można stwierdzić, że model SVM przewyższa pozostałe modele pod względem dokładności przewidywania (najwyższa wartość R, najniższe wartości MAPE i Δ_{max}). Model ANN również wykazuje wysoką dokładność, jednak nieznacznie mniejszą niż SVM. Model RSM, okazał się najmniej dokładny spośród trzech porównywanych modeli, co widać po najniższej wartości R oraz najwyższych wartościach MAPE i Δ_{max} .

Podobne wyniki uzyskano w [32] stosując metody modelowania oparte na ANN oraz RSM do przewidywania wartości chropowatości powierzchni i prędkości cięcia. W obu przypadkach modele ANN okazały się dokładniejsze niż metody statystyczne (RSM).

Te i inne badania, np. [33]–[36], wyraźnie wskazują na istotne korzyści wynikające z wykorzystania metod sztucznej inteligencji do modelowania procesu obróbki elektroerozyjnej. Analiza wyników porównania modeli podkreśla przewagę tych metod nad tradycyjnymi podejściami modelowania, takimi jak Metodologia Powierzchni Odpowiedzi. W związku z tym badacze coraz częściej decydują się na rozszerzenie metody RSM, zastępując regresję wielomianową odpowiednią metodą z zakresu sztucznej inteligencji.

Prace badawcze, w których wykorzystuje się ANN, opierają swoje modele na podobnych architekturach jak przedstawiona na Rys. 4.10 [37]. Niemniej jednak, stosuje się różne modyfikacje w celu dostosowania modeli do specyficznych wymagań badawczych i warunków procesowych.



Rys. 4.10 Architektura wielowarstwowej sieci neuronowej do przewidywania MRR i chropowatości powierzchni (SR) [37]

W pracy [37] wykorzystano sztuczne sieci neuronowe (ANN) ze wsteczną propagacją błędów, w tym trzy różne podejścia: sieć neuronową z algorytmem uczenia Levenberga-Marquardta (LM), sieć neuronową regresji Elmana oraz sieć neuronową uogólnionej regresji (ang. Generalized Regression Neural Network, GRNN) do przewidywania wartości MRR i SR. Opierając się na analizie statystycznej dokładności proponowanych modeli, stwierdzono, że model LM z 10 neuronami w warstwie ukrytej jest najlepszym modelem ANN do przewidywania MRR i chropowatości powierzchni (ang. Surface roughness, SR). Podobne wyniki zaprezentowano w pracy [38], w której porównywano model ANN uczony algorytmem Levenberga-Marquardta z modelem ANN uczonym

metodą spadku gradientu, zastosowany do przewidywania parametru chropowatości Ra. W pracy tej również stwierdzono, że algorytm Levenberga-Marquardta jest najskuteczniejszą techniką uczenia modeli ANN, przewidujących chropowatość powierzchni.

Innym podejściem stosowanym w literaturze jest podejście hybrydowe, integrujące różne algorytmy z sieciami neuronowymi. W pracy [39] do przewidywania chropowatości powierzchni oraz MRR wykorzystano adaptacyjne systemy neuro-rozmyte (ang. Adaptive neuro fuzzy inference system, ANFIS). ANFIS to hybrydowy model łączący sieci neuronowe i logikę rozmytą. Wykorzystuje reguły językowe typu „jeśli-to” oraz dane treningowe do tworzenia systemu wnioskowania rozmytego [40],[41]. Tak utworzony model był porównywany z typowymi sieciami neuronowymi oraz modelem matematycznym opracowanym z użyciem RSM. Wyniki przedstawiono w Tabeli 4.2

Tabela 4.2 Ocena modeli RSM, ANN i ANFIS [39]

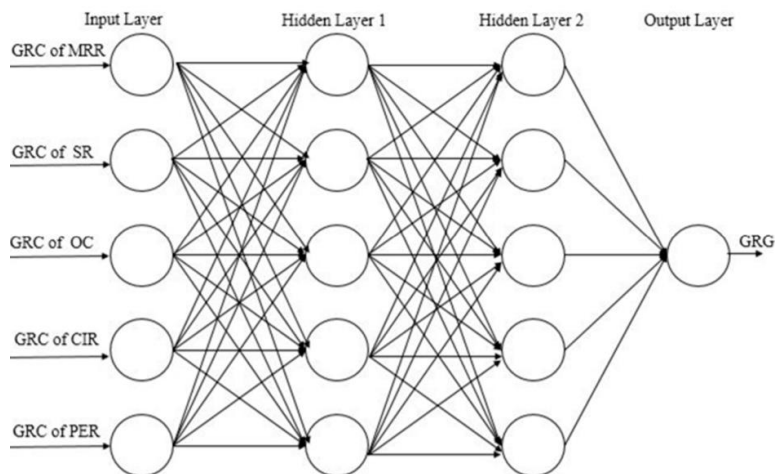
	Predicted (Rsm)		Predicted (ANN)		Predicted (ANFIS)	
	VMRR	Ra	VMRR	Ra	VMRR	Ra
Total percentage error	6.387%	4.301%	0.8738%	2.111%	1.539%	1.069%
R ²	97.34%	89.87%	99.72%	94.77%	99.48%	98.68%
Adj R ²	94.61%	83.74%	99.71%	94.59%	99.46%	98.59%
Predicted R ²	92.29%	72.70%	99.70%	94.23%	97.96%	97.84%

Porównanie oparto na kilku metrykach: całkowitym błędzie procentowym, określającym różnicę procentową między przewidywanymi wynikami a rzeczywistymi wartościami uzyskanymi z eksperymentów, współczynnika determinacji (R²), skorygowanym współczynnika determinacji (Adjusted R²) oraz przewidywanym współczynnika determinacji (Predicted R²). Z analizy wyników można wywnioskować, że oba modele ANN i ANFIS wykazują wysoką dokładność przewidywania. ANN uzyskał niższy błąd przewidywania i wyższe wartości współczynników determinacji dla VMRR, natomiast ANFIS lepiej sprawdza się w przypadku przewidywania chropowatości powierzchni.

W pracy [42] zastosowano integrację ANN z Teorią Szarych Systemów (ang. Grey System Theory, GST), aby opracować hybrydowy model który przewiduje prędkość usuwania materiału, chropowatość powierzchni (SR), odchylenie wymiarowe (CIR), odchyłki okrągłości (OC) oraz odchyłki prostopadłości (PER) wycinanych elementów.

Integracja ANN z GST w tym przypadku rozpoczyna się od normalizacji danych wejściowych, skalując je do zakresu od 0 do 1. Następnie dla każdej znormalizowanej wartości obliczane są współczynniki Grey Relational Coefficient (GRC), które odzwierciedlają relacje między danymi. Współczynniki GRC pełnią rolę zmiennych wejściowych do ANN, który przewiduje Grey Relational Grade (GRG). W ten sposób

model wykorzystuje kombinację metod do przewidywania wyników na podstawie skomplikowanych danych wejściowych. Model opracowany w pracy [42] został przedstawiony na schemacie poniżej.



Rys. 4.11 Schemat sieci neuronowych zintegrowanych z GST [42]

Badania przeprowadzone przez autorów pracy wskazują, że model Grey-ANN osiągnął wysoki współczynnik korelacji (0.9981) oraz niski pierwiastek błędu średniokwadratowego (0.0086).

Poza sieciami neuronowymi i ich integracjami z innymi metodami, stosowane są również inne podejścia wykorzystujące sztuczną inteligencję do modelowania procesu WEDM. Wśród nich znajdują się wcześniej przytoczone maszyny wektorów nośnych SVM, jak również proces Gaussa (ang. Gaussian Process, GP). W pracy [43] wykorzystano te metody oraz ANN, w celu opracowania modelu przewidywania chropowatości powierzchni. W Tabeli 4.3 [43] przedstawiono analizę jakości opracowanych modeli na zbiorze treningowym i testowym.

Tabela 4.3 Porównanie jakości modeli do przewidywania chropowatości powierzchni [43]

Training			Testing	
Model	CC	RMSE	CC	RMSE
GP Poly	0.9613	0.1069	0.8805	0.2270
GP Puk	1.0000	0.0000	0.8554	0.5646
GP RBF	0.9901	0.0533	0.8761	0.2565
SVM Poly	0.9654	0.1048	0.8605	0.3078
SVM PUK	1.0000	0.0013	0.8534	0.5695
SVM RBF	0.8529	0.3629	0.7016	0.5078
ANN	0.9989	0.0266	0.8582	0.2329

Do oceny modeli wykorzystano dwie metryki statystyczne: współczynnik korelacji (CC) i pierwiastek średniokwadratowego błędu (RMSE). Obejmuje ona zarówno etap treningowy, jak i testowy. Wyniki tych badań wykazały, że wszystkie modele (SVM, GP, ANN) prezentują dobre wyniki w przewidywaniu chropowatości powierzchni. W szczególności, proces Gaussa z jądrem wielomianowym (GP Poly) wykazuje najwyższą dokładność na zbiorze testowym.

Kolejnym podejściem jest wykorzystanie metod opartych na drzewach decyzyjnych. W pracy [44] wykorzystano drzewa decyzyjne oraz ich rozszerzenie, tj. metodę lasu losowego, do przewidywania zerwania drutu podczas obróbki. Badacze porównywali modele oparte na drzewach decyzyjnych do modelu regresji logistycznej i naiwnego klasyfikatora Bayesa. Model oparty na metodzie lasu losowego okazał się w tym przypadku najdokładniejszy w przewidywaniu zerwań drutu, osiągając dokładność klasyfikacji 100% dla danych testowych. Las losowy został również wykorzystany w pracy [45], gdzie wraz z drzewem decyzyjnym M5P posłużył do przewidywania MRR, chropowatości powierzchni oraz odchylen wymiarowych. Wyniki eksperymentów wykazały, że las losowy osiągnął najwyższe wartości współczynnika korelacji oraz najmniejsze wartości błędu średniokwadratowego i względnego błędu kwadratowego dla MRR oraz SR, potwierdzając jego wyższość nad modelami M5P. W przypadku odchylenia wymiarowego, najbardziej efektywny okazał się model M5P.

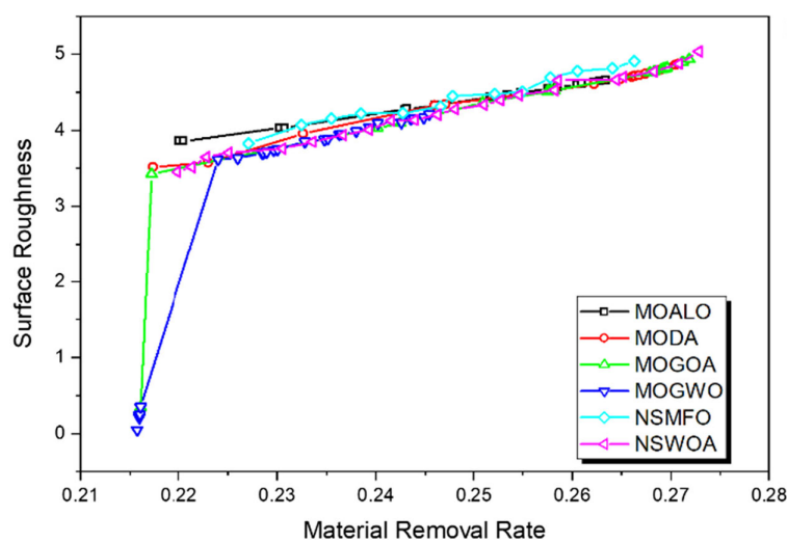
4.2.2. Metody sztucznej inteligencji w optymalizacji procesu

Opracowane modele procesu można następnie wykorzystać do optymalizacji parametrów technologicznych, ze względu na określone kryteria, w szczególności minimalizację chropowatości powierzchni obrabianej oraz maksymalizację prędkości usuwania materiału. Podobnie jak w przypadku badań modelowych, optymalizacja procesu WEDM w literaturze oparta jest często na tradycyjnych metodach. W przypadku optymalizacji jednokryterialnej większość prac badawczych opiera optymalizację procesu na metodach statystycznych, przede wszystkim metodzie Taguchi i RSM [46],[47]. Według pracy przeglądowej [48] metody te wyróżniają się najlepszymi wynikami pod względem efektywności czasowej, wymagań dotyczących mocy obliczeniowej, ograniczeń budżetowych oraz przystępności dla użytkownika. W przypadku optymalizacji wielokryterialnej stosowane są inne metody statystyczne, np. Desirability Function Analysis (DFA) [49], lub metody z zakresu teorii decyzji, takie jak Szara Analiza Relacyjna [50] i logika rozmyta [51] lub Analytic Hierarchy Process (AHP) [52].

Poza tradycyjnymi podejściami, w optymalizacji wielokryterialnej coraz częściej wykorzystuje się metody sztucznej inteligencji, w tym algorytmy inspirowane procesami biologicznymi [53]. Jak wykazano w [54], algorytmy te przewyższają np. metodę Taguchi

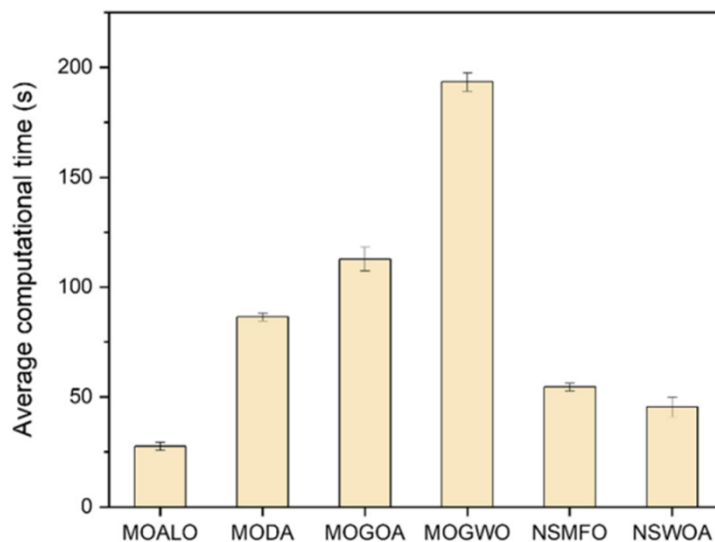
pod względem efektywnego przeszukiwania przestrzeni rozwiązań i unikania lokalnych minimów.

W tej grupie metod sztucznej inteligencji badacze korzystają z szerokiej gamy różnych algorytmów. W pracy [53] przeprowadzono analizę porównawczą sześciu różnych algorytmów rojowych: multi-objective ant lion optimization (MOALO), multi-objective dragonfly algorithm (MODA), multi-objective grasshopper optimization algorithm (MOGOA), multi-objective grey wolf optimizer (MOGWO), non-dominated sorting moth flame optimization (NSMFO) oraz non-dominated sorting whale optimization algorithm (NSWOA). Algorytmy porównywano w problemie optymalizacji parametrów technologicznych ze względu na maksymalizację MRR i minimalizację chropowatości powierzchni. Najlepszy front Pareto dla każdego z algorytmów przedstawiono na Rys. 4.12.



Rys. 4.12 Porównanie frontów Pareto dla SR vs. MRR algorytmów rojowych [53]

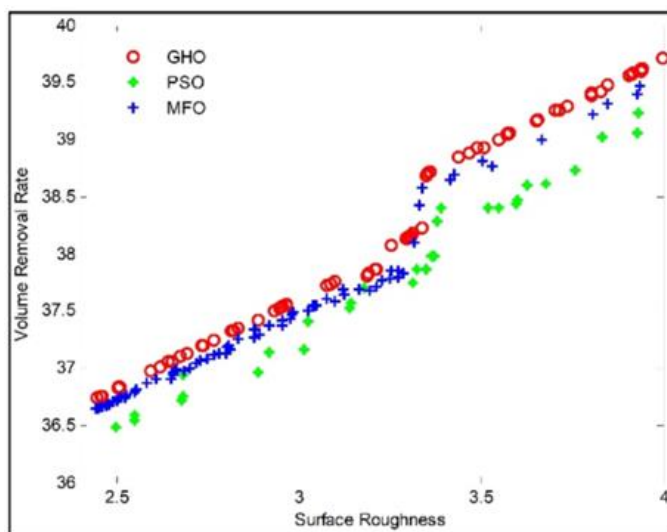
Analiza algorytmów optymalizacji obejmowała również porównanie czasu poszukiwania optymalnego rozwiązania. Wyniki przedstawiono na Rys. 4.13.



Rys. 4.13 Porównanie średniego czasu poszukiwania rozwiązania optymalnego przez algorytmy rojowe

Autorzy pracy wnioskują, że spośród sześciu rozważanych algorytmów, algorytm MOGOA wyróżnia się jako najlepszy pod względem zakresu generowanych frontów Pareto. Z kolei algorytm MOALO okazał się najszybszy pod względem czasu obliczeń [53].

Podobne zadanie badawcze zrealizowano w pracy [55], w której zastosowano algorytmy takie jak Grasshopper Optimization Algorithm (GHO), Particle Swarm Optimization (PSO) oraz Moth-Flame Optimization (MFO), przy uwzględnieniu analogicznych kryteriów optymalizacji. Uzyskane fronty Pareto przedstawiono na Rys. 4.14. Na podstawie analizy wyników stwierdzono, że algorytm GHO przewyższa pozostałe pod względem wydajności.



Rys. 4.14 Porównanie frontów Pareto dla SR vs. MRR algorytmów optymalizacji GHO, PSO i MFO [55]

W pracy przeglądowej [56], autorzy analizują różne algorytmy optymalizacji stosowane w procesach EDM i WEDM, w tym techniki inteligencji roju, takie jak Ant Colony Optimization (ACO), Particle Swarm Optimization (PSO), Genetic Algorithm (GA), Artificial Bee Colony (ABC), Cuckoo Search Algorithm (CSA), oraz różne techniki hybrydowe. W szczególności, w kontekście wielokryterialnej optymalizacji zwrócono uwagę na potencjał algorytmów takich jak GHO (Grasshopper Optimization Algorithm) oraz MFO (Moth-Flame Optimization), które charakteryzują się dużą efektywnością przy jednoczesnym uwzględnieniu kilku sprzecznych kryteriów optymalizacji, takich jak chropowatość powierzchni (SR) oraz wskaźnik usuwania materiału (MRR). Wnioski te potwierdzają wcześniej przytoczone wyniki badań [55].

Autorzy w pracy [56] podkreślają, że algorytm PSO jest jednym z najczęściej stosowanych w WEDM ze względu na prostą implementację oraz wysoką zbieżność wyników. Niemniej jednak, porównanie efektywności PSO z algorytmami GHO i MFO wykazało, że GHO przewyższa PSO pod względem jakości generowanych rozwiązań, co znajduje odzwierciedlenie w bardziej równomiernym pokryciu i lepszej dystrybucji frontów Pareto. Ponadto, analiza czasów obliczeń wykazała, że algorytm MOALO (Multi-Objective Ant Lion Optimization) charakteryzuje się najkrótszym czasem obliczeń spośród analizowanych technik, co potwierdzają wcześniejsze badania [53].

Wyniki tych badań dotyczą jednak przypadków z relatywnie niewielką liczbą zmiennych. W sytuacjach, gdzie liczba zmiennych niezależnych jest większa, jak wykazano w pracach [57][58], algorytmy genetyczne wykazują lepszą efektywność. Algorytmy takie jak PSO i GHO mogą być skuteczne w przypadku mniejszej liczby zmiennych, jednak ich efektywność spada wraz ze wzrostem złożoności problemu, co wymaga większej liczby iteracji i dokładniejszej eksploracji przestrzeni rozwiązań.

4.3. Podsumowanie

Z przeprowadzonej analizy literatury wynika, że problematyka związana z procesem cięcia elektroerozyjnego (WEDM) pozostaje aktualna i wciąż wymaga pogłębionych badań. Na podstawie przeglądu dostępnych publikacji można zidentyfikować kilka istotnych luk badawczych:

- Wiele badań opiera się na analizie niewielkiej liczby obserwacji, co znacząco utrudnia uogólnienie uzyskanych wyników na szerszą skalę. Niewystarczająca liczba próbek badawczych może prowadzić do błędnych wniosków lub zbyt dużej niepewności w przewidywaniach. Konieczne jest przeprowadzenie dalszych badań z wykorzystaniem większej liczby obserwacji, co pozwoli

zweryfikować dotychczasowe wyniki i zwiększyć reprezentatywność uzyskiwanych danych.

- Prace badawcze nad procesem WEDM koncentrują się na analizie wybranych zmiennych niezależnych, takich jak natężenie prądu, czas trwania impulsu czy napięcie elektryczne. Choć każdy z tych czynników ma istotny wpływ na proces, brakuje badań, które uwzględniłyby jednoczesną analizę szerokiego spektrum parametrów technologicznych.
- Optymalizacja procesu WEDM wymaga jednoczesnego rozpatrywania sprzecznych kryteriów, takich jak maksymalizacja prędkości usuwania materiału (MRR) oraz minimalizacja chropowatości powierzchni. Istnieje stosunkowo niewiele badań, które skutecznie rozwiązują problemy wielokryterialne.

Podsumowując, dalsze badania nad procesem WEDM powinny dążyć do zwiększenia reprezentatywności wyników poprzez większą liczbę obserwacji, rozszerzenie analizowanych zmiennych oraz wykorzystanie bardziej zaawansowanych metod modelowania i optymalizacji. Zidentyfikowane w przeglądzie literatury luki wpłynęły na określenie celów niniejszej pracy i jej zakresu. W związku z tym badania zaplanowano tak, aby objęły szeroki zakres parametrów technologicznych oraz różnorodne metody modelowania i optymalizacji, testowane w celu wyłonienia najskuteczniejszych rozwiązań.

5. PRZEGLĄD ZASTOSOWANYCH METOD I ALGORYTMÓW

W niniejszym rozdziale przedstawiono teoretyczne podstawy metod i algorytmów, które zostały wykorzystane w pracy. Punkty 5.1 i 5.2 prezentują odpowiednio metody statystyczne oraz metody uczenia maszynowego zastosowane do konstruowania modeli procesu WEDM. Wyjątek stanowi pierwszy punkt rozdziału, który wprowadza w zagadnienia analizy korelacji badanych zmiennych procesu. Wybór poszczególnych metod statystycznych był podyktowany charakterystyką zgromadzonego zbioru danych, a także wnioskami z analizy literatury.

Ze względu na fakt, że przeprowadzone badania modelowe obejmowały zmienne zależne ciągle oraz nominalne, zgodnie z [59] wyróżniono zagadnienia regresji dla zmiennych ciągłych oraz klasyfikacji dla zmiennych nominalnych. W kontekście regresji, spośród metod statystycznych, omówiono regresję liniową i wielomianową, natomiast dla klasyfikacji opisano regresję logistyczną.

W przypadku metod uczenia maszynowego, takich jak drzewa decyzyjne, las losowy, wzmocnienie gradientowe, maszyny wektorów nośnych oraz sieci neuronowe, przedstawiono odpowiednie warianty każdej z metod zarówno dla przypadków regresji, jak i klasyfikacji.

Na zakończenie rozdziału (p.5.3) przedstawiono algorytmy wykorzystane do optymalizacji procesu, takie jak algorytm genetyczny, ewolucji różnicowej, mrówkowy i optymalizacji rojem cząsteczek.

Szeroki wybór metod zastosowanych w niniejszej pracy ma na celu zapewnienie kompleksowego podejścia do modelowania i optymalizacji procesu. Metody i algorytmy opisane w niniejszym rozdziale zostały zweryfikowane w rozdziałach 9 i 10. W oparciu o analizę wyników przeprowadzoną w tych rozdziałach dokonano selekcji docelowych metod modelowania i optymalizacji procesu.

5.1. Metody statystyczne

5.1.1. Analiza korelacji

W celu identyfikacji korelacji liniowych pomiędzy zmiennymi niezależnymi a ciągłymi zmiennymi zależnymi zastosowano współczynnik korelacji liniowej Pearsona (5.1) [60],[61].

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (5.1)$$

gdzie:

- x_i i y_i - wartości poszczególnych obserwacji odpowiednio zmiennej niezależnej i zmiennej zależnej,

Stosowalność tej metody jest ograniczona założeniami co do rozkładu analizowanych zmiennych. Po pierwsze, w idealnych warunkach zmienne powinny mieć rozkład normalny. Chociaż korelacja Pearsona jest dość odporna na odchylenia od normalności, szczególnie przy dużych próbach, to ekstremalne skośności mogą wpływać na wyniki. Kolejnymi założeniami jest homoskedastyczność rozkładu, brak obserwacji odstających, a także brak zależności autokorelacyjnej w danych [61].

5.1.2. Regresja liniowa

Pierwszym badanym podejściem do modelowania procesu było wykorzystanie regresji liniowej. Regresja liniowa umożliwia modelowanie relacji między zmienną zależną a jedną lub wieloma zmiennymi niezależnymi poprzez dopasowanie liniowej funkcji regresji.

W niniejszej pracy zastosowano model wielokrotnej regresji liniowej, który można przedstawić w notacji macierzowej [62]:

$$y = X\beta + \varepsilon, \quad (5.2)$$

gdzie:

- y - wektor wartości zmiennej zależnej,
- X - macierz zmiennych niezależnych,
- β - wektor współczynników regresji,
- ε - składnik losowy (czynnik stochastyczny równania).

Aby opracować odpowiedni model należy dopasować funkcję regresji (5.2) do empirycznie uzyskanego zbioru danych. Im lepsze dopasowanie modelu, tym mniejsze różnice pomiędzy przewidywanymi wartościami \hat{y} , a rzeczywistymi wartościami y zmiennej zależnej. W regresji wielokrotnej dopasowanie to można uzyskać poprzez estymację współczynników regresji. Estymacja ta polega na minimalizacji funkcji błędu modelu (5.3) [62].

$$\min \sum_i \text{dist}(y_i, \hat{y}_i). \quad (5.3)$$

Standardową metodą minimalizacji powyższej funkcji jest metoda najmniejszych kwadratów (MNK). Metoda ta polega na minimalizowaniu sumy kwadratów reszt szacowanego modelu $S(\beta)$ (5.4) [62].

$$\min_{\beta} S(\beta) = \min_{\beta} \sum_{i=1}^N (y_i - \hat{y}_i)^2 = \min_{\beta} \sum_{i=1}^N e_i^2, \quad (5.4)$$

gdzie β to wektor oszacowań współczynników regresji (5.5).

$$\beta = (X'X)^{-1}X'y. \quad (5.5)$$

Przy czym znalezienie wektora oszacowań na podstawie formuły (5.5) jest możliwe tylko wtedy, gdy macierz $X'X$ jest macierzą odwracalną [62].

5.1.3. Regresja wielomianowa

Rozszerzeniem klasycznej regresji liniowej jest regresja wielomianowa, która umożliwia modelowanie bardziej złożonych zależności między zmiennymi, uwzględniając zarówno nieliniowości, jak i interakcje między zmiennymi niezależnymi. W pracy zastosowano model regresji wielomianowej drugiego stopnia [63]:

$$y_i = \beta_0 + \sum_{j=1}^p \beta_j x_{ij} + \sum_{j=1}^p \beta_{jj} x_{ij}^2 + \sum_{j=1}^p \sum_{k < j}^p \beta_{jk} x_{jk} x_{ik} + \varepsilon_i, \quad (5.6)$$

gdzie:

- x_{ij} - wartość j-tej zmiennej niezależnej dla i-tej obserwacji,
- β_j - współczynnik przy j-tej zmiennej niezależnej,
- β_{jj} - współczynnik przy kwadracie j-tej zmiennej niezależnej,
- β_{jk} - współczynnik interakcji między j-tą a k-tą zmienną niezależną.

Współczynniki β są podobnie jak w regresji liniowej szacowane metodą najmniejszych kwadratów (5.4).

5.1.4. Regresja logistyczna

W przypadku nominalnych zmiennych zależnych, takich jak ciągłość obróbki czy klasa szerokości szczeliny erozyjnej, rozważany jest problemem klasyfikacji. Powszechnie

stosowaną metodą do rozwiązania tego typu problemów jest regresja logistyczna. W regresji logistycznej, zależność między nominalną zmienną zależną Y a jedną lub więcej zmiennymi niezależnymi X jest modelowana za pomocą funkcji logistycznej, która przypisuje prawdopodobieństwo przynależności do określonych klas. Model matematyczny można wyrazić jako [64]:

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n, \quad (5.7)$$

gdzie:

- p - prawdopodobieństwo wystąpienia zdarzenia, gdzie $Y = 1$,

Prawdopodobieństwo p można następnie wyrazić jako [64]:

$$p = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}}, \quad (5.8)$$

Współczynniki β są zazwyczaj estymowane za pomocą funkcji wiarygodności $L(\beta)$:

$$L(\beta) = \prod_{i=1}^n p_i^{y_i} (1 - p_i)^{1-y_i} \quad (5.9)$$

gdzie:

- p_i - estymowane prawdopodobieństwo, że $Y_i = 1$ dla danej obserwacji,
- y_i - rzeczywista klasa obserwacji i .

Estymacja wartości współczynników β , polega na maksymalizacji $L(\beta)$. W praktyce często używa się metod numerycznych, takich jak metoda gradientu prostego (ang. gradient descent).

5.2. Metody uczenia maszynowego

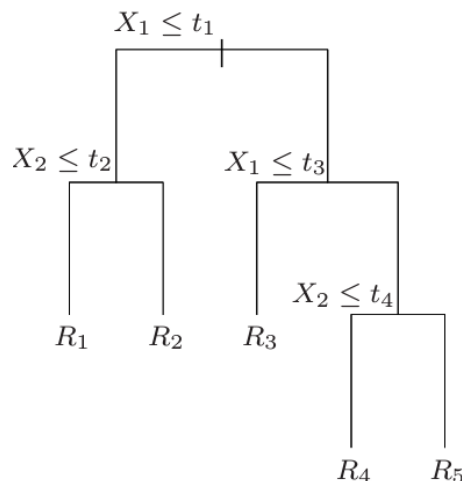
5.2.1. Drzewa decyzyjne

Z punktu widzenia modelowania procesu cięcia elektroerozyjnego istotnym zagadnieniem jest interpretacja modelu, w celu określenia wpływu poszczególnych

zmiennych niezależnych na określone zmienne zależne. Spośród metod uczenia maszynowego, metody oparte na drzewach decyzyjnych bardzo dobrze nadają się do tego zadania, ze względu na łatwo interpretowalne modele [65]. Ponadto, metody te wykazują się znaczną efektywnością w obszarze danych o wysokiej wymiarowości, a także, w przypadku gdy zależności między zmiennymi mogą mieć charakter nieliniowy [65]. Drzewa decyzyjne i ich rozwinięcia, takie jak las losowy czy metody boostingowe, są w stanie modelować złożone zależności bez bezpośredniej konieczności linearyzacji problemu.

Powyższe argumenty były decydujące dla włączenia tych metod do badań modelowych przedstawionych w niniejszej pracy. Poniżej zamieszczono skrócony opis teoretyczny rozpatrywanych metod.

Na Rys. 5.1 zaprezentowano przykładowe drzewo decyzyjne dzielące dwuwymiarową przestrzeń zmiennych niezależnych za pomocą rekurencyjnego podziału binarnego [64]. Drzewo to składa się z węzłów decyzyjnych, które reprezentują testy na zmiennych niezależnych, np. $X_n \leq t_n$, węzłów liści należących do danego regionu R_m , oraz korzenia, który jest węzłem początkowym [64].



Rys. 5.1 Przykładowe drzewo decyzyjne

W zapisie matematycznym powyższy model przyjmuje postać [64]:

$$\hat{f}(X) = \sum_{m=1}^5 c_m I\{(X_1, X_2) \in R_m\}, \quad (5.10)$$

gdzie:

- I - funkcja wskaźnikowa, która przyjmuje wartość 1, gdy para zmiennych X_1, X_2 należy do regionu R_m , i 0 w przeciwnym wypadku,
- c_m - stała modelu przewidująca wartość zmiennej zależnej w regionie R_m .

Poniżej zaprezentowano uogólnioną postać modelu (5.10):

$$\hat{f}(X) = \sum_{m=1}^M c_m I(x \in R_m). \quad (5.11)$$

Jeśli przyjąć jako kryterium minimalizację sumę kwadratów $\sum (y_i - \hat{y}_i)^2$, to wówczas można zauważyć, że najlepszą wartością stałej \hat{c}_m jest średnia wartość z y_i w regionie R_m (5.12) [64].

$$\hat{c}_m = \text{ave}(y_i | x_i \in R_m). \quad (5.12)$$

Proces budowy drzewa polega na dzieleniu zbioru danych na regiony na podstawie kryteriów podziału, aż do osiągnięcia warunku stopu [66]. Uściślając, dla każdej zmiennej w zbiorze danych, algorytm szuka takiego punktu podziału, który minimalizuje błąd średniokwadratowy w wynikających z tego podziału regionach. Dla każdej zmiennej j oraz punktu podziału binarnego s , można zdefiniować dwa nowe regiony [64]:

$$R_1(j, s) = \{X | X_j \leq s\}, \quad R_2(j, s) = \{X | X_j > s\} \quad (5.13)$$

Optymalne punkty podziału otrzymuje się zatem przy użyciu funkcji minimalizacji [64]:

$$\min_{j,s} \left[\min_{c_1} \sum_{x_i \in R_1(j,s)} (y_i - c_1)^2 + \min_{c_2} \sum_{x_i \in R_2(j,s)} (y_i - c_2)^2 \right]. \quad (5.14)$$

Proces ten jest powtarzany rekurencyjnie dla każdego nowo utworzonego regionu, aż do spełnienia jednego z kryteriów zatrzymania, takich jak maksymalna głębokość drzewa, minimalna liczba obserwacji w regionie lub gdy dalszy podział nie przynosi znaczącej poprawy błędu średniokwadratowego.

W przypadku problemów klasyfikacji jedyne zmiany w algorytmie drzewa dotyczą kryteriów podziału węzłów i przycinania drzewa [64]. Zamiast używania błędu średniokwadratowego, stosuje się wówczas np. współczynnik Giniego [64]:

$$I_G(p) = 1 - \sum_{i=1}^j p_i^2, \quad (5.15)$$

gdzie:

- $I_G(p)$ - współczynnik Giniego,
- p_i^2 - prawdopodobieństwo wystąpienia klasy i w danym węźle,
- j - liczba klas.

Tak utworzone drzewo decyzyjne wykazuje również pewne niedoskonałości. Drzewa decyzyjne, są podatne na nadmierne dopasowanie do danych treningowych (ang. overfitting), co ogranicza ich zdolność do generalizacji. Aby zapobiegać overfittingowi, stosuje się techniki takie jak przycinanie drzewa, ograniczanie głębokości drzewa oraz ustalanie minimalnej liczby próbek dla węzłów i liści [64]. Wiąże się to jednak z koniecznością poszukiwania odpowiednich wartości tych hiperparametrów oraz nie zawsze jest wystarczająco skuteczne. Ponadto, metody te są również ograniczone w modelowaniu złożonych relacji nieliniarnych i wrażliwe na zaszumione oraz niezrównoważone dane [65]. W związku z tym, w niniejszej pracy wykorzystano również rozszerzenia tej metody, m.in. metodę lasu losowego (ang. Random Forest).

5.2.2. Las losowy

Zasadę działania lasu losowego można scharakteryzować jako proces agregacji zbioru drzew decyzyjnych generowanych metodą losową. W odróżnieniu od pojedynczego drzewa decyzyjnego, gdzie nacisk kładzie się na optymalizację jednego drzewa, las losowy dąży do połączenia wielu drzew decyzyjnych, które nie muszą być indywidualnie optymalne. Kluczowym aspektem tej metody jest wprowadzenie losowości przy budowie poszczególnych drzew, co umożliwia szeroką eksplorację przestrzeni wszystkich potencjalnych drzew decyzyjnych. Rezultatem tego procesu jest poprawa dokładności przewidywania, wynikająca z dywersyfikacji błędów poszczególnych drzew, co przekłada się na zwiększoną zdolność do generalizacji, poprzez ograniczenie wpływu możliwego overfittingu [67].

Aby uzyskać wysoką dokładność przewidywania za pomocą metody lasu losowego, należy zbudować zbiór drzew decyzyjnych, które spełniają następujące kryteria [67]:

- Są tak różnorodne, jak to możliwe, ponieważ agregacja zestawu drzew, które są bardzo podobne, nie przyniosłaby korzyści w porównaniu do pojedynczego drzewa decyzyjnego.
- Składają się z indywidualnych drzew o akceptowalnej zdolności przewidywania, ponieważ jeśli wszystkie drzewa dostarczają nieprawidłowych przewidywań dla nowej obserwacji x , agregacja tych przewidywań nie poprawi ogólnej dokładności.

Budowa lasu losowego odbywa się w następujących krokach [64]:

1. Próbkowanie bootstrapowe: dla każdego drzewa w lesie, tworzona jest próbka bootstrapowa Z danych treningowych o rozmiarze N . Proces ten polega na losowym wybieraniu obserwacji z pełnego zbioru danych z powtórzeniami, co prowadzi do utworzenia próbki, w której niektóre obserwacje mogą pojawić się wielokrotnie, a inne mogą zostać pominięte.
2. Budowa drzewa decyzyjnego: na podstawie próbki bootstrapowej budowane jest drzewo decyzyjne. W przeciwieństwie do standardowych drzew decyzyjnych, w lesie losowym, na każdym etapie podziału węzła, wybór najlepszego podziału jest dokonywany spośród losowo wybranej podgrupy zmiennych niezależnych, a nie z całego zestawu zmiennych. Liczba zmiennych wybieranych w każdym węźle jest parametrem modelu i jest zwykle oznaczana jako \sqrt{p} gdzie p to liczba całkowita.
3. Budowa wielu drzew: poprzedni proces budowy drzewa jest powtarzany M razy, aby utworzyć las składający się z M drzew decyzyjnych.
4. W przypadku regresji, przewidywania wszystkich drzew są uśredniane lub w przypadku klasyfikacji, występuje głosowanie większościowe.

Korzystając z wcześniej przedstawionego modelu drzewa decyzyjnego (5.11) w przypadku regresji model lasu losowego można przedstawić za pomocą formuły [67]:

$$f_{RF}(x) = \frac{1}{M} \sum_{m=1}^M \hat{f}_m(x), \quad (5.16)$$

gdzie:

- $f_{RF}(x)$ – przewidywane wartości przez las losowy dla obserwacji x ,
- M – liczba drzew decyzyjnych,
- $\hat{f}_m(x)$ – przewidywane wartości dla m -tego drzewa.

W przypadku klasyfikacji formułę modelu można wyrazić następująco [64]:

$$C_{RF}(x) = \text{moda}\{\hat{C}_m(x)\}^M, \quad (5.17)$$

gdzie:

- $C_{RF}(x)$ – przewidywana klasa przez las losowy dla obserwacji x ,
- $\hat{C}_m(x)$ – przewidywana klasa dla m -tego drzewa.

Metoda lasu losowego przewyższa pojedyncze drzewa decyzyjne, zmniejszając overfitting przez uśrednianie wielu drzew [68], lepiej radząc sobie z nieliniowościami i zaszumionymi danymi. Skuteczniej obsługuje nieźrównoważone dane i zmniejsza wrażliwość na zmiany w danych treningowych. Dodatkowo, umożliwia ocenę ważności zmiennych niezależnych poprzez analizę zmiany w błędzie przewidywania, gdy wartości danej zmiennej są permutowane między obserwacjami [64].

5.2.3. Wzmacnianie gradientowe

Alternatywą dla lasu losowego metodą opartą na drzewach decyzyjnych są m.in. metody boostingowe. W odróżnieniu od lasu losowego, w którym wiele drzew jest budowanych niezależnie i ich przewidywania są łączone przez głosowanie lub uśrednianie, w metodach boostingowych drzewa budowane są sekwencyjnie. Każde kolejne drzewo poprawia błędy popełnione przez poprzednie, co prowadzi do stopniowej poprawy jakości modelu.

Ogólną postać algorytmu w metodzie boostingowej można wyrazić w następujący sposób [64]:

$$f_m = \arg \min_f \sum_{i=1}^N L(y_i, F_{m-1}(x_i) + f(x_i)) \quad (5.18)$$

Każde drzewo f_m jest dopasowane tak, aby minimalizować funkcję straty L względem aktualnych reszt, które są różnicami między rzeczywistymi wartościami y_i a przewidywaniami drzewa na bieżącym etapie $F_{m-1}(x_i)$.

W zależności od wybranego algorytmu metody boostingowej funkcja straty L może przyjmować różną postać. W niniejszej pracy wykorzystano algorytm XGBoost (eXtreme Gradient Boosting), który zyskał uznanie w dziedzinie uczenia maszynowego za swoją wydajność i skuteczność [69].

Funkcja straty w algorytmie XGBoost przyjmuje postać [69]:

$$L^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t), \quad (5.19)$$

gdzie:

- $\hat{y}_i^{(t-1)}$ - wartość przewidywana po $t - 1$ iteracjach,
- $\Omega(f_t)$ funkcja regularyzująca, zdefiniowana jako [69]:

$$\Omega(f_t) = \gamma Tl + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2, \quad (5.20)$$

gdzie:

- γ, λ - parametry regularyzacji,
- Tl - ilość liści w drzewie decyzyjnym,
- ω_j - waga przypisana do j - tego liścia.

Regularyzacja składa się z dwóch części: pierwszej, która karze za liczbę liści w drzewie oraz drugiej, która karze za wagi przypisane liściom. Obie te części pomagają kontrolować złożoność modelu i zapobiegać overfittingowi.

Podsumowując, XGBoost stanowi ewolucję metod boostingowych, dostarczając narzędzie, które jest skuteczne w redukowaniu błędu przewidywania, wydajne z punktu widzenia obliczeniowego i skuteczne w zapobieganiu overfittingowi poprzez zastosowanie regularyzacji [69].

5.2.4. Maszyny wektorów nośnych

Innym podejściem, różniącym się od dotychczas rozważanych metod drzew decyzyjnych, które skutecznie radzi sobie z potrzebą efektywnego modelowania złożonych, nieliniowych zależności między zmiennymi, jest metoda maszyn wektorów nośnych (ang. Support Vector Machines, SVM).

SVM to zaawansowana metoda uczenia maszynowego, która pierwotnie została opracowana do rozwiązywania problemów klasyfikacji. Kluczowym aspektem SVM jest znajdowanie hiperpłaszczyzny, która najlepiej oddziela dane, maksymalizując margines między różnymi klasami. Margines jest definiowany jako odległość między hiperpłaszczyzną a najbliższymi próbkami z każdej klasy, zwanymi wektorami nośnymi.

Poszukiwanie odpowiedniej hiperpłaszczyzny w kontekście klasyfikacji binarnej można przedstawić jako problem optymalizacji [64],[70]:

$$\min_{\omega, b} \frac{1}{2} \|\omega\|^2, \quad (5.21)$$

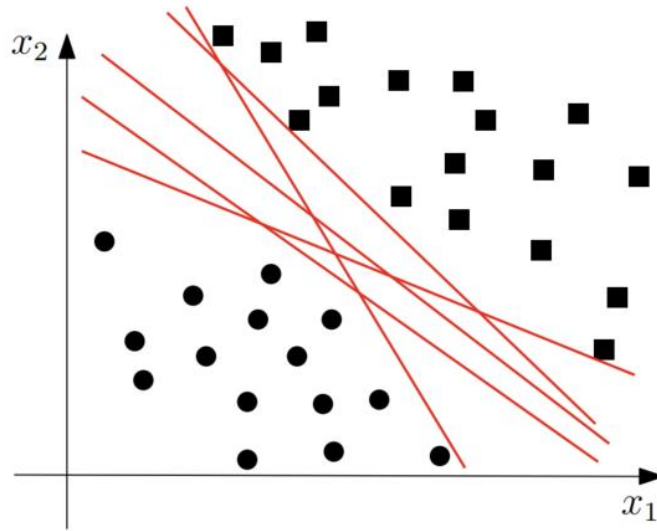
pod warunkiem, że:

$$y_i(\omega \cdot x_i + b) \geq 1. \quad (5.22)$$

gdzie:

- x_i - wektor zmiennych niezależnych i -tej obserwacji
- y_i - etykieta klasy (przyjmuje wartości -1 lub 1),
- ω i b - parametry hiperpłaszczyzny, która separuje klasy z maksymalnym marginesem

Problem dopasowania hiperpłaszczyzny w klasyfikacji binarnej przedstawiono na rysunku poniżej (Rys. 5.2).



Rys. 5.2 Różne hiperpłaszczyzny rozdzielające dla tego samego zbioru danych w problemie klasyfikacji binarnej [70]

W kontekście problemów regresji, wariant tego algorytmu znany jest jako Support Vector Regression (SVR). Zamiast dążyć do maksymalizacji marginesu między klasami, SVR stara się znaleźć hiperpłaszczyznę, która jak najlepiej odzwierciedla rozpatrywane dane, pozwalając jednocześnie na pewne odstępstwa od dokładnych wartości w ramach określonego marginesu błędu [64].

Formuła problemu poszukiwania hiperpłaszczyzny w SVR wygląda następująco [70]:

$$\min_{\omega, b} \frac{1}{2} \|\omega\|^2 + Cr \sum (\xi_i + \xi_i^*), \quad (5.23)$$

pod warunkiem, że:

$$\begin{aligned} y_i - (\omega \cdot x_i + b) &\leq \epsilon + \xi_i, \\ (\omega \cdot x_i + b) - y_i &\leq \epsilon + \xi_i^*, \\ \xi_i^*, \xi_i &\geq 0, \end{aligned} \quad (5.24)$$

gdzie:

- ξ_i^* i ξ_i - zmienne luzu, które pozwalają na przekroczenie marginesu błędu ϵ dla danych punktów,
- Cr - parametr regularyzacji, który kontroluje kompromis między płaskością funkcji a tolerancją na odstępstwa.

Rozwiązanie tego problemu prowadzi do modelu, który przewiduje wartość y z dopuszczalnym błędem ϵ , jednocześnie minimalizując złożoność modelu.

Poza poszukiwaniem odpowiedniej hiperpłaszczyzny, w ramach metod SVM i SVR stosuje się dodatkowo przekształcenie przestrzeni danych wejściowych do przestrzeni o wyższej wymiarowości, w której dane mogą być łatwiej separowane przez hiperpłaszczyznę. Jest to kluczowe dla problemów, gdzie dane w oryginalnej przestrzeni nie są liniowo separowalne. W tym celu stosuje się odpowiednie funkcje jądrowe, których przykłady wymieniono niżej [70]:

- Liniowa

$$K(x, x') = x \cdot x', \quad (5.25)$$

- Wielomianowa

$$K(x, x') = (1 + x \cdot x')^d, \quad (5.26)$$

gdzie d jest stopniem wielomianu,

- Radialna

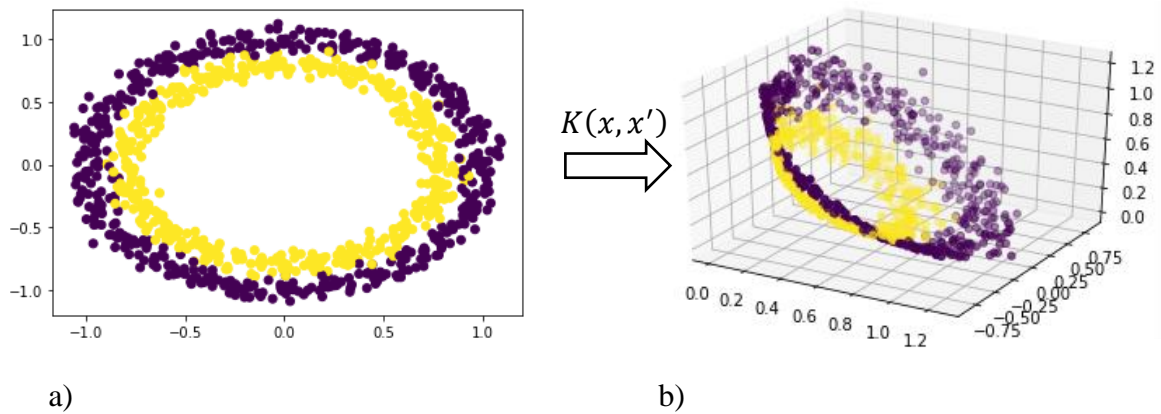
$$K(x, x') = \exp(-\gamma \|x - x'\|^2), \quad (5.27)$$

gdzie γ jest parametrem, który kontroluje zasięg wpływu pojedynczej próbki treningowej.

- Sigmoidalna

$$K(x, x') = \tanh(\gamma x \cdot x'), \quad (5.28)$$

Na Rys. 5.3 przedstawiono wizualizację przykładowej transformacji danych z użyciem radialnej funkcji jądrowej. Pierwotna przestrzeń danych wejściowych Rys. 5.3a jest trudno separowalną dwuwymiarową przestrzenią, natomiast po transformacji przestrzeń na Rys. 5.3b jest wyraźnie łatwiejsza do podziału.



Rys. 5.3 Przykładowa transformacja dwuwymiarowej przestrzeni danych (a) do trójwymiarowej przestrzeni danych (b) z użyciem radialnej funkcji jądrowej

Podsumowując, metoda ta jest szczególnie ceniona za zdolność do efektywnego radzenia sobie z wysokowymiarowymi przestrzeniami danych oraz za elastyczność w modelowaniu złożonych granic decyzyjnych dzięki zastosowaniu różnych funkcji jądrowych [64].

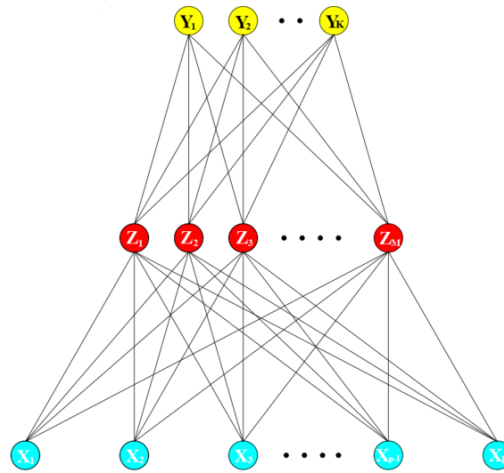
5.2.5. Sztuczne sieci neuronowe

Rozważając algorytmy uczenia maszynowego w kontekście modelowania procesu obróbki elektroerozyjnej, nie sposób pominąć podejście oparte na sztucznych sieciach neuronowych (ang. Artificial Neural Networks, ANN). Sztuczne sieci neuronowe oferują jeszcze większą elastyczność i zdolność do modelowania nieliniowych zależności w danych [68], co czyni je szczególnie przydatnymi w modelowaniu zjawisk, charakteryzujących się wysoką stochastycznością, takich jak obróbka WEDM.

ANN są inspirowane strukturą i funkcjonowaniem ludzkiego mózgu. Składają się w większości przypadków z warstw neuronów, gdzie każdy neuron w jednej warstwie jest połączony z neuronami w następnej warstwie za pomocą sieci komunikacyjnej. Każde połączenie pomiędzy neuronami ma przypisaną wagę, która determinuje siłę i znaczenie przekazywanego sygnału [68].

W niniejszej pracy zastosowano powszechnie stosowaną sieć neuronową typu feed-forward. W odróżnieniu od innych typów sieci neuronowych, takich jak sieci rekurencyjne, sieci feed-forward nie posiadają połączeń zwrotnych. Dane przepływają przez sieć od wejścia do wyjścia [71].

Model sieci neuronowej, zwykle jest reprezentowany przez diagram sieciowy:



Rys. 5.4 Schemat sieci neuronowej typu feed-forward z jedną warstwą ukrytą [24]

Jak przedstawiono na Rys. 5.4. podstawowa architektura sieci neuronowej składa się z następujących elementów [64]:

- Warstwy wejściowej (X), składającej się z neuronów wejściowych X_1, X_2, \dots, X_n , które reprezentują zmienne niezależne
- Warstwy ukrytej (Z), składającej się z neuronów ukrytych Z_1, Z_2, \dots, Z_m , które przetwarzają dane wejściowe. Neurony te wykorzystują ważone sumy danych wejściowych oraz funkcje aktywacji do generowania danych wyjściowych, które będą dalej przekazywane. Liczba warstw ukrytych i liczba neuronów w każdej warstwie może być dostosowana w celu zwiększenia zdolności sieci do modelowania bardziej złożonych relacji.
- Warstwy wyjściowej (Y), składającej się z neuronów wyjściowych Y_1, Y_2, \dots, Y_k , które generują ostateczną odpowiedź sieci (dane wyjściowe).

Każdy neuron opisywany jest funkcją liniową z wykorzystaniem wag (ω) i wartości progowych (biasów, b) oraz funkcją aktywacji (f). Wyjście z i -tego neuronu w warstwie ukrytej można wyrazić jako [64]:

$$Z_i = f \left(\sum_{j=1}^n \omega_{ij} X_j + b_i \right), \quad (5.29)$$

gdzie:

- ω_{ij} - waga połączenia między j -tym neuronem wejściowym a i -tym neuronem ukrytym,
- b_i - bias i -tego neuronu ukrytego.

Następnie, wyjście z warstwy ukrytej jest przekazywane do warstwy wyjściowej, gdzie jest przetwarzane w podobny sposób [64],[71]:

$$Y_k = f \left(\sum_{j=1}^m \omega'_{ij} Z_i + b'_k \right), \quad (5.30)$$

gdzie:

- ω'_{ij} - waga połączenia między i -tym neuronem ukrytym a k -tym neuronem wyjściowym,
- b'_k - bias k -tego neuronu wyjściowego.

Proces uczenia sieci neuronowej polega na dostosowywaniu wag i biasów w taki sposób, aby funkcja straty, mierząca różnicę między danymi wyjściowymi sieci a danymi wzorcowymi, była minimalizowana. Wykorzystuje się do tego algorytm propagacji wstecznej, zwykle w połączeniu z metodą prostego spadku gradientu (SGD) [71].

5.3. Metody optymalizacji

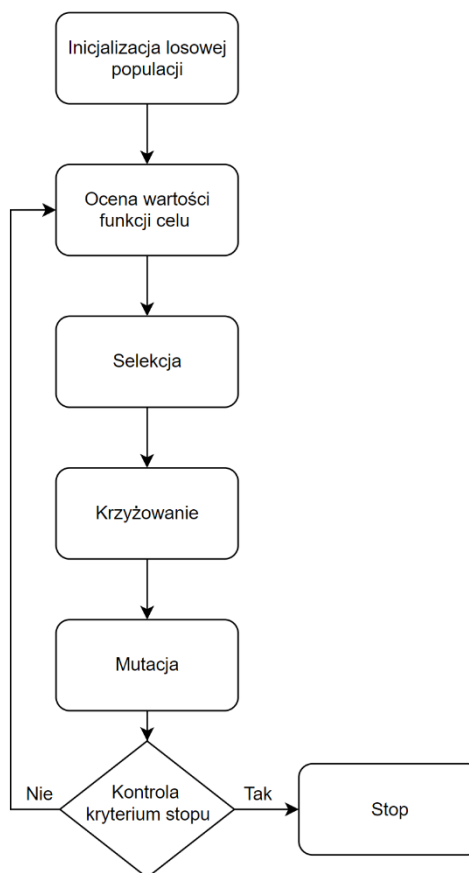
W procesie optymalizacji parametrów WEDM zdecydowano się na zastosowanie algorytmów optymalizacyjnych inspirowanych biologicznie. Wybór ten był podyktowany szeregiem korzyści jakie metody tego typu zapewniają w kontekście niniejszej pracy. Po pierwsze, algorytmy inspirowane biologicznie wykazują wysoką efektywność w znajdowaniu rozwiązań optymalnych lub suboptymalnych w złożonych (wielomodalnych) przestrzeniach wyszukiwania [72]. Ze względu na złożoność procesu WEDM oraz szeroki zakres rozpatrywanych parametrów technologicznych, zdolność algorytmów do efektywnego przeszukiwania rozległego zbioru możliwych rozwiązań jest istotną zaletą. Kolejnym powodem jest wdrożeniowy charakter pracy. Opracowane rozwiązania będą na stałe zintegrowane z systemami przedsiębiorstwa, co z kolei podkreśla potrzebę adaptacyjności w dynamicznym środowisku produkcyjnym. Rozważane algorytmy odpowiadają na to zapotrzebowanie charakteryzując się wysoką elastycznością i możliwością stosunkowo łatwego przystosowania do nowych danych. Podsumowując, decyzja o skoncentrowaniu badań na algorytmach inspirowanych biologicznie jest uzasadniona zarówno z perspektywy teoretycznej, jak i praktycznej.

W związku z powyższym, w tym rozdziale przeprowadzono syntetyczny przegląd teoretyczny zastosowanych w pracy algorytmów. Wyniki badań porównawczych zaprezentowano w rozdziale 10.

5.3.1. Algorytm genetyczny

Najlepszym procesem optymalnego zachowania obserwowanym w naturze jest ewolucja gatunków na przestrzeni ich pokoleń [73]. Każde nowe pokolenie dowolnego gatunku będzie zmieniać się w stosunku do poprzedniego i będzie starało się jak najlepiej dostosować do zmieniającego się środowiska. Cechy pojawiające się w nowym pokoleniu sprawiają, że będzie ono bardziej zdolne do przetrwania.

Algorytm genetyczny (ang. Genetic Algorithm, GA), próbuje zaimplementować ten naturalny proces w celu znalezienia optymalnego rozwiązania różnych problemów optymalizacyjnych, poprzez symulację procesów takich jak selekcja, krzyżowanie (rekombinacja) i mutacja osobników w danej populacji. GA inicjuje proces z losowo wygenerowaną populacją osobników, które są ewaluowane za pomocą funkcji przystosowania. Najlepiej przystosowane osobniki mają większe szanse na przekazanie swoich cech do kolejnych pokoleń poprzez krzyżowanie i mutacje, co z czasem prowadzi do akumulacji korzystnych cech i zwiększenia ogólnego przystosowania populacji. Proces ten jest realizowany iteracyjnie, aż do spełnienia określonego kryterium zatrzymania algorytmu, na przykład osiągnięcie założonej liczby generacji lub znalezienie rozwiązania o zadawalającym poziomie przystosowania [73]. Mechanizm ten został zaprezentowany w postaci schematu blokowego na Rys. 5.5.



Rys. 5.5 Schemat blokowy algorytmu genetycznego [30]

Jak można zauważyć, w opisie algorytmu genetycznego występuje specyficzna terminologia. Podstawowe pojęcia występujące w kontekście tego algorytmu to [72],[73]:

- Chromosom – reprezentuje jedno z możliwych rozwiązań problemu optymalizacyjnego, zawierające zakodowane wartości parametrów (geny) definiujących to rozwiązanie.
- Selekcja – mechanizm wykorzystywany do identyfikacji i wyboru chromosomów posiadających pożądane cechy genetyczne, które zostaną przekazane do kolejnego etapu reprodukcji. Selekcja jest realizowana na podstawie funkcji przystosowania, która ocenia jakość rozwiązania reprezentowanego przez chromosom.
- Krzyżowanie – proces rekombinacji genetycznej, w którym dochodzi do wymiany fragmentów chromosomów pomiędzy parami wyselekcjonowanych chromosomów rodzicielskich w celu wytworzenia nowego zestawu chromosomów potomnych. Ten mechanizm pozwala na eksplorację przestrzeni rozwiązań poprzez generowanie nowych konfiguracji genetycznych.
- Mutacja – operacja modyfikacji genów w chromosomie, mająca na celu wprowadzenie losowych zmian genetycznych, co umożliwia zwiększenie różnorodności genetycznej populacji. Mutacje są kluczowe dla zachowania zdolności adaptacyjnych systemu ewolucyjnego, umożliwiając eksplorację nowych obszarów przestrzeni rozwiązań.

5.3.2. Algorytm ewolucji różnicowej

Algorytm ewolucji różnicowej (ang. Differential Evolution, DE), podobnie do algorytmu genetycznego polega na ewoluowaniu populacji rozwiązań, przy czym każde rozwiązanie reprezentowane jest przez ciąg liczb rzeczywistych. Dystynktywną cechą ewolucji różnicowej w stosunku do tradycyjnych algorytmów ewolucyjnych jest wykorzystanie operatorów arytmetycznych do generowania nowych rozwiązań zamiast standardowych operatorów krzyżowania i mutacji. Mechanizmy mutacji w ewolucji różnicowej opierają się na operacjach dodawania i odejmowania wektorów (potencjalnych rozwiązań optymalizacji), co sprawia, że metoda ta jest stosowana wyłącznie w przestrzeniach metrycznych, takich jak przestrzenie logiczne, całkowite lub rzeczywiste [74].

Termin "różnicowy" odnosi się do procesu tworzenia różnic pomiędzy losowo wybranymi wektorami (chromosomami) z populacji, które następnie kierują procesem poszukiwania optymalnego rozwiązania. W trakcie mutacji liczba wektorów różnicowych, które są używane do modyfikacji istniejącego rozwiązania, może się różnić. Zwykle wykorzystuje się jeden ($y = 1$) lub dwa ($y = 2$) wektory różnicowe. Te wektory różnicowe są używane do tworzenia nowego zmutowanego wektora x_m , który jest określany jako "wektor dawcy" podczas krzyżowania. Wektor dawcy x_m jest następnie krzyżowany

z istniejącym wektorem docelowym, który pochodzi z bieżącej populacji, w celu utworzenia nowego potencjalnie lepszego rozwiązania w kolejnej generacji populacji [75].

Ramy operacyjne klasycznych wariantów algorytmu (DE) mogą być podsumowane w następujących krokach [75]:

1. Inicjalizacja parametrów kontrolnych oraz losowe wygenerowanie początkowej populacji N osobników w n -wymiarowej przestrzeni poszukiwań D , przy czym dla każdego osobnika wyznaczana jest jego wartość funkcji przystosowania (fitness).
2. Sekwencyjne wybieranie osobników, od pierwszego do N -tego, które będą pełniły rolę wektorów docelowych x_t . Cykl generacyjny uznaje się za zakończony, gdy każdy z N osobników przejdzie poniższe etapy:
 - a. Wektor docelowy x_t jest poddawany mutacji w celu wygenerowania powiązanego z nim wektora zmutowanego x_m .
 - b. x_t i x_m są poddawane rekombinacji (krzyżowaniu) w celu wygenerowania osobników potomnych (potomków) x_0 .
 - c. Ocena wartości funkcji przystosowania potomstwa x_0 jest porównywana z wartością funkcji przystosowania wektora docelowego x_t . W zależności od wyniku, lepszy z tych dwóch wektorów zostaje wybrany i przeniesiony do tymczasowej populacji.
3. Po wypełnieniu tymczasowej populacji N wybranymi osobnikami, zastępuje ona bieżącą populację. Następnie inicjowany jest kolejny cykl generacyjny, zaczynający się od kroku 2.

Proces ten jest kontynuowany w sposób iteracyjny, podobnie jak w przypadku algorytmu genetycznego.

5.3.3. Algorytm optymalizacji rojem cząstek

Algorytm optymalizacji rojem cząstek (ang. particle swarm optimization, PSO) to metaheurystyczny algorytm optymalizacyjny inspirowany pewnymi aspektami zachowania organizmów żyjących w grupach, takich jak ławice ryb czy stada ptaków. Został on zaproponowany przez Kennedy'ego i Eberharta w 1995 roku jako technika optymalizacji globalnej [72]. W algorytmie PSO, potencjalne rozwiązania, nazywane "cząstkami", poruszają się w przestrzeni poszukiwań, która reprezentuje wszystkie możliwe rozwiązania problemu. Każda cząstka posiada własną pozycję, którą można uznać za kandydujące rozwiązanie problemu, oraz prędkość, która określa, jak szybko cząstka przemieszcza się w tej przestrzeni.

Pozycja i -tej cząstki w d -wymiarowej przestrzeni poszukiwań jest reprezentowana jako wektor $x_i = (x_{i1}, \dots, x_{id})$, a prędkość jako $v_i = (v_{i1}, \dots, v_{id})$.

Na podstawie informacji o własnych doświadczeniach oraz doświadczeniach sąsiadujących cząstek, każda cząstka aktualizuje swoją prędkość i pozycję według poniższych równań [72]:

- Aktualizacja prędkości:

$$v_i(t + 1) = \omega v_i(t) + c_1 r_1 (p_i - x_i(t)) + c_2 r_2 (g - x_i(t)), \quad (5.31)$$

- Aktualizacja pozycji:

$$x_i(t + 1) = x_i(t) + v_i(t + 1), \quad (5.32)$$

gdzie:

- t - iteracja algorytmu,
- ω - współczynnik bezwładności, który wpływa na prędkość cząstki,
- c_1 i c_2 - współczynniki uczenia osobistego i społecznego, odpowiednio,
- r_1 i r_2 - losowe liczby z zakresu $[0, 1]$,
- p_i - najlepsza dotychczas znaleziona pozycja przez i -tą cząstkę (tzw. najlepsze osobiste),
- g - najlepsza pozycja znaleziona przez sąsiedztwo cząstki lub przez całą populację (tzw. najlepsze globalne).

W trakcie działania algorytmu, wszystkie cząstki są inicjalizowane z losowymi pozycjami i prędkościami, a następnie iteracyjnie poprawiają swoje rozwiązania. Cząstki są "przyciągane" zarówno do swoich własnych najlepszych pozycji, jak i do najlepszej pozycji znalezionej przez całą populację [72].

5.3.4. Algorytm mrówkowy

Algorytm mrówkowy, znany również jako Ant Colony Optimization (ACO), jest metaheurystyczną techniką optymalizacyjną inspirowaną zachowaniem mrówek w poszukiwaniu pokarmu. Algorytm ten został wprowadzony przez Marco Dorigo w latach 90. XX wieku jako podejście do rozwiązywania problemów optymalizacyjnych, szczególnie tych związanych z wyznaczaniem tras, takich jak problem komiwojażera [76].

ACO symuluje zachowanie mrówek podczas poszukiwania najkrótszej ścieżki od ich gniazda do źródła pokarmu. Mrówki eksplorują przestrzeń poszukiwań i zostawiają ślad feromonowy, który pomaga im w nawigacji i komunikacji z innymi mrówkami. Im silniejszy ślad feromonowy, tym większe prawdopodobieństwo, że kolejne mrówki pójdą tą samą trasą [77].

Kiedy mrówki przemierzają ścieżki, uaktualniają ilość feromonów zgodnie z formułą [76]:

$$\tau_i(t + 1) = (1 - p)\tau_{xy}(t) + \Delta\tau_{xy}, \quad (5.33)$$

gdzie:

- $\tau_{xy}(t)$ - ilość feromonu na ścieżce od punktu x do y w chwili t ,
- p - współczynnik parowania feromonu,
- $\Delta\tau_{xy}$ - ilość feromonu dołożonego przez mrówki, która zależy od jakości znalezionej ścieżki.

Prawdopodobieństwo, że mrówka wybierze przejście z punktu i do j , jest zazwyczaj wyrażane jako [76],[77]:

$$p_{ij}(t) = \frac{[\tau_{ij}(t)]^\alpha [\eta_{ij}]^\beta}{\sum_{k \in A_i} [\tau_{ik}(t)]^\alpha [\eta_{ik}]^\beta}, \quad (5.34)$$

gdzie:

- α i β - parametry kontrolujące wpływ feromonu τ i widoczności η odpowiednio,
- η_{ij} to atrybut heurystyczny, np. odwrotność odległości, sugerujący preferencję dla krótszych ścieżek,
- A_i to zbiór dopuszczalnych ruchów z pozycji i .

W algorytmie ACO, wiele mrówek jest rozmieszczanych losowo na różnych ścieżkach. Po każdej iteracji, ścieżki te są oceniane, a ślady feromonowe są aktualizowane, co ma odzwierciedlać jakość znalezionych rozwiązań. Wraz z upływem czasu, mrówki mają tendencję do wybierania ścieżek z silniejszym śladem feromonowym, co prowadzi do zwiększenia prawdopodobieństwa znalezienia optymalnych rozwiązań [76].

6. ZASADY I TECHNIKI SYMULACJI KOMPUTEROWYCH

Niniejszy rozdział został poświęcony omówieniu procesów związanych z przygotowaniem danych pomiarowych do celów modelowania, przedstawieniu zastosowanych metod konstruowania modeli oraz metod oceny ich jakości.

W pierwszej części rozdziału (p. 6.1.1) omówiono różne strategie radzenia sobie z niekompletnymi danymi, w tym zarówno klasyczne metody, jak i te oparte na technikach uczenia maszynowego.

Następnie, w p. 6.1.2, przedstawiono metody transformacji danych, które miały na celu dostosowanie ich do wymogów różnych technik modelowania. Omówiono techniki normalizacji, standaryzacji oraz inne przekształcenia, niezbędne do zapewnienia optymalnych warunków dla procesu modelowania.

W przypadku problemów klasyfikacji szczególną uwagę poświęcono technikom balansowania klas (p. 6.1.3), które miały na celu zapewnienie równomiernego rozkładu klas w zbiorze danych. Przedstawiono różne metody nadpróbkowania, podpróbkowania oraz techniki hybrydowe.

Kolejna część rozdziału skoncentrowała się na procesach trenowania i ewaluacji modeli (p. 6.2). Omówiono metodę podziału danych treningowych (p. 6.2.1), walidację krzyżową a także optymalizację hiperparametrów (p.6.2.2), która pozwoliła na maksymalne wykorzystanie potencjału wybranych metod modelowania.

Ostatnia część rozdziału została poświęcona metodom oceny jakości modeli (p. 6.3), z uwzględnieniem specyfiki oceny zarówno modeli regresyjnych, jak i klasyfikacyjnych. Przedstawiono kryteria oceny, takie jak dokładność, precyzja, czułość, specyficzność, a także miary błędów dla modeli regresyjnych. Omówiono również techniki analizy błędów. Na koniec przedstawiono metodę interpretacji modeli SHAP, która pozwala na ocenę wpływu poszczególnych cech na wyniki modelu, oraz ogólny schemat badań modelowych, który podsumowuje całość procesu modelowania.

Wszystkie wyniki opisanych procesów zostały szczegółowo przedstawione i omówione w rozdziale 9

6.1. Przygotowanie danych do modelowania

Dane wejściowe do modeli zazwyczaj wymagają odpowiedniego przetworzenia, aby zagwarantować ich użyteczność w kontekście budowy modelu. W niniejszej pracy proces

przygotowanie danych obejmuje kilka etapów: uzupełnianie brakujących wartości, transformację danych oraz balansowanie klas. W poniższej sekcji opisano każdy z tych etapów.

6.1.1. Postępowanie z niekompletnymi danymi

W zgromadzonym zbiorze danych liczba dostępnych wyników pomiarowych nie była równomierna dla wszystkich zmiennych zależnych. Plan eksperymentu zakładał, że dla każdej kombinacji wartości zmiennych niezależnych zostanie wykonany odpowiedni pomiar zmiennej zależnej, jednak w praktyce nie udało się zgromadzić pełnego zestawu pomiarów dla wszystkich kombinacji. Oznacza to, że niektóre zmienne zależne mają mniej wyników pomiarowych (obserwacji) niż założono, co określane jest jako niekompletność danych.

Przyczyny tych niekompletności zostały omówione w rozdziale 8.1. Brakujące dane mogą stanowić istotny problem w procesie modelowania, ponieważ wiele metod uczenia maszynowego wymaga pełnych zestawów danych do prawidłowego funkcjonowania. W związku z tym podjęto decyzję o przetestowaniu różnych metod uzupełniania (imputacji) brakujących danych oraz rozważeniu możliwości usunięcia obserwacji zawierających niepełne pomiary.

Usuwanie danych niesie ze sobą nieuchronną utratę informacji, a zatem z perspektywy modelowania procesu jest to działanie niepożądane. Dlatego zdecydowano się w pierwszej kolejności brakujące dane uzupełnić. W tym celu wykorzystano metody klasyczne, tj. uzupełnienie danych wartością średnią lub medianą, a także metodami wykorzystującymi uczenie maszynowe, które opisano w dalszej części tej sekcji. Dobór odpowiedniej metody imputacji był podyktowany przede wszystkim jej wpływem na jakość modelu, a także rozkładem danej zmiennej zależnej. Po przetestowaniu klasycznych metod uzupełniania danych (szczegółowe wyniki wpływu poszczególnych metod zostaną zaprezentowane w rozdziale 9) uwagę poświęcono imputacji opartej na metodach uczenia maszynowego. Pierwszą metodą uczenia maszynowego, którą zdecydowano się zastosować, była metoda k-najbliższych sąsiadów (K-NN, od ang. k-nearest neighbors).

K-NN jest jedną z podstawowych metod stosowanych w uczeniu maszynowym, zarówno do klasyfikacji, jak i regresji. Jej działanie opiera się na prostym założeniu, że obiekty (tutaj pomiary) podobne do siebie znajdują się w przestrzeni zmiennych blisko siebie. Metodę można wyrazić formułą [64]:

$$\hat{Y}(x) = \frac{1}{k} \sum_{x_i \in N_k(x)} y_i, \quad (6.1)$$

gdzie:

- $N_k(x)$ - sąsiedztwo x zdefiniowane przez k najbliższych punktów x_i w zbiorze danych.

Bliskość sąsiedniego punktu jest obliczana na podstawie różnego rodzaju miar, z czego najpopularniejszą i zastosowaną w tej pracy jest odległość Euklidesowa.

W pracy zastosowano algorytm realizowany w następujących krokach [78]:

1. Dla każdej obserwacji, w której brakuje wartości zmiennej zależnej, obliczono odległości do innych obserwacji w zbiorze danych, wykorzystując metrykę euklidesową. Odległość ta opiera się na różnicach w dostępnych wartościach zmiennych niezależnych między obserwacjami, ponieważ imputacja dotyczy wyłącznie zmiennych zależnych.
2. Dla każdej obserwacji z brakującymi wartościami zmiennej zależnej, algorytm wybiera k obserwacji (sąsiadów) o najmniejszych odległościach, czyli najbardziej podobnych pod względem dostępnych wartości zmiennych niezależnych.
3. Brakujące wartości zmiennej zależnej są uzupełniane przez obliczenie średniej z odpowiednich wartości zmiennej zależnej u wybranych k sąsiadów. W ten sposób brakujące wartości zmiennej zależnej są estymowane na podstawie obserwacji o podobnych wartościach zmiennych niezależnych, które zawierają pełne dane.

Proces ten opiera się jednak na dość silnym założeniu, że podobieństwo między obserwacjami może być wykorzystane do estymacji brakujących wartości zmiennej zależnej.

Inną testowaną metodą była iteracyjna imputacja (znana również jako MICE, od ang. Multivariate Imputation by Chained Equations). Metoda ta opiera się na iteracyjnym modelowaniu każdej z brakujących wartości jako zmienną zależną, wykorzystując pozostałe kompletne wartości jako zmienne niezależne. Do tego celu stosuje się wybraną funkcję estymacji [79]. W rozpatrywanym przypadku jako funkcję estymacji brakujących wartości przyjęto metodę bazującą na regresji liniowej (5.2), na drzewach decyzyjnych (5.11), a także na regresji Bayesowskiej, opisanej niżej.

Regresja Bayesowska (ang. Bayesian Ridge Regression) jest metodą regresji, która opiera się na podejściu bayesowskim do estymacji parametrów modelu regresji liniowej. W przeciwieństwie do tradycyjnej regresji liniowej, która szacuje pojedyncze wartości dla parametrów modelu, regresja Bayesowska dostarcza rozkład prawdopodobieństwa dla tych parametrów, co umożliwia ocenę ich niepewności [80]. Takie podejście jest często stosowane w kontekście imputacji danych, umożliwiając dostosowanie się do częściowo

kompletnych zbiorów danych oraz kwantyfikowanie niepewności związanej z imputowanymi wartościami [81].

Cel regresji Bayesowskiej polega na znalezieniu rozkładu a posteriori współczynników β po uwzględnieniu obserwowanych danych, co można zapisać jako [80]:

$$p(\beta|y, X) \propto p(y|X, \beta)p(\beta), \quad (6.2)$$

gdzie:

- $p(\beta|y, X)$ - rozkład a posteriori współczynników regresji β ,
- $p(y|X, \beta)$ - funkcja wiarygodności, która mierzy prawdopodobieństwo obserwacji y przy danych współczynnikach β i danych X ,
- $p(\beta)$ - rozkład a priori współczynników β .

Regresja Bayesowska używa teorematu Bayesa do aktualizacji początkowego przekonania o rozkładzie współczynników (a priori) w oparciu o dostępne dane, prowadząc do rozkładu a posteriori, który uwzględnia zarówno informacje a priori jak i informacje płynące z danych [80].

6.1.2. Transformacja danych

Kolejnym istotnym etapem przygotowania danych jest odpowiednia transformacja danych, mająca na celu dostosowanie zakresu wartości zmiennych niezależnych tak, aby były one porównywalne i nie wprowadzały nieproporcjonalnego wpływu na proces uczenia. Praktyka ta jest szczególnie istotna w uczeniu maszynowym, gdzie różnice w skalach mogą prowadzić do problemów z konwergencją algorytmu, wydłużając czas potrzebny na trenowanie modeli lub nawet uniemożliwiając osiągnięcie optymalnych rozwiązań [82].

Ze względu na różnorodność zastosowanych modeli uczenia maszynowego i modeli statystycznych oraz ze względu na ich specyficzne wymagania co do rozkładu danych wejściowych, postanowiono przetestować różne metody transformacji.

Testowanymi metodami transformacji danych w niniejszej pracy są:

- a) Metoda standaryzacji Z - polegająca na przekształceniu danych tak, aby miały średnią równą 0 i odchylenie standardowe równe 1 [83]. Metoda ta jest wyrażana formułą:

$$z' = \frac{x - \mu}{\sigma}, \quad (6.3)$$

gdzie:

- x – wartość pojedynczego przykładu (obserwacji) w zbiorze danych.
- μ – średnia arytmetyczna wszystkich wartości zmiennej w zbiorze danych.
- σ – odchylenie standardowe dla tej zmiennej.

- b) Normalizacja min-max - stosowana w celu przekształcenia danych do określonego zakresu, zwykle między 0 a 1 [83]:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)}, \quad (6.4)$$

Normalizacja min-max stosowana jest najczęściej, kiedy wymagane są dane w określonym zakresie, jak w przypadku sieci neuronowych [83],[84].

- c) Normalizacja według normy - polega na przeskalowaniu wektora zmiennych niezależnych tak, aby jego norma wynosiła 1 (lub inną wybraną wartość). Najczęściej stosuje się normę L2, gdzie każdy wektor cech x jest przekształcany do:

$$x' = \frac{x}{\|x\|_2}, \quad (6.5)$$

gdzie:

- $\|x\|_2$ - norma L2 wektora x .

Metoda ta jest szczególnie przydatna w przypadku metod opierających się na obliczeniach odległości, jak k-NN czy SVM [84].

- d) Normalizacja według maksymalnej wartości - w tej metodzie każda wartość zmiennej jest dzielona przez maksimum absolutne wśród wszystkich wartości tej zmiennej. Formuła wygląda następująco:

$$x' = \frac{x}{\max(|x|)}, \quad (6.6)$$

Używana często w przypadku potrzeby zachowania odpowiednich proporcji w zbiorze danych [84].

6.1.3. Balansowanie klas

Metody opisane w p. 6.1.2 dotyczą zmiennych ciągłych i problemów regresji. W kontekście przygotowania danych do modelowania, innym zagadnieniem jest przetwarzanie zmiennych nominalnych. Zasadnicze wyzwanie przy pracy ze zmiennymi nominalnymi objawia się nierównowagą klas. Jest to stan, w którym jedna z klas jest znacznie mniej liczna niż druga. Taka dysproporcja może generować problemy w procesie modelowania, ponieważ modele uczenia maszynowego mogą wykazywać tendencję do faworyzowania liczniejszej klasy, co obniża skuteczność klasyfikacji dla klasy mniejszościowej.

W odpowiedzi na problem nierównowagi, w literaturze [85] opisane są różne techniki balansowania klas, które można sklasyfikować do czterech głównych kategorii:

- Oversampling klas mniejszościowych - techniki polegające na zwiększeniu liczby obserwacji w klasie mniejszościowej. Realizowane np. poprzez proste duplikowanie istniejących próbek.
- Techniki syntetycznego oversamplingu klas mniejszościowych (ang. synthetic minority oversampling technique, SMOTE), które polegają na generowaniu syntetycznych obserwacji na bazie interpolacji między istniejącymi obserwacjami.
- Undersampling klas większościowych - techniki polegające na redukcji liczby próbek w klasie dominującej. Realizowane np. poprzez losowe eliminowanie próbek z klasy większościowej.
- Kombinowane podejścia - metody te integrują elementy oversamplingu i undersamplingu w celu równoważenia wpływu obu klas na model, minimalizując jednocześnie utratę istotnych informacji.

Inną stosowaną metodą jest przypisanie wag klas, co jest szczególnie popularne w modelach uczenia maszynowego umożliwiającym uwzględnienie wag w procesie trenowania. Metoda ta polega na przypisaniu większej wagi klasom mniej licznie reprezentowanym w zbiorze danych, co pozwala zwiększyć ich wpływ podczas uczenia modelu [85].

6.2. Trenowanie i ewaluacja modeli

Procedura trenowania i ewaluacji modeli składała się z kilku etapów. W pierwszej fazie dane podzielono na zbiór treningowy i testowy. Proces optymalizacji hiperparametrów oraz trenowania modeli przeprowadzono wyłącznie na zbiorze treningowym, natomiast zbiór testowy posłużył do ostatecznej oceny jakości modelu. Do wyboru optymalnych wartości hiperparametrów zastosowano walidację krzyżową (k-krotną), co zapewniło bardziej wiarygodną ocenę modeli i zminimalizowało ryzyko

przeuczenia. Następnie, wytrenowane modele z dobranymi hiperparametrami zostały ocenione metodami opisanymi w punkcie 6.3, również z wykorzystaniem walidacji krzyżowej, aby zagwarantować spójną ocenę ich jakości.

6.2.1. Podział danych na zbiór treningowy i testowy

Zastosowany w pracy podział danych prezentuje się następująco:

$$\begin{aligned}(X, Y) &= (X, Y)^{(train)} \cup (X, Y)^{(test)}, \\ (X, Y)^{(train)} \cap (X, Y)^{(test)} &= \emptyset,\end{aligned}\tag{6.7}$$

gdzie

- (X, Y) – zbiór danych empirycznych, gdzie X to wartości zmiennych niezależnych, a Y to wartości zmiennej zależnej,
- $(X, Y)^{(train)}$ – zbiór danych treningowych,
- $(X, Y)^{(test)}$ – zbiór danych testowych,

przy czym proporcje podziału zostały zdefiniowane jako:

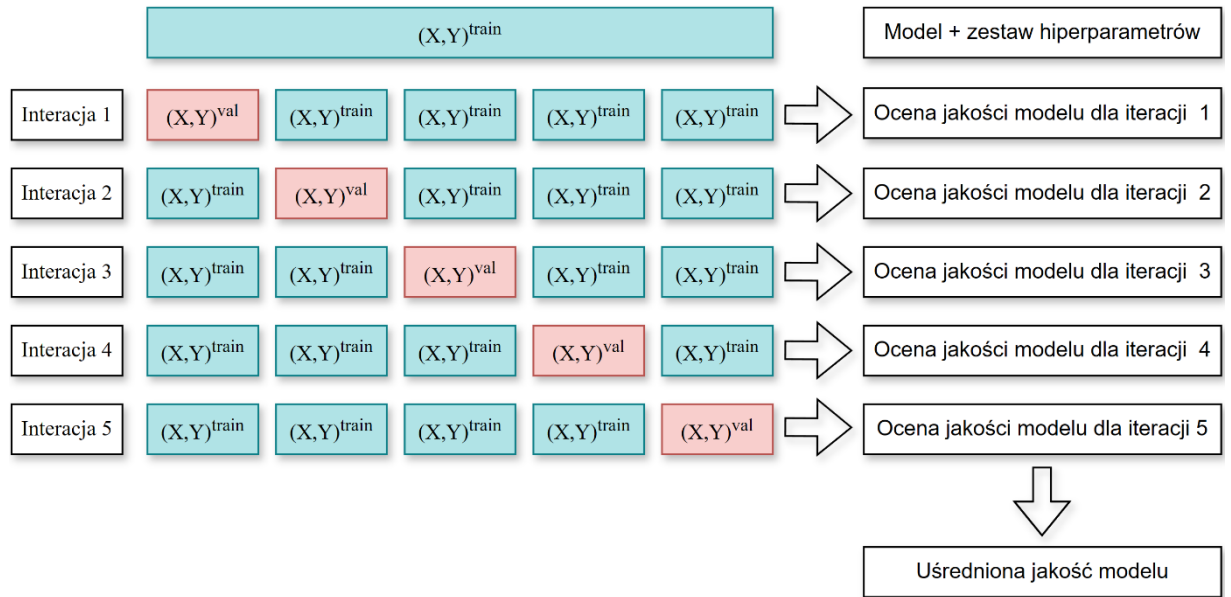
$$\begin{aligned}|(X, Y)^{(train)}| &= 0.8 \times |(X, Y)|, \\ |(X, Y)^{(test)}| &= 0.2 \times |(X, Y)|,\end{aligned}\tag{6.8}$$

Aby zapewnić, że zbiory są reprezentatywne, zastosowano procedurę losowego wyboru próbek do każdego zbioru, przy czym w przypadku ominimalnej zmiennej zależnej dbano o zachowanie zrównoważonej proporcji klas występujących w całym zbiorze danych (procedura balansowania klas została szczegółowo opisana w punkcie 0).

6.2.2. Walidacja krzyżowa i optymalizacja hiperparametrów

Do trenowania modeli wraz z równoczesną optymalizacją hiperparametrów, a także do późniejszej analizy porównawczej, zastosowano metodę k-krotnej walidacji krzyżowej (ang. K-fold cross-validation), zgodnie ze schematem przedstawionym na Rys. 6.1. Procedura ta zakładała podział zbioru danych treningowych na K równych segmentów (w badaniach przyjęto $K=5$), określanych jako "foldy". W każdej iteracji, model był trenowany na $K-1$ z tych segmentów, a testowany na pozostałym segmencie walidacyjnym. Proces ten był powtarzany K razy, przy czym za każdym razem inny segment danych pełnił funkcję zestawu walidacyjnego. Po zakończeniu każdej iteracji,

dokonywano oceny jakości modelu na zbiorze walidacyjnym. Ostateczna jakość modelu była oceniana na podstawie średniej oceny ze wszystkich iteracji.



Rys. 6.1 Proces trenowania i oceny modeli z wykorzystaniem K-krotnej walidacji krzyżowej

Przedstawioną procedurę można wyrazić matematycznie [86]:

$$\psi_{CV}^{(val)} = \frac{1}{k} \sum_{i=1}^k \frac{1}{|(X, Y)^{(val)}|} \sum_{x \in (X, Y)_i^{(val)}} \mathcal{L}(y, M_{\lambda}((X, Y)_i^{(train)})), \quad (6.9)$$

gdzie:

- $\psi_{CV}^{(val)}$ - uśredniona jakość modelu w walidacji krzyżowej,
- \mathcal{L} - funkcja oceniająca jakość modelu (różna dla problemów regresji i klasyfikacji)
- M_{λ} – model z określoną konfiguracją hiperparametrów (λ).

Przy czym w pracy podczas optymalizacji hiperparametrów funkcja \mathcal{L} w przypadku regresji przyjmowała postać [86]:

$$\mathcal{L}(y, \hat{y}) = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|, \quad (6.10)$$

gdzie:

- N - całkowita liczba próbek w zbiorze danych,
- y_i - rzeczywista wartość zmiennej zależnej dla i-tego przykładu,

- \hat{y}_i - przewidywana wartość zmiennej zależnej dla i-tego przykładu.

Modele regresji były zatem oceniane ze względu na wartość średniego błędu absolutnego (ang. mean absolute error, MAE). Wybór tej metryki był podyktowany powszechnością jej użycia oraz łatwością interpretacji (zob. p.6.3.1).

Z kolei w przypadku klasyfikacji funkcja \mathcal{L} oceniała dokładność klasyfikacji i przyjmowała postać [86]:

$$\mathcal{L}(y, \hat{y}) = \frac{1}{N} \sum_{i=1}^N I(y_i \neq \hat{y}_i), \quad (6.11)$$

gdzie:

- $I(y_i \neq \hat{y}_i)$ - funkcja wskaźnikowa, która przyjmuje wartość 1, jeśli przewidywana wartość zmiennej zależnej \hat{y}_i jest równa rzeczywistej wartości zmiennej zależnej y_i , i 0, jeśli są one różne.

Do optymalizacji hiperparametrów wykorzystano metodę przeszukiwania siatki (ang. Grid search), polegającej na ocenie jakości modelu (6.9) we wszystkich punktach zdefiniowanej przestrzeni hiperparametrów. W Tabeli 6.1. przedstawiono przykładową przestrzeń hiperparametrów, przy czym dla każdego modelu przestrzeń ta różniła się, uwzględniając specyficzne parametry i wymagania charakterystyczne dla każdego z modeli. Nazewnictwo zastosowane w tabeli ściśle odzwierciedla terminologię stosowaną w dokumentacji poszczególnych bibliotek programistycznych.

Tabela 6.1 Zestawienie testowanych modeli i hiperparametrów

Model	Hiperparametr	Opis	Zakres testowanych wartości	Rozmiar skoku
XGBoost	n_estimators	Liczba drzew w modelu	200 do 500	100
	learning_rate	Tempo uczenia się modelu	0.01 do 0.05	0.01
	max_depth	Maksymalna głębokość drzew	1 do 6	3
	colsample_bytree	Procent użytych cech przy budowie drzewa	0.3 do 0.7	0.1
SVR	kernel	Funkcja jądra	linear, rbf, poly	-
	C	Parametr regularyzacji	1 do 100	2
	gamma	Współczynnik funkcji jądra	scale, auto	-
	epsilon	Tolerancja na błąd	0.001 do 1	0.001
Las losowy	n_estimators	Liczba drzew w lesie	50 do 400	50
	max_depth	Maksymalna głębokość drzew	3 do 30	2
	min_samples_split	Minimalna liczba próbek do podziału w węźle	2, 5, 10	-
	min_samples_leaf	Minimalna liczba próbek w liściu	1, 2, 4	-
	max_features	Maksymalna liczba cech rozważanych przy podziale	auto, sqrt, log2	-
	max_depth	Maksymalna głębokość drzewa	10 do 30	5
	min_samples_split	Minimalna liczba próbek do podziału w węźle	2, 5, 10	-

Drzewa decyzyjne	min_samples_leaf	Minimalna liczba próbek w liściu	1, 2, 4	-
	max_features	Liczba cech używanych do podziału	auto, sqrt, log2	-
Sieci neuronowe	n_layers	Liczba warstw ukrytych w modelu	2, 3, 4	-
	n_neurons	Liczba neuronów w każdej warstwie	64, 128, 256	-
	learning_rate	Tempo uczenia się sieci	0.001, 0.01, 0.1	-
	dropout_rate	Prawdopodobieństwo wyłączenia neuronów w czasie treningu	0.1, 0.2, 0.3	-
	activation	Funkcja aktywacji	relu, elu, tanh, sigmoid	-
	optimizer	Optymalizator używany do treningu	adam, rmsprop, sgd, adagrad	-
	batch_size	Liczba próbek przetwarzanych na raz	16, 32, 64	-
	epochs	Liczba epok treningowych	100, 200, 300	-
	kernel_regularizer	Regularyzacja L2 dla wag	0.001, 0.01	-
	bias_regularizer	Regularyzacja L2 dla biasów	0.001, 0.01	-
	early_stopping	Zastosowanie wczesnego zatrzymania treningu	True, False	-
	patience	Liczba epok bez poprawy przed zatrzymaniem	10, 20, 30	-
	batch_normalization	Zastosowanie normalizacji batchowej	True, False	-

Pełna procedura trenowania modeli i optymalizacji hiperparametrów przebiegała zatem w następujących krokach [86]:

1. Wybór hiperparametrów w iteracji t (w pierwszej iteracji $t = 1$).
 - Wybierany jest zestaw hiperparametrów z przestrzeni hiperparametrów $\lambda(t) \in \Lambda$.
2. Trening i ocena modelu.
 - Następuje trenowanie modelu $M_{\lambda(t)}$ na danych uczących $(X, Y)^{(train)}$ i jego walidacja na danych walidacyjnych $(X, Y)^{(val)}$, zgodnie z metodą walidacji krzyżowej (6.12).
3. Aktualizacja hiperparametrów.
 - Dobierany jest kolejny zestaw hiperparametrów do wypróbowania, $\lambda(t + 1) \in \Lambda$.
4. Wykonywanie powyższych kroków w pętli aż do kryterium zatrzymania – w tym przypadku wyczerpania możliwych kombinacji hiperparametrów z przestrzeni hiperparametrów (Λ).

Procedura ta została zastosowana do wszystkich rozpatrywanych modeli. Wykorzystanie walidacji krzyżowej było użyteczne, ze względu na możliwość

minimalizacji problemu przeuczenia. Pozwoliła również na lepsze wykorzystanie dostępnych danych, ponieważ każda obserwacja została użyta zarówno do treningu, jak i walidowania modelu.

Po przeprowadzeniu opisanej procedury, wytrenowane modele z optymalną konfiguracją hiperparametrów (λ^*) zostały porównane przy użyciu metod oceny jakości modeli. Do oceny zastosowano walidację krzyżową, a funkcje oceniające jakość modeli \mathcal{L} były zgodne z metodami opisanymi w punkcie 6.3.

6.3. Metody oceny modeli

W niniejszym podrozdziale omówione zostaną wykorzystane w pracy metody oceny jakości modeli zarówno regresyjnych, jak i klasyfikacyjnych. Zastosowanie wielu metryk (wskaźników) umożliwia uzyskanie pełniejszego obrazu działania modeli. Modele regresyjne, przewidujące parametry takie jak chropowatość powierzchni i prędkość usuwania materiału, wymagają dokładnej oceny pod kątem trafności przewidywań, identyfikacji potencjalnych błędów oraz analizy reszt w celu sprawdzenia ich losowości. W tym celu zastosowano metryki takie jak średni błąd kwadratowy (MSE), średni błąd bezwzględny (MAE), współczynnik determinacji (R^2) oraz test Durбина-Watsona.

W przypadku modeli klasyfikacyjnych, których celem jest przewidywanie ciągłości procesu obróbki oraz klasy szczeliny erozyjnej (w klasyfikacji wieloklasowej), kluczowa jest ocena zarówno ogólnej skuteczności modelu, jak i jego zdolności do poprawnego rozróżniania między klasami. Dlatego stosuje się metryki takie jak dokładność, czułość, precyzja, a także analizę krzywej ROC oraz pole pod krzywą ROC (AUC).

Kombinacja tych metryk umożliwia bardziej kompleksowe modelowanie procesu WEDM, uwzględniając szeroki zakres wymagań oraz scenariuszy zastosowania w technologii obróbki elektroerozyjnej.

6.3.1. Ocena modeli regresji

Jedną z podstawowych miar oceny jakości dopasowania modeli regresyjnych jest współczynnik determinacji R^2 (6.13) [61].

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \geq 0, \quad (6.13)$$

gdzie \bar{y} to średnia arytmetyczna zmierzonych wartości zmiennej zależnej.

Współczynnik ten informuje o tym, jaka część zmienności (wariancji) zmiennej zależnej jest wyjaśniana przez zmienne niezależne zawarte w modelu. Innymi słowy, R^2 pokazuje, w jakim stopniu zmiany w zmiennej zależnej można przypisać wpływowi zmiennych niezależnych zastosowanych w modelu regresji.

Inną popularną metryką jest błąd średniokwadratowy (ang. Mean Squared Error, MSE) (6.14) [86], która wyraża średnią kwadratową różnic między przewidywanymi a rzeczywistymi wartościami zmiennej zależnej.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad (6.14)$$

Kolejną metodą oceny modelu jest wspomniany wcześniej (p.6.2.2) średni błąd bezwzględny MAE [86] (6.15), czyli średnia wartość bezwzględnych różnic między przewidywanymi a rzeczywistymi wartościami. Metryka ta jest mniej wrażliwa na odstające wartości niż MSE, co czyni ją bardziej odporną na wartości odstające.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (6.15)$$

MAPE [86] (średni procentowy błąd bezwzględny), wyraża średnią procentową różnicę między przewidywanymi a rzeczywistymi wartościami. MAPE, w odróżnieniu od MAE, wyraża błąd w postaci względnej, co pozwala lepiej ocenić jakość przewidywania w kontekście różnych skali danych. Metryka ta, podobnie jak MAE, jest odporna na wartości odstające, ale jej użycie jest bardziej adekwatne w przypadkach, gdy istotna jest analiza proporcjonalnych błędów w danych.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i}, \quad (6.16)$$

W pracy wskaźnik ten wyrażany jest w skali od 0 do 1.

Poza oszacowaniem błędów modeli stosuje się również analizę reszt. W pracy posłużono się m.in. analizą wizualną wykresów reszt modelu, gdzie dodatkowym narzędziem diagnostycznym ułatwiającym ocenę rozkładu danych była krzywa Lowess (ang. locally weighted scatterplot smoothing) [63]. Metoda ta opiera się na lokalnym przybliżeniu danych, co oznacza, że dla każdego punktu danych tworzy się osobne ważone

modele regresji liniowej, przy użyciu metod najmniejszych kwadratów, ale w ograniczonym sąsiedztwie tego punktu. Wagi dla lokalnej regresji liniowej określane są przy pomocy funkcji wagowej, wyrażoną wzorem (6.17) [63].

$$\omega_i = \left(1 - \left|\frac{x_i - x}{h}\right|^3\right)^3, \quad (6.17)$$

dla $|x_i - x| < h$, gdzie h jest parametrem wygładzania określającym "szerokość okna" lokalnej regresji, a ω_i to waga przypisana i -temu punktowi.

Lokalnie ważona regresja liniowa dla punktu x_i minimalizuje ważoną sumę kwadratów reszt [63]:

$$\sum_i \omega_i (y_i - \beta_0 - \beta_1 x_i)^2. \quad (6.18)$$

Poza wizualną oceną, przeprowadzono testy statystyczne rozkładu reszt. W celu zbadania obecności autokorelacji składnika resztowego wykorzystano test Durбина-Watsona (DW) (6.19) [87].

$$DW = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2}, \quad (6.19)$$

gdzie:

- e_t oznacza reszty modelu regresji w czasie t
- n jest liczbą obserwacji

Interpretacja wyniku testu DW opiera się na wartościach granicznych, zależnych od liczby obserwacji oraz zmiennych. Ogólne zasady prezentują się następująco [61]:

- Wartość DW bliska 2 sugeruje brak autokorelacji w resztach.
- Wartość DW mniejsza niż 2 wskazuje na dodatnią autokorelację.
- Wartość DW większa niż 2 sugeruje ujemną autokorelację.

Kolejnym testem było sprawdzenie normalności rozkładu reszt. Aby zbadać rozkład pod tym względem posłużono się testem Jarque-Bera (JB) (6.20) [88], który pozwala ocenić, czy reszty mają rozkład zbliżony do normalnego pod względem skośności i kurtozy.

$$JB = \frac{n}{6} \left(S^2 + \frac{1}{4} (K - 3)^2 \right), \quad (6.20)$$

Gdzie:

- S – skośność rozkładu
- K – kurtoza rozkładu.

Wysoka wartość statystyki JB , w połączeniu z niską p -wartością (w pracy przyjęto $p < 0.05$), wskazuje na odchylenie reszt od rozkładu normalnego, co sugeruje odrzucenie hipotezy zerowej o normalności rozkładu reszt. Z kolei niska wartość statystyki JB oraz wysoka p -wartość ($p > 0.05$) sugerują brak podstaw do odrzucenia hipotezy zerowej, co oznacza, że rozkład reszt może być zbliżony do normalnego [88].

6.3.2. Ocena modeli klasyfikacyjnych

Podstawowym narzędziem w ocenie modeli klasyfikacji binarnej, jest macierz pomyłek (Tabela 6.2). Macierz składa się z czterech różnych wartości: prawdziwie pozytywnych (ang. true positive, TP), fałszywie pozytywnych (ang. false positive, FP), prawdziwie negatywnych (ang. true negative, TN) oraz fałszywie negatywnych (ang. false negative, FN).

Tabela 6.2 Przykładowa macierz pomyłek [89]

	Klasyfikacja pozytywna	Klasyfikacja negatywna
Rzeczywiste pozytywne	TP	FN
Rzeczywiste negatywne	FP	TN

Interpretacja powyższej macierzy prezentuje się następująco [89]:

- Prawdziwie pozytywne (TP) - liczba przypadków, które są pozytywne i zostały poprawnie zidentyfikowane przez model jako pozytywne.
- Fałszywie pozytywne (FP) - liczba przypadków, które są negatywne, ale błędnie zidentyfikowane przez model jako pozytywne. Jest to błąd typu I.
- Prawdziwie negatywne (TN) - liczba przypadków, które są negatywne i zostały poprawnie zidentyfikowane jako negatywne.
- Fałszywie negatywne (FN) - liczba przypadków, które są pozytywne, ale błędnie zidentyfikowane jako negatywne. Jest to błąd typu II.

W kontekście klasyfikacji ciągłości obróbki kategorie wyników klasyfikacji wyjaśnia Tabela 6.3.

Tabela 6.3 Kategorie wyników klasyfikacji modelu ciągłości obróbki

Klasyfikacja	Opis
Prawdziwie pozytywna (TP)	Model poprawnie klasyfikuje przypadek wystąpienia przerwania ciągłości obróbki jako przerwanie ciągłości obróbki.
Fałszywie pozytywna (FP)	Model błędnie klasyfikuje przypadek, w którym ciągłość obróbki nie została przerwana, jako przerwanie ciągłości obróbki
Prawdziwie negatywna (TN)	Model poprawnie klasyfikuje przypadek, w którym ciągłość obróbki została zachowana, jako brak przerwania ciągłości obróbki
Fałszywie negatywna (FN)	Model błędnie klasyfikuje przypadek przerwania ciągłości obróbki jako zachowanie ciągłości obróbki

Na podstawie macierzy pomyłek można wyprowadzić różnorodne metryki oceniające różne aspekty modeli. Poniżej zamieszczono krótką charakterystykę metryk wykorzystanych w badaniach [89].

- Dokładność klasyfikacji (ang. Accuracy, ACC) (6.21), mierzy stosunek liczby poprawnych klasyfikacji do całkowitej liczby przypadków.

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}. \quad (6.21)$$

W pracy dokładność klasyfikacji wyraża się w skali od 0 do 1 lub zamiennie, w skali procentowej.

- Czulość (ang. Sensitivity, SEN) (6.22), mierzy zdolność modelu do poprawnego identyfikowania przypadków pozytywnych.

$$SEN = \frac{TP}{TP + FN}. \quad (6.22)$$

- Precyzja (ang. Precision, PRE) (6.23), mierzy stosunek prawdziwie pozytywnych przypadków do wszystkich przypadków uznanych przez model za pozytywne.

$$PRE = \frac{TP}{TP + FP}. \quad (6.23)$$

- Specyficzność (ang. Specificity, SPE) (6.24), mierzy zdolność modelu do poprawnego identyfikowania przypadków negatywnych.

$$SPE = \frac{TN}{TN + FP}. \quad (6.24)$$

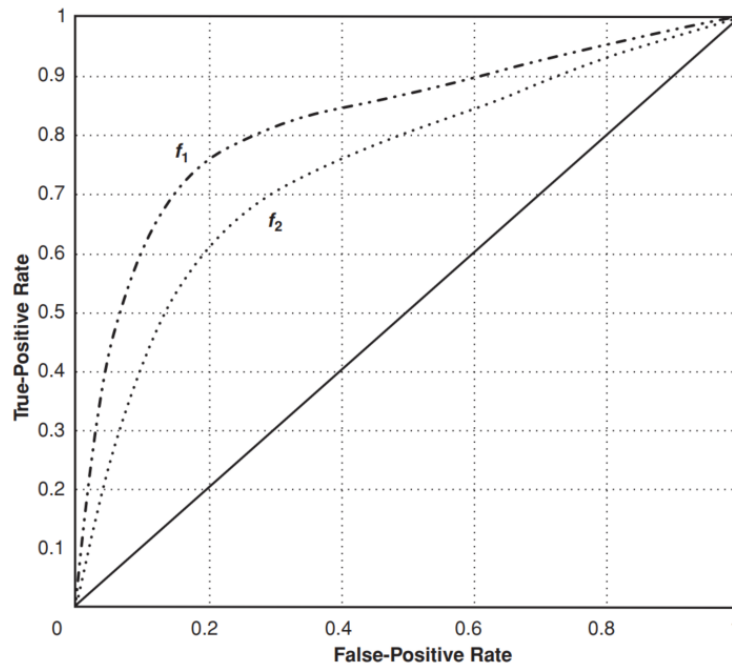
Wykorzystując powyższe metryki, można zwizualizować zdolności klasyfikacyjne modelu przy pomocy tzw. krzywej ROC (ang. Receiver Operating Characteristic). Krzywa ROC jest graficzną reprezentacją zdolności klasyfikatora do różnicowania między klasami przy różnych progach decyzyjnych. Próg decyzyjny (ang. decision threshold) to wartość graniczna, która określa, przy jakim poziomie prawdopodobieństwa model przypisuje obserwację do danej klasy. Zmiana progu decyzyjnego wpływa na liczbę prawdziwie i fałszywie pozytywnych oraz negatywnych klasyfikacji, co ma bezpośredni wpływ na działanie modelu. Krzywa ROC jest tworzona poprzez wykreślenie wartości TPR (ang. True Positive Rate, czyli czułości) (6.25) względem FPR (ang. False Positive Rate) (6.26) dla różnych wartości progowych, co pozwala zrozumieć, jak model zachowuje się w zależności od wyboru progu decyzyjnego.

$$TPR = SEN. \quad (6.25)$$

$$FPR = 1 - SPE. \quad (6.26)$$

Zmieniając próg decyzyjny klasyfikatora, otrzymujemy różne wartości dla TPR i FPR, co pozwala na wykreślenie krzywej ROC. Krzywa ta zaczyna się w punkcie (0,0) (idealna specyficzność, brak czułości) i kończy w punkcie (1,1) (maksymalna czułość, brak specyficzności). Idealny klasyfikator osiągnąłby punkt (0,1), co oznacza brak fałszywych pozytywów (FPR=0) i doskonałą czułość (TPR=1) [89].

Poniżej na Rys. 6.2 przedstawiono przykładową krzywą ROC dla dwóch hipotetycznych modeli klasyfikacyjnych f_1 i f_2 .



Rys. 6.2 Krzywe ROC dla dwóch hipotetycznych klasyfikatorów f_1 i f_2 [89]

Z krzywą ROC najczęściej stosuje się powiązaną z nią metrykę AUC (ang. Area Under the Curve). AUC, czyli pole powierzchni pod krzywą ROC, jest skalarną miarą zdolności klasyfikatora do rozróżniania między klasami. Wartość AUC może być interpretowana jako prawdopodobieństwo, że klasyfikator przypisze wyższą wartość obserwacji pozytywnej niż obserwacji negatywnej, gdy jedna z każdej klasy jest losowo wybrana. Wartości AUC wahają się od 0 do 1, gdzie [89]:

- $AUC = 0.5$ oznacza brak zdolności dyskryminacyjnej, czyli działanie na poziomie losowego zgadywania.
- $AUC = 1$ oznacza idealną zdolność do rozróżniania między klasami bez popełnienia żadnych błędów.
- $AUC < 0.5$ teoretycznie sugeruje, że klasyfikator działa gorzej niż losowe zgadywanie.

Aby obliczyć wartość AUC, często stosuje się metody numeryczne, takie jak reguła trapezów, gdzie pary punktów na krzywej ROC są używane do aproksymacji małych trapezów pod krzywą, a ich pola są sumowane.

Kolejną wykorzystaną miarą służącą do oceny jakości klasyfikacji binarnej jest korelacja Matthews'a, znana również jako Matthews Correlation Coefficient (MCC) [90]. Jest to wskaźnik, który bierze pod uwagę prawdziwe pozytywy, prawdziwe negatywy, fałszywe pozytywy i fałszywe negatywy, stanowiąc bardziej wyważoną miarę niż np.

precyzja lub dokładność klasyfikacji, szczególnie w przypadkach niezrównoważonych zbiorów danych [90].

MCC przyjmuje wartość w przedziale od -1 do 1:

- Wartość 1 oznacza idealną klasyfikację.
- Wartość 0 wskazuje na losową klasyfikację.
- Wartość -1 oznacza całkowicie błędną klasyfikację

MCC można wyrazić formułą [90]:

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (6.27)$$

Powyższe metryki oceny modeli można również uogólnić i zastosować do klasyfikacji wieloklasowej, co znajduje zastosowanie w niniejszej pracy w przypadku klasyfikacji szerokości szczeliny erozyjnej. W przypadku klasyfikacji wieloklasowej, metryki takie jak dokładność, czułość, precyzja czy specyficzność są obliczane osobno dla każdej klasy, a następnie odpowiednio uśredniane. Najczęściej stosuje się dwa rodzaje uśredniania [91]:

- Uśrednianie makro (ang. macro averaging) – każda klasa ma taką samą wagę, niezależnie od liczby obserwacji. Miary są obliczane dla każdej klasy osobno, a następnie wyciągana jest średnia arytmetyczna.
- Uśrednianie mikro (ang. micro averaging) – wyniki są ważone proporcjonalnie do liczności klas, co oznacza, że bardziej liczne klasy mają większy wpływ na wynik końcowy.

W niniejszej pracy zastosowano uśrednianie makro, ze względu na równomierny rozkład obserwacji pomiędzy klasami.

Dodatkowo, w klasyfikacji wieloklasowej zastosowano krzywe ROC i obliczono wartość AUC dla każdej klasy w podejściu one-vs-all (jedna klasa kontra wszystkie pozostałe). Następnie wartości AUC zostały uśrednione w celu uzyskania globalnej miary zdolności modelu do klasyfikacji

6.4. Interpretacja modeli

Aby zrealizować cel szczegółowy pracy, polegający na analizie wpływu parametrów technologicznych na wybrane wielkości charakteryzujące proces WEDM, konieczne jest zastosowanie odpowiednich narzędzi do interpretacji modeli procesu. W uczeniu

maszynowym analiza wpływu poszczególnych zmiennych niezależnych na przewidywaną wartość modelu bywa trudna, zwłaszcza w przypadku złożonych modeli o charakterze „czarnej skrzynki” takich jak modele oparte na lasach losowych, metodzie gradient boosting czy sieciach neuronowych. Jedną z najskuteczniejszych metod interpretacji takich modeli jest metoda SHAP (ang. SHapley Additive exPlanations) [92] [93], która opiera się na teorii wartości Shapleya z teorii gier.

Podstawą metody SHAP jest wyznaczenie wartości Shapleya dla każdej zmiennej niezależnej. Wartości te oblicza się, analizując marginalne wpływy zmiennych niezależnych, czyli różnice w wartościach przewidywanych przez model przy obecności i nieobecności danej zmiennej w różnych kombinacjach z innymi zmiennymi. Nieobecność zmiennej rozumiane jest tutaj jako zastąpienie jej wartości próbkami referencyjnymi (np. średnią lub losową próbką z rozkładu), co symuluje sytuację, w której model nie uwzględnia informacji o tej zmiennej. Analiza tych różnic pozwala oszacować wkład zmiennej na wynik modelu [92].

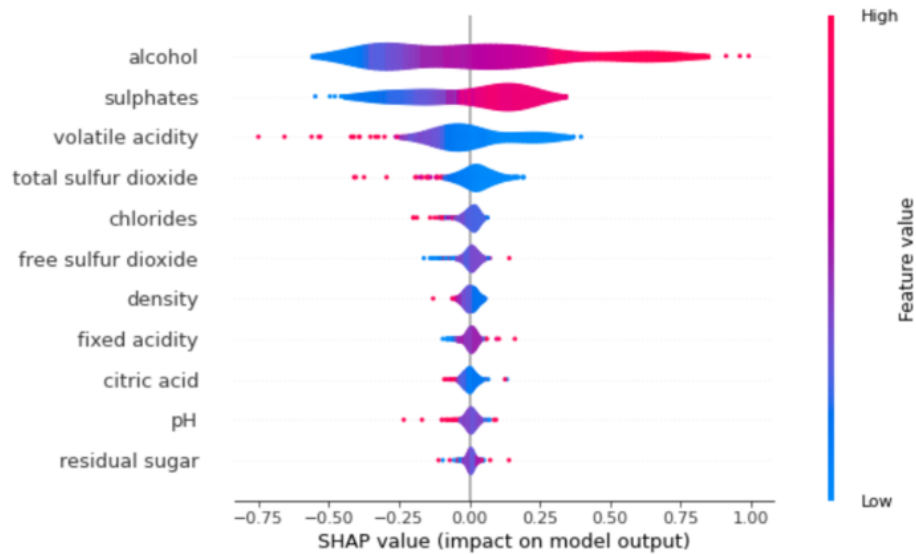
Wartość SHAP dla danej zmiennej niezależnej i w przypadku konkretnej obserwacji x można zapisać jako [92]:

$$\phi_i(x) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|! (|N| - |S| - 1)!}{|N|!} [f(S \cup \{i\}) - f(S)], \quad (6.28)$$

gdzie:

- S – zbiór zmiennych niezależnych nie zawierający zmiennej niezależnej i ,
- N – zbiór wszystkich zmiennych niezależnych,
- $f(S)$ – wartość przewidywana modelu na podstawie zbioru zmiennych niezależnych S .

Wartości SHAP, obliczone dla każdej zmiennej niezależnej i każdej obserwacji, zostały wykorzystane do wizualizacji wpływu zmiennych z użyciem wykresu punktowego. Przykładowy wykres punktowy wartości SHAP wraz jego interpretacją przedstawiono poniżej.



Rys. 6.3 Przykładowy wykres punktowy wartości SHAP [92]

Wykres punktowy SHAP ilustruje wpływ poszczególnych zmiennych na wynik modelu.

- Oś pionowa przedstawia zmienne niezależne uporządkowane według globalnego wpływu na wynik modelu. Zmienna „alcohol” ma największy wpływ, a „residual sugar” najmniejszy.
- Oś pozioma pokazuje wartości SHAP, które wskazują, jak dana zmienna wpływa na wynik. Wartości dodatnie (na prawo) zwiększają wynik, a ujemne (na lewo) go zmniejszają.
- Kolory odpowiadają wartościom zmiennych w danych obserwacjach: różowy oznacza wysokie wartości, niebieski niskie.
- Kształt punktów pokazuje gęstość obserwacji – szersze obszary oznaczają większą liczbę punktów.

7. OPIS BADAŃ EKSPERYMENTALNYCH

Rozdział ten został poświęcony opisowi warunków oraz metodyki badań eksperymentalnych przeprowadzonych w ramach niniejszej pracy. Badania te zostały zrealizowane w warunkach przemysłowych, w przedsiębiorstwie SM Hydro, które wybrano jako miejsce wdrożenia opracowanych rozwiązań.

W pierwszej części rozdziału (p.7.1) omówiono materiały oraz aparaturę użyte podczas badań. Opis ten obejmuje charakterystykę obrabiarki elektroerozyjnej, materiału obrabianego, wycinanych próbek, stanowiska do badań chropowatości powierzchni oraz mikroskopów współrzędnościowych.

Druga część rozdziału (p.7.2) została poświęcona metodom badań eksperymentalnych, gdzie przedstawiono plan badań, zmienne procesu oraz zastosowane plany eksperymentu.

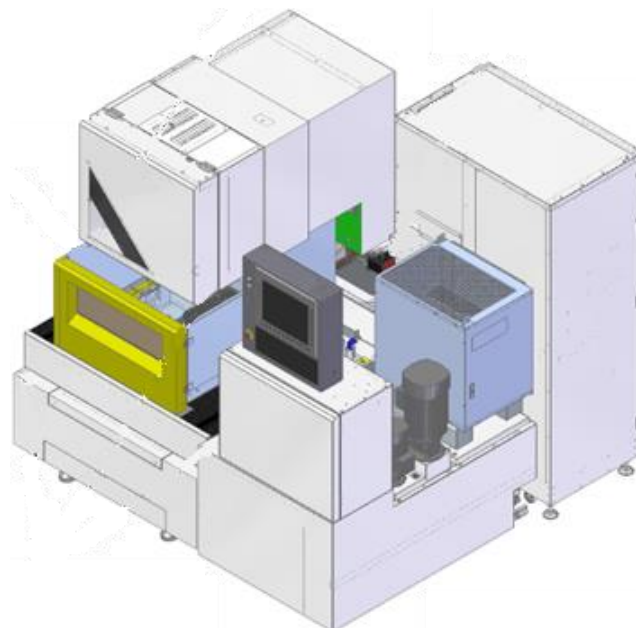
7.1. Materiały i aparatura badawcza

7.1.1. Obrabiarka elektroerozyjna

Badania eksperymentalne cięcia elektroerozyjnego realizowane były na elektrodrążarce drutowej Fanuc α – C600iC przedstawionej na Rys. 7.1. Obrabiarka ta w firmie SM Hydro przeznaczona jest do wycinania elementów mechanizmu satelitowego silnika hydraulicznego. W Tabeli 7.1 przedstawiono dane techniczne maszyny.

Tabela 7.1 Dane techniczne elektrodrążarki drutowej Fanuc α – C600iC

Dane techniczne	Fanuc α – C600iC
Maksymalna waga przedmiotu obrabianego	1000 kg
Zakres posuwu stołu roboczego	600 x 400 mm
Maksymalna prędkość posuwu stołu	2000 mm/min
Minimalny przyrost pozycjonowania napędu	0.0001 mm
Maksymalna waga elektrody	16 kg

Rys. 7.1 Widok 3D elektrodrążarki drutowej Fanuc α – C600iC

7.1.2. Materiał obrabiany

Do badań wykorzystano stal stopową o nazwie handlowej IMPAX. Wybór tego materiału był podyktowany wymogami wdrożeniowymi - stal ta stanowi główny materiał wykorzystywany do produkcji silników hydraulicznych w przedsiębiorstwie.

Szczegóły dotyczące składu chemicznego oraz właściwości mechanicznych wybranego materiału przedstawiono w Tabeli 7.2 i Tabeli 7.3.

Tabela 7.2 Skład chemiczny stali użytej w badaniach [%]

C	Si	Mn	Cr	Ni	Mo	S
0.37	0.3	1.4	2.0	1.0	0.2	<0.01

Tabela 7.3 Wybrane właściwości mechaniczne stali użytej w badaniach

Gęstość [g/cm ³]	Twardość [HB]	Wytrzymałość na rozciąganie R _m , [Mpa]	Granica plastyczności R _e , [Mpa]
7.8	330	1020	900

7.1.3. Wycinane próbki

W ramach badań wycinano próbki w kształcie prostopadłościanów o różnych wymiarach w zakresie od 65x5x25 do 85x5x25 mm. Przykładowe próbki przedstawiono na zdjęciach poniżej.

a)



b)



Rys. 7.2 Widok próbek wycinanych w ramach badań eksperymentalnych, na których przeprowadzono a) pomiary chropowatości powierzchni oraz b) szerokości szczeliny erozyjnej

W początkowej fazie badań, próbki były wykorzystywane wyłącznie do pomiarów chropowatości powierzchni. W miarę postępu badań, zmodyfikowano wycinany kształt próbek, dodając dodatkową ścieżkę przeznaczoną do pomiarów szerokości szczeliny erozyjnej.

7.1.4. Stanowisko do badań chropowatości powierzchni

Badania chropowatości powierzchni odbywały się przy pomocy chropowatościomierza Mitutoyo SJ-210. Stanowisko do badań przedstawiono na Rys. 7.3



Rys. 7.3 Stanowisko do badań chropowatości powierzchni

7.1.5. Mikroskopy współrzędnościowe

Pomiary szerokości szczeliny erozyjnej zrealizowano przy pomocy mikroskopów współrzędnościowych przedstawionych na Rys. 7.4.

a)



b)



Rys. 7.4 Mikroskopy współrzędnościowe Nikon iNEXIV VMA-2520 a) i KEYENCE IM b)

7.2. Metody badań eksperymentalnych

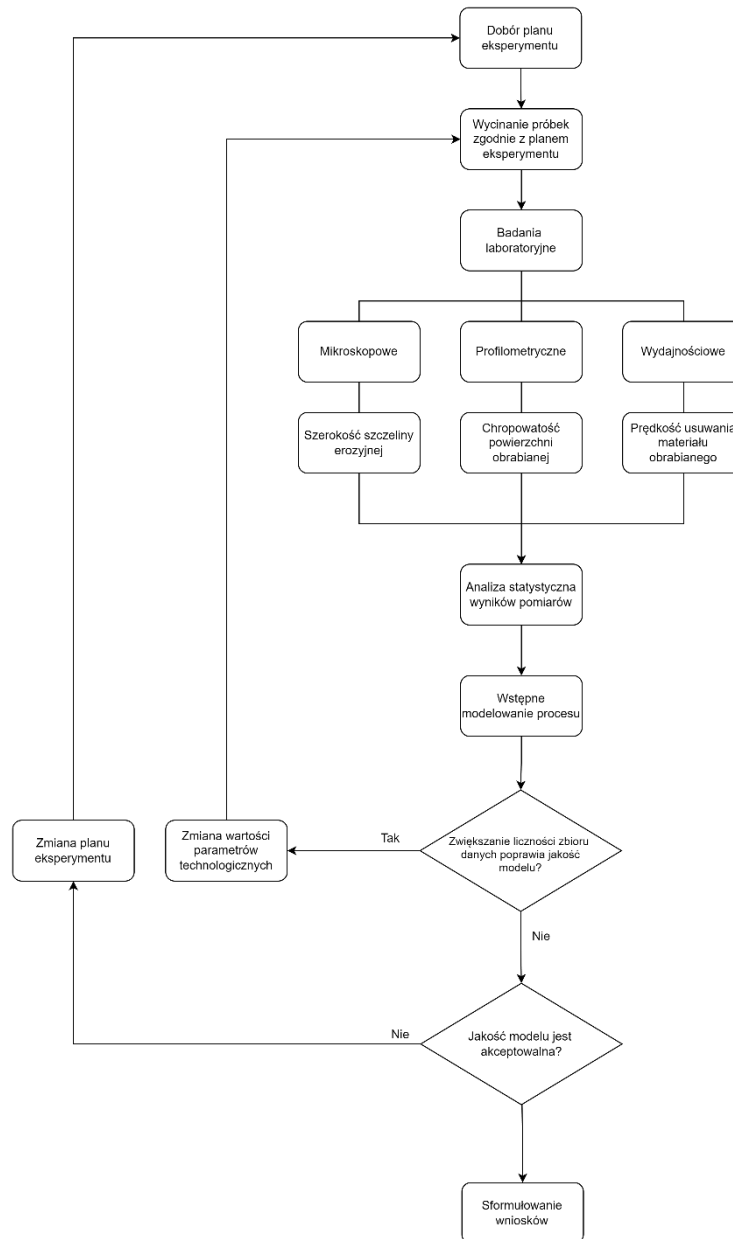
7.2.1. Plan badań

Na Rys. 7.5 przedstawiono ogólny schemat, prezentujący w skróconej formie poszczególne kroki realizowane w ramach badań eksperymentalnych. Pierwszym etapem był odpowiedni dobór planu eksperymentu oraz zakresów wartości zmiennych niezależnych (tutaj parametrów technologicznych maszyny), rozpatrywanych w ramach planu eksperymentu (zob. p. 7.2.2). Kolejnym i kluczowym elementem badań było wycinanie stalowych próbek (scharakteryzowanych w p. 7.1.2 i p. 7.1.3) za pomocą elektrodrażarki drutowej (p. 7.1.1) z różnymi kombinacjami parametrów technologicznych, zgodnie z wcześniej wybranym planem eksperymentu. Po wycięciu wszystkich przewidzianych próbek, przeprowadzono badania laboratoryjne. Badania te obejmowały pomiary profilometryczne chropowatości powierzchni próbek oraz pomiary mikroskopowe szerokości szczeliny erozyjnej. Dodatkowo, czas cięcia każdej z próbek był rejestrowany automatycznie przez obrabiarkę, co pozwoliło na obliczenie prędkości usuwania materiału. Istotnym aspektem procesu badawczego było również monitorowanie ciągłości obróbki, w tym obserwacja ewentualnego zerwania elektrody roboczej.

W tym miejscu istotne jest uszczegółowienie współbieżności fazy badań eksperymentalnych i fazy badań modelowych. W trakcie badań eksperymentalnych, wraz z gromadzeniem danych pomiarowych, konstruowano wstępne modele procesu. Modele te były na bieżąco poddawane ocenie pod względem dokładności przewidywania wartości rozpatrywanych zmiennych zależnych, przy użyciu metod opisanych w p. 6.3. W ramach oceny modeli sprawdzano również, jak zwiększenie zakresu danych pomiarowych wpływa na jakość modeli, wykorzystując w tym celu krzywe uczenia [94] (wyniki przedstawiono w rozdziale 9). W przypadku gdy jakość modeli wykazywała trend wzrostowy wraz ze zwiększaniem się zakresu danych pomiarowych, wówczas kontynuowano badania eksperymentalne bazując na wcześniej opracowanym planie eksperymentu, jednakże zmieniając zakresy wartości parametrów technologicznych. W przypadku gdy jakość modeli była nieakceptowalna i nie notowano wzrostu jakości wraz ze wzrostem zakresu danych, dokonywano zmiany planu eksperymentu i badania były powtarzane.

Cały proces miał zatem charakter iteracyjny – był kontynuowany do momentu, gdy zwiększanie zakresu danych nie wpływało na jakość modelu lub wpływało w stopniu minimalnym. Ponowienie serii badawczych było również rozpatrywane na etapie analizy statystycznej pomiarów, gdy analizy wyraźnie wskazywały na niską jakość uzyskanych wyników, a także podczas wycinania próbek, w przypadku problemów z ciągłością procesu (np. przez zerwania elektrody roboczej).

Po zakończeniu badań eksperymentalnych sformułowano wnioski, które posłużyły do opracowania docelowych modeli procesu



Rys. 7.5 Ogólny plan badań eksperymentalnych

7.2.2. Wybór zmiennych procesu

Jako zmienne niezależne przyjęto wszystkie parametry technologiczne elektrodrażarki drutowej Fanuc. Decyzja ta wynika z faktu, że parametry te są ustawiane przez technologa lub operatora podczas przygotowania procesu cięcia elektroerozyjnego w rzemieślniczym przedsiębiorstwie,

gdzie wdrażane są opracowane rozwiązania. Zestawienie tych parametrów zostało przedstawione w Tabeli 7.4.

Tabela 7.4 Zestawienie parametrów technologicznych przyjętych jako zmienne niezależne w badaniach

Symbol w sterowniku Fanuc	Jednostka	Opis
VS	V	Napięcie otwartej przerwy
CC	A	Natężenie prądu
VM	V	Napięcie wyładowania
ON	μ s	Czas trwania impulsu
OFF	μ s	Czas wyłączenia impulsu
SV	V	Napięcie serwomechanizmu - określa napięcie referencyjne dla ruchów serwomechanizmu w celu utrzymania stałego napięcia w przerwie między elektrodami.
SVG		Wzmocnienie sprzężenia zwrotnego - określa jak szybko układ jest w stanie reagować na zmiany napięcia wywołane przez zmiany odległości między elektrodą a obrabianym materiałem
WP1A		Parametry układu regulacji Fanuc
WP1B		
WP2A		
WP2B		
T	gf	Naciąg drutu
WF	m/min	Prędkość przewijania elektrody
FR	l/min	Natężenie przepływu dielektryka

Badania zostały przeprowadzone z wykorzystaniem stałych wartości warunków obróbki, które zestawiono w Tabeli 7.5.

Tabela 7.5 Stałe warunki obróbki w badaniach eksperymentalnych

Opis warunku	Wartość
Wysokość materiału obrabianego	25 [mm]
Średnica elektrody drutowej	0.2 [mm]
Oporność elektryczna dielektryka	8.1 [$M\Omega \cdot cm$]
Odległość górnej i dolnej głowicy od obrabianego materiału	0.1 [mm]

W odpowiedzi na cele pracy zdefiniowane w rozdziale 2, wybrano zmienne zależne rozpatrywane w badaniach, zaprezentowane w Tabeli 7.6.

Tabela 7.6 Zmienne zależne rozpatrywane w badaniach

Zmienna zależna	Jednostka	Opis
MRR	mm ³ /min	Prędkość usuwania materiału
Ra, Rq, Rz	μm	Parametry chropowatości powierzchni
Kw	-	Szerokość szczeliny erozyjnej
C	-	Ciągłość procesu cięcia

Jako zmienną określającą wydajność procesu obróbki przyjęto prędkość usuwania materiału (MRR), obliczaną według formuły:

$$MRR = F \cdot D_w \cdot H, \quad (7.1)$$

gdzie:

- F [mm/min] - prędkość posuwu głowicy maszyny,
- D_w [mm] - średnica drutu,
- H [mm] - grubość obrabianego materiału.

Prędkość posuwu (F) obliczano ze wzoru:

$$F = \frac{60l}{t}, \quad (7.2)$$

gdzie:

- l [mm] - długość cięcia,
- t [s] - czas cięcia.

W ramach badań jako kryterium oceny ciągłości procesu cięcia (C), przyjęto występowanie zdarzeń przerywających ten proces. Do zjawisk zakłócających ciągłość zaliczono: przerwanie drutu, alarm związany ze zwarcie elektrycznym oraz alarm sygnalizujący problemy z inicjacją wyładowania elektrycznego. W przypadku wystąpienia któregokolwiek z tych zdarzeń, zmienna opisująca ciągłość procesu (C) przyjmowała wartość 1, co oznaczało przerwanie ciągłości. Z kolei, w sytuacji, gdy proces przebiegał bez zakłóceń, zmienna ta przyjmowała wartość 0. Wartości zmiennej zależnej C przedstawiono w Tabeli 7.7.

Tabela 7.7 Wartości zmiennej zależnej określającej stabilność procesu cięcia

Ciągłość obróbki	Wartość zmiennej C
Zachowana	0
Przerwana	1

Dodatkowe badania, o charakterze uzupełniającym, obejmowały analizę szerokości szczeliny erozyjnej (Kw). Zmienna ta została rozpatrzona jako zmienna nominalna. Na podstawie analizy zgromadzonych danych pomiarowych wyróżniono trzy klasy szerokości szczeliny, odpowiadające wartościom minimalnym, średnim oraz maksymalnym. Szczegóły dotyczące klasyfikacji przedstawiono w Tabeli 7.8.

Tabela 7.8 Zakresy wartości klas szczeliny erozyjnej

Klasa szerokości szczeliny erozyjnej	Zakres wartości [mm]
1	0.2 - 0.22
2	0.23 - 0.257
3	0.258 - 0.3

Decyzja o traktowaniu szerokości szczeliny erozyjnej jako zmiennej nominalnej, a nie ciągłej, była uzasadniona zarówno względami ekonomicznymi, jak i praktycznymi aspektami wdrożeniowymi. Jak wykazały wstępne badania, opracowanie modelu do przewidywania konkretnych wartości szerokości szczeliny (Kw), który jednocześnie zachowywałby zdolność do generalizacji, wymagałoby zgromadzenia dużej liczby danych. Dla porównania, w badaniach nad chropowatością powierzchni oraz prędkością usuwania materiału (MRR) przeprowadzono kilkaset pomiarów (zob. Tabela 8.1). Tego rodzaju badania są kosztowne i czasochłonne. Ponadto, praktyczne zastosowanie przewidywanych wartości szerokości szczeliny erozyjnej w środowisku przemysłowym byłoby ograniczone.

Po pierwsze, w przedsiębiorstwie korekta szerokości szczeliny erozyjnej odbywa się głównie poprzez edycję „offsetu”, czyli wartości przesunięcia programowanej ścieżki elektrody roboczej względem nominalnego zarysu obrabianego elementu. W związku z tym kontrola wpływu parametrów technologicznych na szerokość szczeliny sprowadza się do utrzymania stałej wartości tej zmiennej.

Po drugie, w realiach produkcyjnych firmy SMHydro nie stosuje się precyzyjnych pomiarów geometrii mechanizmów satelitowych, ze względu na ich wysoki poziom złożoności. Uniemożliwia to dokładne ustalenie wymaganej szerokości szczeliny erozyjnej. W praktyce stosuje się korektę szerokości szczeliny na podstawie parametrów

technologicznych, obserwacji współpracy elementów mechanizmów satelitowych oraz doświadczenia operatorów.

Powyższe czynniki zdecydowały o ograniczeniu zakresu badań do liczby próbek wystarczającej do opracowania klasyfikatora szerokości szczeliny erozyjnej. Takie podejście spełnia wymagania przedsiębiorstwa, jednocześnie minimalizując koszty i czasochłonność badań.

7.2.3. Plan eksperymentu

Mimo licznych badań dotyczących teorii planowania eksperymentów, wciąż brakuje jednoznacznych wytycznych dotyczących doboru przestrzeni badawczej dla różnych procesów produkcyjnych [95]. W obliczu tej luki, przyjęto zróżnicowane plany eksperymentu, które poddano weryfikacji zgodnie z procedurą opisaną w punkcie 7.2.1. W dalszej części tego podrozdziału scharakteryzowano zastosowane plany eksperymentu, przedstawiając w skróconej formie matryce eksperymentalne, z wartościami zmiennych niezależnych wyrażonymi w skali znormalizowanej. Zakres wartości zmiennych niezależnych w skali znormalizowanej przedstawiono w Tabeli 7.9.

Tabela 7.9 Wartości zmiennych niezależnych w skali znormalizowanej

Wartość zmiennej niezależnej	Skala znormalizowana
Minimalna	-1
Średnia	0
Maksymalna	1

Mając na uwadze liczną grupę zmiennych niezależnych (łącznie 14), które podlegały analizie w ramach badań, kryterium wyboru odpowiednich planów eksperymentalnych oparto na minimalizacji ilości obserwacji wymaganych do efektywnego przeprowadzenia eksperymentu w ramach wybranego planu. Badania rozpoczęto zatem od zastosowania dwupoziomowego planu selekcyjnego [96]. Plan ten umożliwia skuteczne przetestowanie wielu czynników przy minimalnej liczbie obserwacji, co pozwala na ograniczenie kosztów badań. Plany selekcyjne dwupoziomowe oznacza się symbolem PS/DS 2^{n-k} , gdzie: n – liczba zmiennych niezależnych, k - krotność połowienia planu kompletnego PS/DS 2^n , z którego utworzono dany plan selekcyjny [96][97]. Zdecydowano się na sześciokrotne połowienie planu kompletnego, uzyskując matrycę planu eksperymentu z 256 kombinacjami parametrów, przedstawioną w skróconej formie w Tabeli 7.10.

Tabela 7.10 Skrócona matryca planu PS/DS 2^8 dla 14 zmiennych niezależnych

Lp.	VS	CC	VM	ON	OFF	SV	SVG	WP1A	WP1B	WP2A	WP2B	T	WF	FR
0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0
1	1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	1.0	1.0	1.0	1.0	1.0	1.0
2	-1.0	1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	1.0	1.0	1.0	1.0	1.0	1.0
...
253	1.0	-1.0	1.0	1.0	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0
254	-1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0
255	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Analiza wyników przeprowadzonych badań według dwupoziomowego selekcyjnego planu eksperymentu wykazała jednak słabą zmienność pomiarów chropowatości pomiędzy obserwacjami. Jakość modelu przewidującego chropowatość powierzchni, opracowanego na podstawie uzyskanych danych, okazała się nieakceptowalna.

W związku z tym w kolejnym kroku zdecydowano się użyć planu Placketta-Burmana [98], który podobnie jak PS/DS 2^{n-k} pozwala na analizę wpływu wielu czynników, przy jednoczesnym ograniczeniu liczby potrzebnych obserwacji. Zdecydowano się również na rozszerzenie zakresu wartości parametrów, w celu zwiększenia prawdopodobieństwa uzyskania różnych wyników pomiarów zmiennych zależnych.

Plan Placketta-Burmana umożliwia badanie wpływu $4c - 1$ badanych parametrów na zmienną zależną przy realizacji $N = 4c$ prób, gdzie: c – liczba całkowita. Jeśli liczba parametrów jest mniejsza niż $4c - 1$, to w matrycy planowania ujmuję się tzw. czynniki fikcyjne [97],[98]. W rozważanym przypadku, przy 14 parametrach należy dodać jeden czynnik fikcyjny, planując tym samym 15 kombinacji parametrów. Opracowana w ten sposób matryca eksperymentu została zaprezentowana w skróconej formie w Tabeli 7.11.

Tabela 7.11 Skrócona matryca planu eksperymentu Placketta-Burmana dla 14 parametrów technologicznych

Lp.	VS	CC	VM	ON	OFF	SV	SVG	WP1A	WP1B	WP2A	WP2B	T	WF	FR
0	-1.0	-1.0	1.0	-1.0	1.0	1.0	-1.0	-1.0	1.0	1.0	-1.0	1.0	-1.0	-1.0
1	1.0	-1.0	-1.0	-1.0	-1.0	1.0	1.0	-1.0	-1.0	1.0	1.0	1.0	1.0	-1.0
2	-1.0	1.0	-1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	1.0	-1.0	1.0
...
12	-1.0	-1.0	1.0	1.0	-1.0	-1.0	1.0	1.0	-1.0	-1.0	1.0	1.0	-1.0	-1.0
13	1.0	-1.0	-1.0	1.0	1.0	-1.0	-1.0	1.0	1.0	-1.0	-1.0	1.0	1.0	-1.0
14	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0

Niestety, zastosowane rozwiązanie spowodowało problemy z ciągłością procesu (zerwania elektrody drutowej) ze względu na większy rozrzut wartości parametrów technologicznych co wymusiło wielokrotne powtórzenie eksperymentu i zmiany wartości

parametrów. Badania bazujące na tym planie eksperymentu zakończono po pięciu próbach, gromadząc 75 obserwacji.

Ostateczne wyniki badań okazały się jednak wstępnie zadowalające. Uzyskano oczekiwaną zmienność wyników pomiarów oraz udało się wykazać istotny statystycznie wpływ poszczególnych parametrów na badane zmienne zależne i tym samym dokonać wstępnej selekcji parametrów.

Po usunięciu z badań parametrów o mniejszej istotności w kontekście wpływu na zmienną zależną, opracowano nowy plan eksperymentu, bazując na planie Boxa-Behnkena (BB) [96]. Plan ten uwzględnia parametry na trzech poziomach, co pozwoliło na rozszerzenie badanego zakresu i poprawę jakości modeli procesu.

Liczba obserwacji (N) wymaganych do opracowania planu BB jest zdefiniowana przez formułę [99]:

$$N = 2k(k - 1) + C_0, \quad (7.3)$$

gdzie:

- k - liczba zmiennych niezależnych,
- C_0 - liczba punktów centralnych

W przeprowadzonych badaniach przyjęto 8 parametrów oraz jeden punkt centralny, co dało 113 obserwacji. Skrócona matryca omawianego planu została zaprezentowana w tabeli Tabela 7.12.

Tabela 7.12 Skrócona matryca planu eksperymentu Boxa-Behnkena dla 8 zmiennych niezależnych

Lp.	VS	CC	VM	ON	OFF	SV	SVG	T
0	-1.0	-1.0	0.0	0.0	0.0	0.0	0.0	0.0
1	1.0	-1.0	0.0	0.0	0.0	0.0	0.0	0.0
2	-1.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0
...
110	0.0	0.0	0.0	0.0	0.0	0.0	-1.0	1.0
111	0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0
112	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

W kolejnym kroku dokonano dalszej selekcji zmiennych ze względu na ich istotność wpływu na badane zmienne zależne. Następnie powtórzono 8-krotnie badanie z ograniczonym zakresem zmiennych, według planu BB, w celu dalszego zwiększania jakości modeli. Zgromadzono w ten sposób 245 kolejnych obserwacji. Ostatecznie przeprowadzono badania eksperymentalne, które obejmowały łącznie 731 obserwacji.

8. ANALIZA STATYSTYCZNA WYNIKÓW POMIARÓW

Celem niniejszego rozdziału jest przeprowadzenie wstępnej analizy statystycznej wyników pomiarów uzyskanych w ramach badań eksperymentalnych, w celu identyfikacji najważniejszych właściwości zbioru danych. Analiza rozpoczyna się od ogólnej charakterystyki zbioru danych (p.8.1), w której przedstawiono podstawowe statystyki opisowe, takie jak średnie wartości, odchylenia standardowe, mediany, kwartyle oraz minimalne i maksymalne wartości zmiennych zależnych. Zestawienie tych statystyk umożliwia wstępne rozpoznanie struktury danych oraz ich zmienności.

W kolejnej części (p.8.2) przeprowadzono analizę obserwacji odstających, które mogą wpływać na wyniki pomiarów i ich interpretację. W końcowej sekcji (p.8.3) zaprezentowano wyniki analizy korelacji, mającej na celu identyfikację liniowych zależności pomiędzy zmiennymi, co stanowi podstawę do dalszych badań nad wpływem zmiennych niezależnych na zmienne zależne.

8.1. Ogólna charakterystyka zbioru danych

W Tabeli 8.1 zaprezentowano syntetyczne zestawienie podstawowych statystyk opisowych zgromadzonych wyników pomiarów.

Tabela 8.1 Podstawowe statystyki opisowe wyników pomiarów

	MRR [mm ³ /min]	Ra [μm]	Rq [μm]	Rz [μm]	Kw [mm]
Liczba obserwacji	685	563	563	563	80
Średnia arytmetyczna	7.37	3.51	4.37	21.82	0.242
Odchylenie standardowe	2.456	0.75	0.92	4.27	0.025
Minimum	1.09	1.99	2.43	12.12	0.2
Dolny kwartył	5.14	2.87	3.61	18.48	0.221
Mediana	7.25	3.47	4.34	21.82	0.247
Górny kwartył	9.71	4.1	5.11	25.06	0.258
Maksimum	13.88	5.68	6.76	33.12	0.3

W Tabeli 8.2 przedstawiono liczbę obserwacji dla poszczególnych wartości dotyczących ciągłości obróbki.

Tabela 8.2 Liczba obserwacji dotyczących ciągłości obróbki

Ciągłość obróbki (C)	zachowana	przerwana
Liczba obserwacji	685	46

Tabela 8.1 przedstawia istotne rozbieżności w liczbie dostępnych obserwacji dla zmiennych zależnych, co wynika głównie z trudności w utrzymaniu ciągłości procesu cięcia elektroerozyjnego podczas eksperymentów.

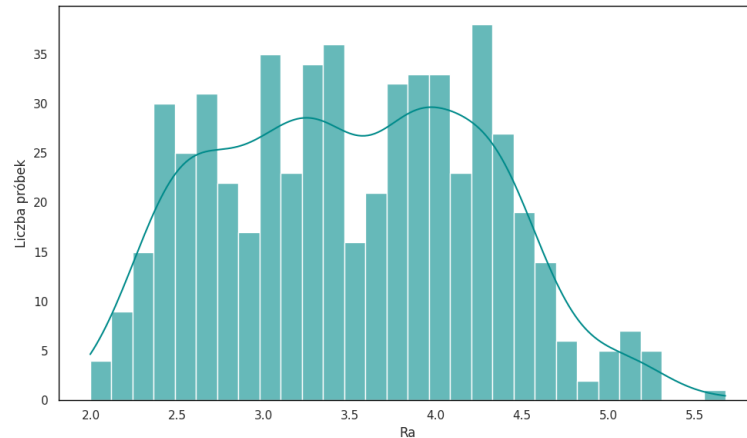
W przypadku przerwania procesu wycinania próbki, rejestracja czasu obróbki przez maszynę była utrudniona z uwagi na konieczność wprowadzenia zmian w parametrach technologicznych, co miało na celu wznowienie procesu. Tego rodzaju interwencje uniemożliwiały obliczenie prędkości usuwania materiału (MRR), gdyż warunki obróbki ulegały zmianie, co prowadziło do niejednorodności danych. Z tego samego powodu nie było możliwe przeprowadzenie pomiaru chropowatości powierzchni. Procentowy udział brakujących danych dla poszczególnych zmiennych zależnych, wynikających z opisanych powyżej przyczyn, został przedstawiony w Tabeli 8.3.

Tabela 8.3 Procentowy udział brakujących danych dla poszczególnych zmiennych zależnych

Zmienna zależna	Ra	MRR	Kw	C
Procent brakujących danych	8.6%	6.2%	0%	0%

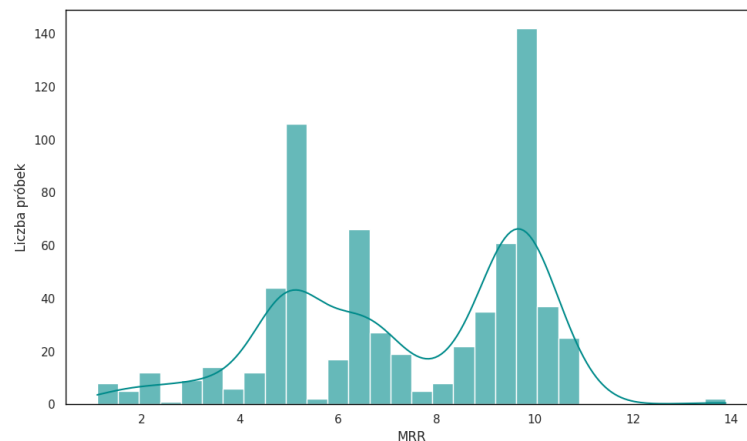
Przy pomiarze szerokości szczeliny erozyjnej (Kw) nie wykonywano pomiarów chropowatości z powodu geometrii próbek, która uniemożliwiała ich przeprowadzenie. Rejestrowano jednak czas obróbki. W rezultacie liczba dostępnych obserwacji dla zmiennych MRR i Ra jest zróżnicowana.

Poniżej przedstawiono serię histogramów dla rozpatrywanych ciągłych zmiennych zależnych, które ilustrują rozkład zgromadzonych danych. Ze względu na bardzo duże podobieństwo rozkładów wartości parametrów chropowatości Ra, Rq i Rz, zdecydowano się na zaprezentowanie tylko rozkładu zmiennej Ra.



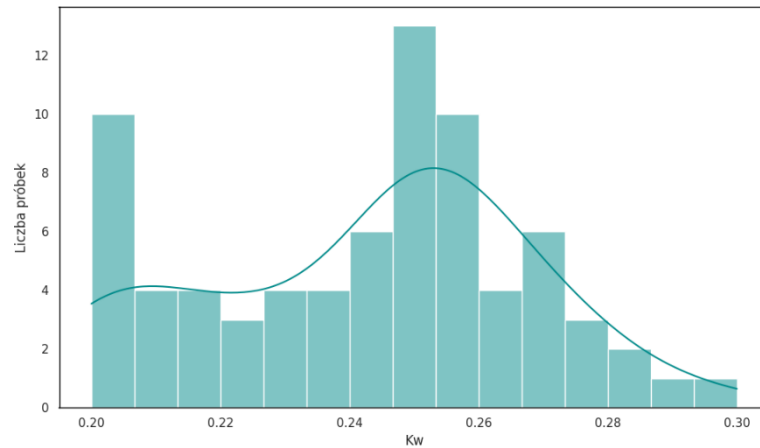
Rys. 8.1 Histogram wartości parametru chropowatości Ra

Rozkład wartości zmiennej zależnej Ra (Rys. 8.1) wykazuje lekką prawostronną asymetrię i wielomodalność, z dwoma lokalnymi maksimumami przy 3,5 μm i 4,5 μm . Zakres Ra wynosi od 2 do 5,5 μm , z zauważalnym spadkiem liczby próbek powyżej 4,5 μm . Rozkład wskazuje na relatywnie dużą zmienność danych i tendencję do wyższych wartości chropowatości.



Rys. 8.2 Histogram wartości wskaźnika prędkości usuwania materiału MRR

Zakres wartości MRR (Rys. 8.2) wynosi od około 1 do 14 mm^3/min . Histogram wskazuje na trzy główne piki. Pierwszy istotny pik występuje przy około 4 jednostkach, drugi w przedziale 7-8 mm^3/min , a największy przy 10 mm^3/min . Wartości powyżej 12 mm^3/min są marginalne. Rozkład jest wielomodalny, co może wskazywać na istnienie kilku podgrup w danych.



Rys. 8.3 Histogram wartości szerokości szczeliny erozyjnej

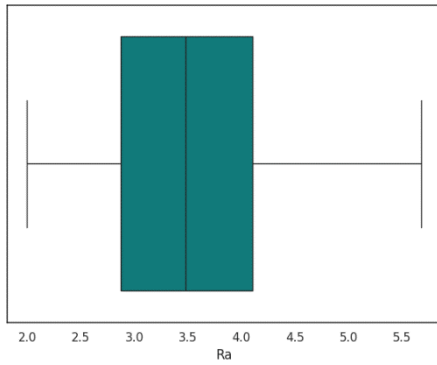
Zakres wartości zmiennej Kw (Rys. 8.3) wynosi od 0,20 do 0,3 mm. Histogram wskazuje na dwa wyraźne piki: pierwszy, mniejszy, przy 0,20 mm oraz drugi, wyraźniejszy, w przedziale 0,24–0,26 mm, z maksimum przy 0,26 mm. Wartości powyżej 0,26 mm są mniej liczne, a powyżej 0,28 mm występują sporadycznie. Rozkład danych jest lekko prawoskośny, z większą liczbą próbek o niższych wartościach Kw.

8.2. Analiza pomiarów pod względem obserwacji odstających

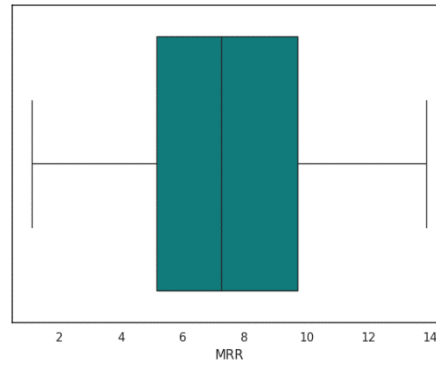
W procesie badawczym ważnym elementem jest identyfikacja i analiza wartości odstających, które mogą wskazywać na potencjalne błędne pomiary lub zjawiska mogące mieć znaczenie dla wyników badań. W kontekście przeprowadzonych pomiarów, wykresy pudełkowe (Rys. 8.4 -

Rys. 8.6) prezentują dystrybucję wartości badanych zmiennych dając możliwość oceny obecności obserwacji odstających.

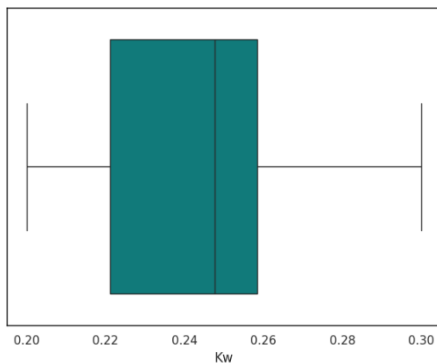
Na podstawie wizualnej oceny wykresów można stwierdzić, że w przypadku wszystkich zmiennych zależnych, nie zaobserwowano obserwacji leżących poza granicami określonymi przez „wąsy” wykresu, co sugeruje brak wyraźnych wartości odstających.



Rys. 8.4 Wykres pudełkowy wartości parametru chropowatości Ra



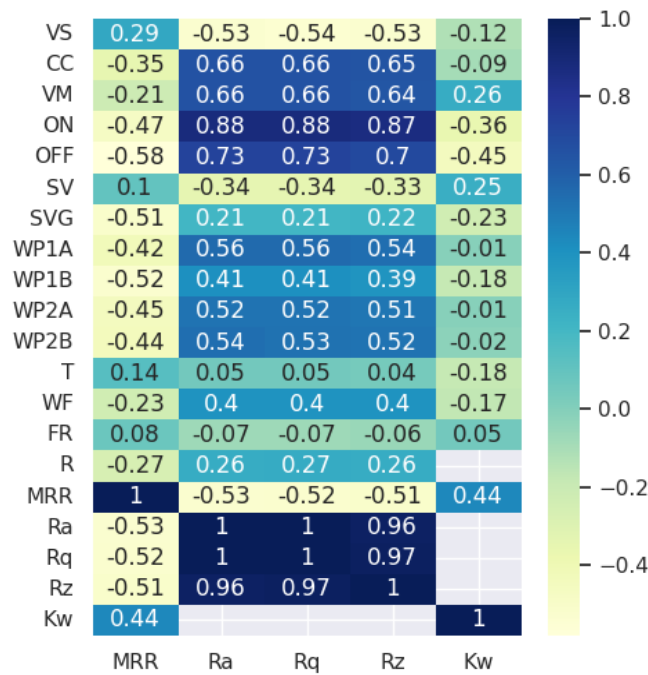
Rys. 8.5 Wykres pudełkowy wartości prędkości usuwania materiału MRR



Rys. 8.6 Wykres pudełkowy wartości szerokości szczeliny erozyjnej Kw

8.3. Wyniki analizy korelacji

W celu określenia zależności liniowych pomiędzy zmiennymi procesu przeprowadzono analizę korelacji z zastosowaniem współczynnika korelacji liniowej Pearsona (5.1) dla wszystkich możliwych par badanych zmiennych. Wyniki przedstawiono w formie macierzy korelacji (Rys. 8.7).



Rys. 8.7 Korelacja ciągłych zmiennych niezależnych i zależnych

Na podstawie przeprowadzonych obliczeń sformułowano następujące wnioski:

- Występuje niemal pełna korelacja pomiędzy parametrami chropowatości (Ra, Rq, Rz), co jest zgodne z oczekiwaniami, ponieważ wszystkie te zmienne opisują właściwości chropowatości powierzchni. Z tego powodu w dalszej analizie zostanie uwzględniony jedynie parametr Ra jako reprezentatywny.
- Parametry chropowatości są najsilniej skorelowane z parametrem ON (0.88), następnie z OFF (0.73) oraz z VM i CC (0.66).
- Na podstawie tych wyników można stwierdzić, że chropowatość powierzchni wykazuje silną korelację liniową z parametrami elektrycznymi procesu obróbki. Nie wykazano jednak istotnej zależności liniowej pomiędzy chropowatością powierzchni a analizowanymi parametrami hydromechanicznymi, takimi jak T, WF, czy FR.
- Wyniki wskazują, że prędkość usuwania materiału (MRR) nie wykazuje istotnej korelacji liniowej z żadnym z rozpatrywanych parametrów, co może sugerować bardziej złożoną zależność w tym zakresie, wymagającą dalszej analizy.
- Podobnie jak w przypadku MRR, szerokość szczeliny erozyjnej (Kw) również nie jest wyraźnie skorelowana liniowo z żadnym z badanych parametrów.

9. BADANIA MODELOWE PROCESU WEDM

Niniejszy rozdział przedstawia wyniki badań modelowych procesu cięcia elektroerozyjnego, z wykorzystaniem wybranych modeli statystycznych oraz uczenia maszynowego. W ramach tego rozdziału zaprezentowane zostaną cztery modele przewidujące wartości (lub klasy) rozpatrywanych zmiennych zależnych: chropowatość powierzchni (R_a), prędkość usuwania materiału (MRR), szerokość szczeliny erozyjnej (K_w) oraz ciągłość procesu obróbki (C).

Badania modelowe prowadzono według powtarzalnej procedury (p. 9.1 - 9.4). W pierwszym kroku przeprowadzono wstępne przygotowanie danych, które obejmowało standaryzację Z zmiennych niezależnych oraz uzupełnianie brakujących wartości zmiennych zależnych metodą K-NN. W przypadku klasyfikacji dodatkowo zrealizowano balansowanie licznosci klas zmiennej zależnej, wykorzystując metodę oversamplingu klasy mniejszościowej.

Następnie, przygotowane dane podzielono na zbiór treningowy i testowy (zob. p.6.2.1). Dane treningowe wykorzystano do trenowania modeli z użyciem walidacji krzyżowej. Jednocześnie przeprowadzono analizę porównawczą w celu wyłonienia modelu o najwyższych wskaźnikach jakości, który następnie zakwalifikowano do dalszych badań.

Po wyborze optymalnego modelu przeprowadzono dodatkową analizę wpływu metod przygotowania danych na jakość modelu, uwzględniając specyficzne wymagania wynikające z wybranego modelu. Wyniki tej analizy umożliwiły wybór docelowych metod przetwarzania danych. Następnie model został ponownie wytrenowany na danych przetworzonych zgodnie z wybranymi metodami, co miało na celu maksymalizację jego jakości.

W kolejnym kroku przeprowadzono szczegółową analizę jakości modelu, wykorzystując wcześniej wydzielone dane testowe. W przypadku modeli regresyjnych dokonano analizy rozrzutu wartości rzeczywistych (testowych) i przewidywanych, analizy statystycznej reszt oraz obliczono wskaźniki jakości modelu na zbiorze testowym. W przypadku modeli klasyfikacyjnych przeprowadzono analizę jakości klasyfikacji, stosując metody opisane w p.6.3.2. W każdym przypadku dodatkowo zbadano wpływ liczby obserwacji treningowych na wybrany wskaźnik jakości modelu.

W dalszej części rozdziału zaprezentowano analizę wpływu parametrów zmiennych niezależnych na badane zmienne zależne z wykorzystaniem zidentyfikowanych modeli (p.0). W analizie zastosowano metodę interpretacji modeli SHAP (zob. p.6.4) oraz trójwymiarowe wizualizacje wyników.

Na zakończenie rozdziału przedstawiono podsumowanie badań modelowych (p.0).

Dla uproszczenia opisu w tym rozdziale, zastosowano konwencję nazewnictwa modeli, polegającą na użyciu terminu „model” oraz symbolu danej zmiennej zależnej (zob. Tabela 7.6), której wartości model przewiduje. Zatem np. model przewidujący wartość parametru chropowości Ra, jest określany terminem „model Ra”.

Zastosowano również oznaczenia modeli odpowiednim skrótem, co ułatwia odniesienie się do nich w dalszej części pracy (Tabela 9.1).

Tabela 9.1 Oznaczenia modeli zastosowane w rozdziale

Model	Skrót
Sztuczne sieci neuronowe	ANN
Regresja liniowa	LR
Regresja wielomianowa	PR
Maszyny wektorów nośnych	SVR (regresja), SVC (klasyfikacja)
XGBoost	XGB
Las losowy	RF
Drzewo decyzyjne	DT
Regresja logistyczna	LGR

W analizach porównawczych, w których wykorzystano wskaźniki jakości modeli, zastosowano skalę kolorów w celu wyróżnienia poszczególnych wartości tych wskaźników. Kolory te reprezentują stopniowanie od najmniej korzystnych wartości (w odcieniu jasno szarym), przez wartości mediany (w odcieniu szarym), do najbardziej korzystnych wartości (w odcieniach zieleni), co ułatwia interpretację wyników porównawczych.

Tabela 9.2 Legenda skali kolorystycznej wskaźników jakości

Najkorzystniejsza wartość
Mediana
Najmniej korzystna wartość

9.1. Model Ra

9.1.1. Dobór metody modelowania

Wyniki analizy porównawczej przeprowadzonej z użyciem metod omówionych w p.6.3 przedstawiono w Tabela 9.3.

Tabela 9.3 Porównanie opracowanych modeli Ra

Model	\overline{R}_{cv}^2	σR_{cv}^2	\overline{MAE}_{cv}	σMAE_{cv}	\overline{MSE}_{cv}	σMSE_{cv}	\overline{MAPE}_{cv}	$\sigma MAPE_{cv}$
ANN	0.849	0.036	0.222	0.033	0.082	0.012	0.064	0.007
LR	0.873	0.03	0.168	0.013	0.067	0.015	0.049	0.003
PR	0.741	0.081	0.232	0.022	0.134	0.038	0.069	0.007
SVR	0.879	0.029	0.158	0.01	0.064	0.015	0.046	0.003
XGB	0.866	0.027	0.188	0.01	0.07	0.013	0.055	0.002
RF	0.86	0.035	0.197	0.013	0.073	0.017	0.058	0.003
DT	0.822	0.026	0.227	0.012	0.093	0.011	0.066	0.004

Powyższe wyniki można podsumować następująco:

- Najkorzystniejsze wartości wszystkich ocenianych wskaźników (\overline{R}_{cv}^2 , \overline{MAE}_{cv} , \overline{MSE}_{cv} , \overline{MAPE}_{cv}) uzyskano przy zastosowaniu SVR
- Korzystne wyniki uzyskano również przy zastosowaniu LR, które były nieznacznie niższe niż w przypadku SVR.
- Najmniej korzystne wyniki uzyskano przy zastosowaniu PR.

W procesie poszukiwania optymalnych hiperparametrów SVR zbadano m.in. różne rodzaje funkcji jądrowych. Ostatecznie jako najlepsza okazała się funkcja liniowa. Z tego względu podobieństwa w dokładności przewidywania wartości zmiennej zależnej pomiędzy SVR a LR nie są zaskakujące. W przeprowadzonej analizie statystycznej (rozdział 8) wykazano istotne korelacje liniowe pomiędzy zmiennymi niezależnymi, w szczególności parametrem ON, a zmienną zależną Ra. Modele takie jak regresja liniowa i SVR z liniowym jądrem są na ogół skuteczne w modelowaniu tego typu zależności, co potwierdziła przeprowadzona analiza porównawcza. Nieznacznie lepsze wskaźniki osiągnięte przez SVR można przypisać jego mniejszej wrażliwości na szum w danych.

W związku z tym, na podstawie przeprowadzonej analizy, zdecydowano się na zastosowanie modelu SVR do przewidywania wartości Ra.

W Tabeli 9.4 zestawiono wartości hiperparametrów zidentyfikowanego modelu SVR dostosowane w procesie walidacji krzyżowej.

Tabela 9.4 Hiperparametry zidentyfikowanego modelu SVR

Parametr	C	cache_size	coef0	degree	epsilon	gamma	kernel	max_iter	shrinking	tol
Wartość	0.1	200	0.0	3	0.001	scale	linear	1000000	True	0.001

Dzięki zastosowaniu liniowej funkcji jądrowej model SVR można przedstawić jako łatwo interpretowalny model liniowy. Poniżej zamieszczono jego reprezentację:

$$\begin{aligned}
 Ra = & 3.496 + 0.049 \cdot VS - 0.003 \cdot CC + 0.217 \cdot VM + 0.487 \cdot ON \\
 & + 0.092 \cdot OFF - 0.042 \cdot SV + 0.037 \cdot SVG - 0.005 \\
 & \cdot WP1A - 0.043 \cdot WP1B + 0.046 \cdot WP2A - 0.005 \\
 & \cdot WP2B - 0.093 \cdot T + 0.011 \cdot WF + 0.005 \cdot FR
 \end{aligned} \tag{9.1}$$

9.1.2. Przygotowanie danych treningowych

W Tabeli 9.5 przedstawiono ocenę metod postępowania z brakującymi wartościami zmiennej zależnej R_a na jakości modelu R_a , ocenianą na podstawie średniej wartości współczynnika determinacji w walidacji krzyżowej \bar{R}_{cv}^2 oraz odchylenia standardowego $\sigma \bar{R}_{cv}^2$.

Tabela 9.5 Wpływ metody postępowania z brakującymi wartościami na jakość modelu R_a

Metoda postępowania z brakującymi wartościami	\bar{R}_{cv}^2	$\sigma \bar{R}_{cv}^2$
Usunięcie obserwacji z brakującymi wartościami	0.587	0.169
Imputacja brakujących wartości medianą	0.675	0.062
Imputacja brakujących wartości średnią arytmetyczną	0.674	0.059
Imputacja metodą K-najbliższych sąsiadów (k=2)	0.826	0.036
Imputacja wielokrotna za pomocą regresji liniowej (MICE)	0.879	0.029
Imputacja wielokrotna za pomocą drzewa decyzyjnego (MICE)	0.699	0.078
Imputacja wielokrotna za pomocą regresji Bayesowskiej (MICE)	0.877	0.030

Biorąc pod uwagę, że model SVR z liniową funkcją jądrową jest w tym przypadku najbardziej odpowiedni do przewidywania wartości R_a , można przypuszczać, że metody imputacji oparte na regresji liniowej będą skuteczne w przewidywaniu brakujących wartości R_a . Jak wynika z powyższego zestawienia, wysokie wartości współczynnika determinacji oraz niskie odchylenia standardowe dla metod imputacji wielokrotnej za pomocą regresji liniowej oraz regresji Bayesowskiej potwierdzają to przypuszczenie.

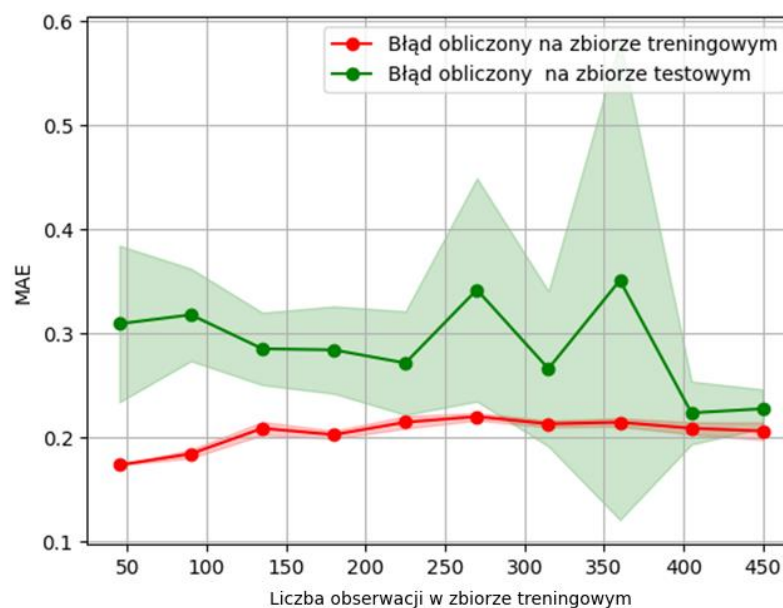
W związku z tym jako metodę uzupełniania brakujących wartości pomiarów wybrano metodę imputacji wielokrotnej za pomocą regresji liniowej (MICE).

Kolejnym etapem przygotowania danych treningowych była ich odpowiednia transformacja. Przeprowadzone badania wykazały, że metody takie jak standaryzacja Z, normalizacja min-max oraz normalizacja według maksymalnej wartości zapewniają najwyższą jakość modelu. W związku z tym zdecydowano się zastosować powszechnie używaną metodę standaryzacji Z.

9.1.3. Analiza jakości modelu

Na załączonym wykresie (Rys. 9.1) przedstawiono analizę średniego błędu absolutnego (MAE) modelu w zależności od liczby obserwacji użytych do treningu i testowania. Na wykresie przedstawiono dwie krzywe: czerwoną, która reprezentuje błąd obliczony na zbiorze treningowym, oraz zieloną, która pokazuje błąd obliczony na zbiorze testowym. Obie krzywe zostały wyliczone z wyników walidacji krzyżowej. Widoczne obszary zacienione wokół krzywych reprezentują odchylenie standardowe wyników, które zostało obliczone na podstawie rozproszenia wyników MAE w ramach tej walidacji, zarówno dla zbioru treningowego, jak i testowego.

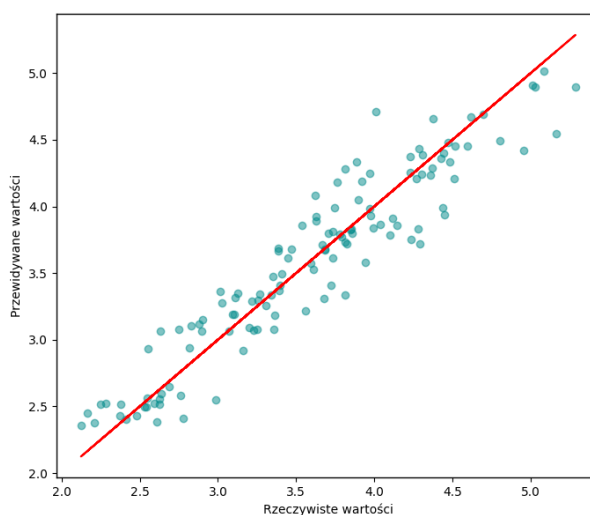
W celu przeprowadzenia tej analizy, zbiór danych został podzielony na pięć części, z których każda pełniła kolejno rolę zbioru testowego, podczas gdy pozostałe były używane do treningu modelu. Rozmiar zbioru treningowego był stopniowo zwiększany od 10% do 100%, przy jednoczesnym utrzymaniu stałej liczby danych testowych. Dla każdego etapu oceniano jakość modelu, obliczając błędy MAE zarówno dla zbioru treningowego, jak i testowego.



Rys. 9.1 Krzywa uczenia modelu

Stabilność wielkości błędu wyznaczonego dla zbioru treningowego (czerwona krzywa) wskazuje na zdolność modelu do adekwatnego dopasowania się do danych, niezależnie od liczby obserwacji. Błąd obliczony dla zbioru testowego (zielona krzywa) wykazuje trend malejący wraz ze wzrostem liczby danych. W przedziale około 250-350 obserwacji, można jednak zauważyć wyraźne wahania i wzrost wielkości błędu wyznaczonego dla zbioru testowego. Te skoki mogą być rezultatem kilku czynników, takich jak nadmierne dopasowanie modelu do danych treningowych lub obecność szumu w danych testowych. Powyżej 400 obserwacji treningowych wielkość błędu uzyskana przy użyciu zbioru testowego ulega jednak stabilizacji i staje się jedynie minimalnie większa niż wielkość błędu uzyskanego przy użyciu zbioru treningowego, co sugeruje, że model lepiej generalizuje, gdy dostępna jest większa liczba danych.

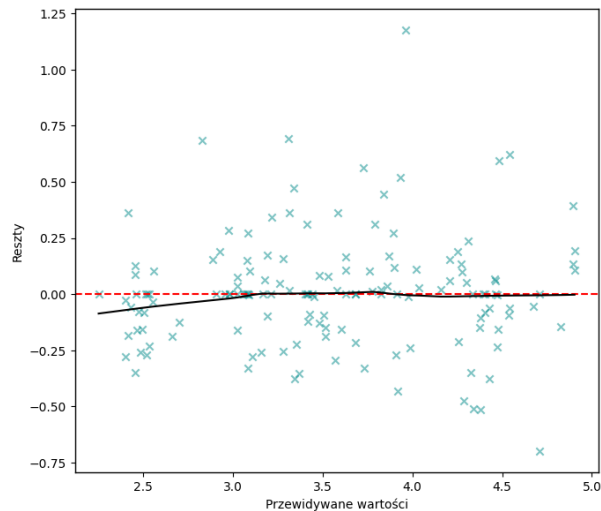
Na Rys. 9.2 przedstawiono wykres rozrzutu pomiędzy rzeczywistymi wartościami R_a ze zbioru testowego, a przewidywanymi wartościami R_a , uzyskanymi za pomocą zidentyfikowanego modelu.



Rys. 9.2 Wykres rozrzutu wartości przewidywanych i rzeczywistych wartości R_a

Punkty na wykresie zdają się układać wzdłuż linii idealnej, co sugeruje prawidłowe dopasowanie modelu do rzeczywistych wartości R_a . Potwierdza to obliczony współczynnik korelacji liniowej Pearsona dla wartości przewidywanych i wartości rzeczywistych, wynoszący 0.943. Rozproszenie punktów staje się większe przy wyższych wartościach R_a (około 4.0 do 5.0), co sugeruje, że model może być mniej dokładny przy takich wartościach. Widać, że przy niższych wartościach R_a (około 2.0 do 3.5), punkty są bliżej linii idealnej, co wskazuje na większą jakość modelu w tym zakresie.

W celu dalszej analizy na Rys. 9.3 zaprezentowano wykres reszt analizowanego modelu.



Rys. 9.3 Reszty modelu w stosunku do przewidywanych wartości Ra

Na wykresie można zauważyć, że reszty są rozproszone w sposób względnie losowy wokół linii poziomej na poziomie zero, co sugeruje, że model nie wykazuje wyraźnych tendencji do systematycznych błędów. Krzywa Lowess (6.17) jest bliska linii zero i ma stosunkowo płaski przebieg na większości zakresu przewidywanych wartości, co dodatkowo potwierdza brak systematycznych błędów w modelu. Nie zaobserwowano również wzorców sugerujących heteroskedastyczność.

W celu pogłębienia analizy rozkładu reszt obliczono statystyki rozkładu przedstawione w Tabeli 9.6

Tabela 9.6 Statystyki rozkładu reszt modelu

Kurtoza	Skośność	DW	JB	p-wartość dla testu JB
3.122	0.879	2.04	0.326	0.849

Analiza statystyczna reszt ujawnia kilka istotnych cech:

- Wartość kurtozy wynosząca 3.122 wskazuje na niewielkie odchylenie [100] od normalności w kierunku leptokurtycznego rozkładu.
- Skośność rozkładu reszt wynosząca 0.879 wskazuje na umiarkowaną dodatnią asymetrię [100].
- Statystyka Durbina-Watsona (DW) wynosząca 2.04 wskazuje na brak autokorelacji w resztach.

- Wartość statystyki Jarque-Bera (JB) wynosząca 0.326 oraz p-wartość równa 0.849 wskazują, że nie ma istotnych odchyień od rozkładu normalnego w rozkładzie reszt.

Analiza rozrzutu wartości przewidywanych i rzeczywistych Ra oraz rozkładu reszt wskazuje na poprawną konstrukcję modelu. Niemniej jednak zaobserwowano zwiększenie błędu przewidywania dla wyższych wartości Ra. Niedokładności te prawdopodobnie wynikają z relatywnie niewielkiej liczby obserwacji wysokich wartości Ra w zbiorze danych treningowych, co jest widoczne na Rys. 8.1. Problemy te nie wpływają jednak w sposób znaczący na ogólną jakość modelu, co potwierdzają wskaźniki przedstawione w Tabeli 9.7, prezentujące ocenę jakości modelu ustalone na zbiorze testowym.

Tabela 9.7 Ocena jakości modelu na zbiorze testowym

R^2_{test}	MAE_{test}	MSE_{test}	$MAPE_{test}$
0.888	0.170	0.063	0.048

Podsumowując, model może zostać przyjęty do dalszych badań ze względu na relatywnie niewielkie odchylenia od normalności oraz dobre statystyki oceny jakości.

9.2. Model MRR

9.2.1. Dobór metody modelowania

W Tabeli 9.8 przedstawiono wyniki analizy porównawczej modeli MRR.

Tabela 9.8 Porównanie opracowanych modeli MRR

Model	$\overline{R^2}_{cv}$	σR^2_{cv}	\overline{MAE}_{cv}	σMAE_{cv}	\overline{MSE}_{cv}	σMSE_{cv}	\overline{MAPE}_{cv}	$\sigma MAPE_{cv}$
ANN	0.747	0.066	0.703	0.069	1.474	0.34	0.139	0.028
LR	0.627	0.023	1.087	0.054	2.253	0.31	0.233	0.045
PR	0.213	0.412	1.066	0.123	4.579	2.174	0.202	0.023
SVR	0.786	0.064	0.582	0.071	1.296	0.414	0.118	0.03
XGB	0.796	0.051	0.599	0.066	1.233	0.341	0.132	0.038
RF	0.793	0.052	0.588	0.098	1.264	0.405	0.128	0.043
DT	0.729	0.078	0.643	0.068	1.619	0.405	0.118	0.025

Powyższe wyniki można podsumować następująco:

- Najkorzystniejsze wyniki rozpatrywanych wskaźników uzyskano przy zastosowaniu modeli SVR, RF oraz XGB.
- Przy zastosowaniu modelu SVR uzyskano wysokie \bar{R}_{cv}^2 i niskie \overline{MAE}_{cv} , ale wyższe \overline{MSE}_{cv} w porównaniu do XGB i RF.
- Zastosowanie RF pozwala na osiągnięcie wysokiego \bar{R}_{cv}^2 oraz najniższe \overline{MAE}_{cv} . Jednakże, biorąc pod uwagę wszystkie metryki, zastosowanie XGB pozwala na uzyskanie najkorzystniejszych wyników.

Wstępne badania statystyczne zaprezentowane w rozdziale 8 wskazywały na konieczność zastosowania modeli nieliniowych do modelowania wpływu zmiennych niezależnych na MRR. Wnioski te znajdują swe potwierdzenie w powyższych wynikach analizy porównawczej. Zastosowanie modelu regresji liniowej spowodowało niemal dwukrotnie wyższy średni błąd bezwzględny \overline{MAE}_{cv} , niż modele nieliniowe, takie jak SVR (z radialną funkcją jądrową), RF oraz XGB. Korzystne wyniki uzyskane przy zastosowaniu modeli drzewiastych RF i XGB, można prawdopodobnie przypisać wielomodalności rozkładu danych pomiarowych. Modele te wykorzystują progi decyzyjne, które dzielą przestrzeń danych na mniejsze, bardziej jednorodne regiony. Dzięki temu są w stanie identyfikować skupienia danych, co jest szczególnie ważne w przypadku rozkładów wielomodalnych. Ponadto, modele te posiadają wbudowane mechanizmy redukcji ryzyka przeuczenia, co jest częstym problemem w innych modelach, takich jak sztuczne sieci neuronowe czy pojedyncze drzewa decyzyjne. W przypadku RF, losowe podzbiory danych oraz losowy wybór zmiennych niezależnych do każdego drzewa pomagają w zapobieganiu przeuczeniu. W XGB, iteracyjne dodawanie drzew, które poprawiają błędy poprzedników, oraz mechanizmy regularyzacyjne, takie jak shrinkage i subsampling, minimalizują ryzyko nadmiernego dopasowania do danych treningowych [64],[65].

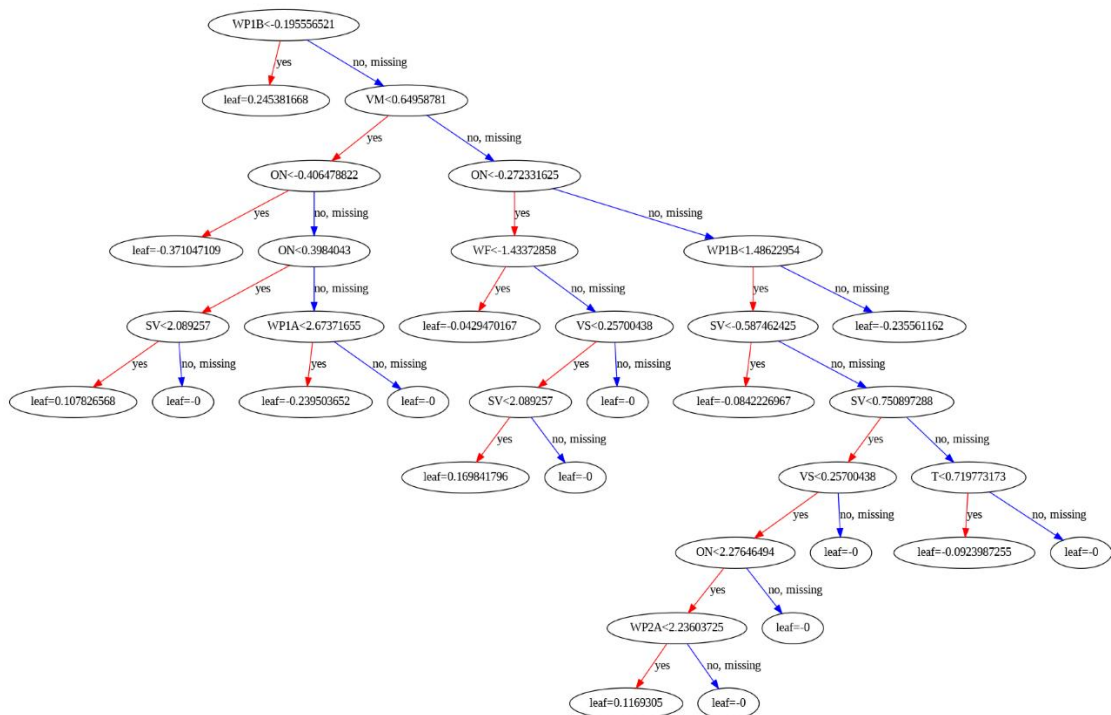
Na podstawie powyższej analizy zdecydowano o wyborze modelu XGB do modelowania wpływu zmiennych niezależnych na MRR.

Dokładna konfiguracja dobranych wartości hiperparametrów modelu została zaprezentowana w Tabeli 9.9. Hiperparametry nieujęte w tabeli zostały ustawione na wartości domyślne według biblioteki programistycznej XGBoost.

Tabela 9.9 Hiperparametry modelu XGB

Parametr	booster	colsampl e_bytree	learning_ rate	max_d epth	min_child_ weight	n_esti mators	reg_alp ha	subsam ple	eta	min_spl it_loss
Wartość	dart	0.8	0.1	12	1	500	6	0.8	0.5	1

Na Rys. 9.4 przedstawiono graficzną reprezentację jednego z drzew decyzyjnych wygenerowanych przez model XGB. Na diagramie można zauważyć różne węzły decyzyjne, które definiują progi decyzyjne stosowane do podziału przestrzeni danych na mniejsze, bardziej jednorodne regiony. Linie czerwone i niebieskie odpowiadają odpowiedziom "tak" i "nie" na pytania decyzyjne w węzłach, prowadząc do kolejnych podziałów aż do osiągnięcia liści drzewa, które reprezentują wartości przewidywane przez drzewo. Ostateczna wartość przewidywana przez model jest wynikiem agregacji wyników wszystkich drzew wchodzących w skład lasu.



Rys. 9.4 Graficzna reprezentacja jednego z drzew decyzyjnych wygenerowanych przez model XGB.

9.2.2. Przygotowanie danych treningowych

W przypadku zmiennej zależnej MRR zgromadzony zbiór danych posiada 6.2% obserwacji z brakującymi pomiarami. Zidentyfikowany model XGBoost wykorzystuje wbudowany mechanizm obsługi brakujących wartości, znany jako podział uwzględniający rzadkość danych (ang. Sparsity Aware Splitting). Mechanizm ten automatycznie identyfikuje brakujące wartości i traktuje je jako odrębną kategorię. Podczas budowy drzew decyzyjnych brakujące wartości są przypisywane do lewej lub prawej gałęzi drzewa w sposób, który minimalizuje funkcję straty [69].

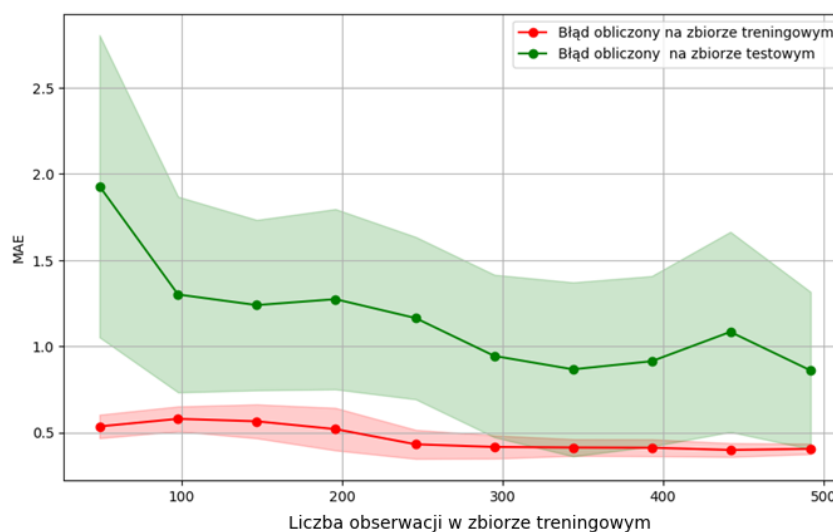
Zastosowane metody uzupełniania brakujących wartości zmiennej zależnej (zob. p. 6.1.1) nie poprawiły w istotny sposób jakości modelu, z czego można wnioskować, że

wyżej opisany mechanizm okazał się wystarczający. **Zdecydowano zatem nie stosować dodatkowych metod uzupełniania brakujących wartości.**

Ponadto, w przypadku modeli drzewiastych, takich jak XGBoost, nie ma potrzeby normalizacji danych, ze względu na niewrażliwość tego typu modeli na skalę i rozkład danych [101].

9.2.3. Analiza jakości modelu

Na Rys. 9.5 przedstawiono zależność średniego błędu absolutnego (MAE) modelu od liczby obserwacji zbioru treningowego oraz testowego.

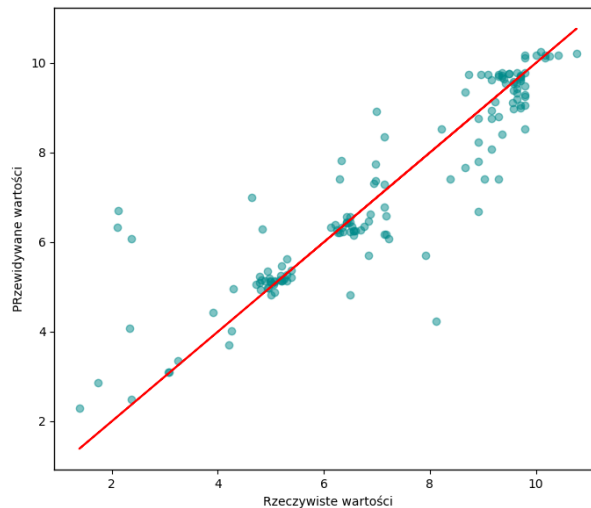


Rys. 9.5 Krzywa uczenia modelu

Błąd obliczony dla zbioru treningowego wykazuje niskie wartości, oscylując wokół $0.5 \text{ mm}^3/\text{min}$, co sugeruje, że model dobrze dopasowuje się do danych treningowych już przy relatywnie niewielkiej liczbie obserwacji.

Błąd obliczony dla zbioru testowego początkowo wynosi około $2 \text{ mm}^3/\text{min}$, co sugeruje ograniczoną zdolność modelu do generalizacji przy małej liczbie obserwacji. Wraz ze wzrostem liczby obserwacji, błąd obliczony dla zbioru testowego systematycznie maleje, osiągając wartość około $1 \text{ mm}^3/\text{min}$ przy 300 obserwacjach. Powyżej tej liczby danych błąd obliczony dla zbioru testowego stabilizuje się, co sugeruje, że dalsze zwiększanie liczby obserwacji powyżej 400 nie przynosi już istotnych korzyści dla poprawy zdolności generalizacyjnych modelu. Porównanie wartości błędów obliczonych dla obu zbiorów wskazuje, że model jest nieznacznie przeuczony.

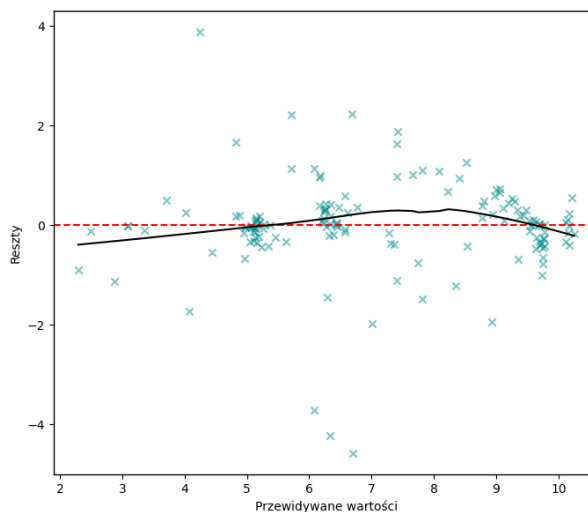
Na Rys. 9.6 przedstawiono zależność pomiędzy wartościami rzeczywistymi a wartościami przewidywanymi przez model MRR.



Rys. 9.6 Wykres rozrzutu przewidywanych i rzeczywistych wartości MRR

Na wykresie widać relatywnie dobre dopasowanie modelu do danych rzeczywistych MRR. Współczynnik korelacji liniowej Pearsona, wynoszący 0.909, potwierdza tę obserwację, co wskazuje na skutecznie przewidywanie wartości MRR przez model.

Na Rys. 9.7 zaprezentowano wykres reszt analizowanego modelu.



Rys. 9.7 Reszty modelu w stosunku do przewidywanych wartości MRR

Na powyższym wykresie reszt można zauważyć, że większość punktów jest rozmieszczona w miarę równomiernie wokół linii zero, co sugeruje, że model w ogólnym ujęciu działa poprawnie. Jednakże, można dostrzec pewne większe rozproszenie reszt w zakresie przewidywanych wartości między 6 a 9 mm³/min, co może wskazywać na większe błędy przewidywania w tym przedziale. Przy niższych wartościach przewidywanych (około 3-5 mm³/min) oraz wyższych wartościach (około 9-10 mm³/min),

reszty są bardziej zgrupowane wokół linii zero, co wskazuje na lepszą jakość modelu w tych zakresach. Szczególnie wartości w wyższym przedziale wydają się być przewidywane z mniejszym błędem, choć nadal można zauważyć lekkie przesunięcie w stronę ujemnych reszt dla najwyższych wartości, co sugeruje pewną tendencję do przeszacowywania w tym zakresie.

Statystyki rozkładu reszt modelu przedstawiono w Tabeli 9.10.

Tabela 9.10 Statystyki rozkładu reszt modelu MRR

Kurtoza	Skośność	DW	JB	p-wartość dla testu JB
8.123	-1.17	1.945	437.78	8.61E-96

Na podstawie analizy statystycznej rozkładu reszt modelu można wyciągnąć następujące wnioski:

- Wartość kurtozy wynosząca 8.123 wskazuje na znaczące odchylenie od normalności w rozkładzie reszt. Wysoka kurtoza sugeruje obecność dłuższych ogonów i bardziej wyraźnego szczytu niż w rozkładzie normalnym [100].
- Skośność reszt wynosząca -1.17 wskazuje na znaczną ujemną asymetrię [100].
- Statystyka Durbina-Watsona (DW) wynosząca 1.945 jest bliska wartości 2, co wskazuje na brak istotnej autokorelacji w resztach.
- Wartość statystyki Jarque-Bera (JB) wynosząca 437.78 oraz bardzo niska p-wartość (8.61E-96) jednoznacznie wskazują na istotne odchylenie rozkładu reszt od normalnego.

W celu ostatecznej oceny modelu obliczono wskaźniki statystyczne błędu przewidywania na zbiorze testowym. Wyniki przedstawiono w Tabeli 9.11.

Tabela 9.11 Ocena jakości modelu na zbiorze testowym

R^2_{test}	MAE_{test}	MSE_{test}	$MAPE_{test}$
0.828	0.546	0.892	0.113

Podsumowując, model wykazuje dobrą zdolność przewidywania zmiennej zależnej, co potwierdzają wysoki współczynnik determinacji oraz umiarkowane wartości błędów MAE i MAPE. Mimo, że test Jarque-Bera oraz analiza rozkładu reszt wskazały na odchylenia od normalności oraz asymetrię rozkładu, nie musi stanowić to problemu, ponieważ XGBoost, jako metoda oparta na drzewach decyzyjnych, nie wymaga spełnienia założeń dotyczących normalności reszt, które są istotne w klasycznych modelach regresji [64]. Warto jednak zaznaczyć, że nienormalność rozkładu reszt może mieć wpływ na interpretację niektórych wyników.

Model może być wykorzystany w dalszych badaniach, jednak należy uwzględnić jego ograniczenia. W szczególności zwraca uwagę większe rozproszenie przewidywanych wartości w przypadku wyższych rzeczywistych wartości MRR (powyżej 8 mm³/min), co sugeruje, że model może być mniej precyzyjny w tym zakresie.

9.3. Model C

9.3.1. Dobór metody modelowania

W Tabeli 9.12 przedstawiono wyniki analizy porównawczej zidentyfikowanych modeli C, z wykorzystaniem metryk zdefiniowanych w p.6.3.2.

Tabela 9.12 Porównanie opracowanych modeli C

Model	\overline{AUC}_{cv}	σAUC_{cv}	\overline{ACC}_{cv}	σACC_{cv}	\overline{PRE}_{cv}	σPRE_{cv}	\overline{SEN}_{cv}	σSEN_{cv}
LGR	0.888	0.005	0.811	0.020	0.826	0.028	0.793	0.004
SVC	0.974	0.005	0.930	0.014	0.928	0.021	0.933	0.025
XGB	0.992	0.005	0.979	0.005	0.960	0.009	1.000	0.000
RF	0.963	0.012	0.889	0.019	0.876	0.024	0.904	0.018

Wyniki zamieszczone w Tabeli 9.12 wyraźnie wskazują, że model XGB jest najlepszym wyborem spośród analizowanych modeli. Model ten charakteryzuje się najkorzystniejszymi wynikami we wszystkich rozpatrywanych wskaźnikach.

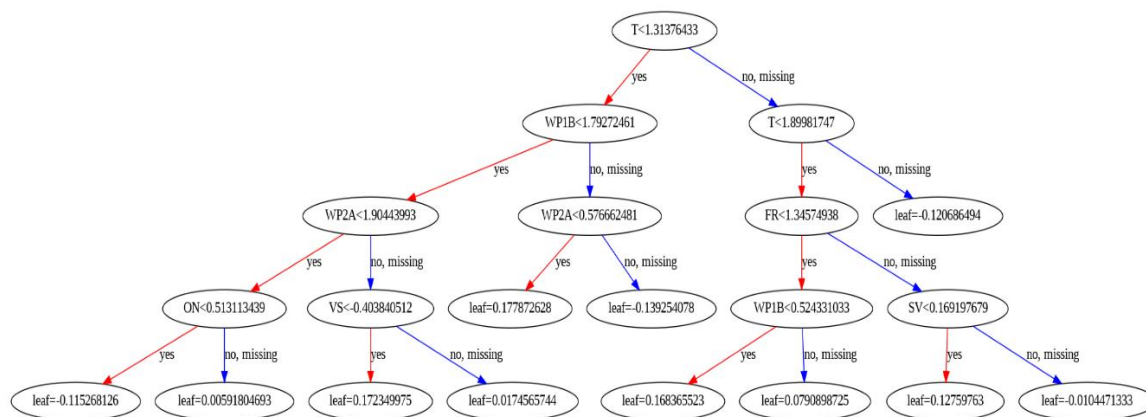
W związku z tym zdecydowano o wyborze modelu XGB do modelowania wpływu zmiennych niezależnych na ciągłość obróbki.

Podobnie jak w przypadku modelowania MRR, Tabela 9.13 przedstawia hiperparametry zidentyfikowanego modelu, które różnią się od domyślnych, zaimplementowanych w bibliotece programistycznej XGBoost.

Tabela 9.13 Hiperparametry modelu XGB

Parametr	colsample_bytree	learning_rate	max_depth	n_estimators	eta
Wartość	0.7	0.1	4	300	0.3

Na Rys. 9.8 zilustrowano graficzną reprezentację jednego z drzew decyzyjnych wygenerowanych przez model XGB.



Rys. 9.8 Graficzną reprezentacją jednego z drzew decyzyjnych wygenerowanych przez model XGB.

9.3.2. Przygotowanie danych treningowych

Dane dotyczące ciągłości charakteryzują się kompletnym zbiorem danych. W związku z tym, metody uzupełniania brakujących danych, stosowane w innych modelach, nie były w tym przypadku konieczne. Ponadto, zastosowanie modelu XGBoost eliminuje potrzebę normalizacji danych.

W analizowanej klasyfikacji binarnej istotnym zagadnieniem jest odpowiednia reprezentacja klas w zbiorze danych treningowych. W celu zapewnienia odpowiedniego zrównoważenia klas przeprowadzono analizę wpływu modyfikacji rozkładu danych przy użyciu metod resamplingowych (zob. p.0) na wskaźniki jakości modelu oraz porównano wyniki z modelem, w którym nie ingerowano w strukturę zbioru danych. Wyniki analizy przedstawiono w Tabeli 9.14.

Tabela 9.14 Wpływ metody równoważenia klas na jakość modelu S

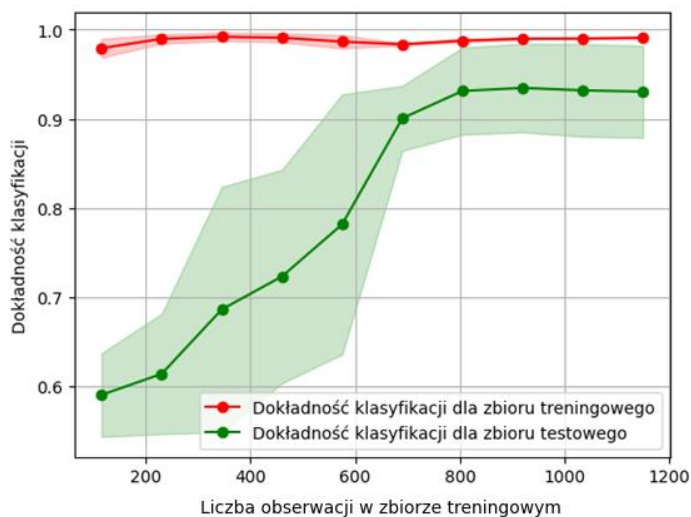
Metoda	\overline{AUC}_{cv}	σAUC_{cv}	\overline{ACC}_{cv}	σACC_{cv}	\overline{PRE}_{cv}	σPRE_{cv}	\overline{SEN}_{cv}	σSEN_{cv}
Oversampling klasy mniejszościowej	0.992	0.005	0.979	0.005	0.960	0.009	1.000	0.000
SMOTE	0.994	0.004	0.962	0.016	0.960	0.013	0.963	0.024
Undersampling klasy większościowej	0.895	0.080	0.754	0.026	0.800	0.110	0.752	0.102
Brak ingerencji w rozkład danych	0.916	0.033	0.940	0.014	0.387	0.317	0.211	0.194

Wyniki jednoznacznie wskazują, że model wytrenowany na danych ze zmodyfikowanym rozkładem przy użyciu metody oversamplingu cechuje się najkorzystniejszymi wskaźnikami jakości. Model wytrenowany na danych o niezmiennym rozkładzie wykazuje porównywalną, choć minimalnie niższą dokładność

klasyfikacji. Niemniej jednak, wskaźniki czułości i precyzji klasyfikacji w tym przypadku są znacznie niższe, co sugeruje, że model prawdopodobnie klasyfikuje głównie obserwacje należące do klasy większościowej. **Na tej podstawie zdecydowano o wyborze metody oversamplingu klasy mniejszościowej do modyfikacji rozkładu danych treningowych.**

9.3.3. Analiza jakości modelu

Rys. 9.9 przedstawia zależność dokładności modelu od liczby próbek treningowych, osobno dla zbioru treningowego i testowego.

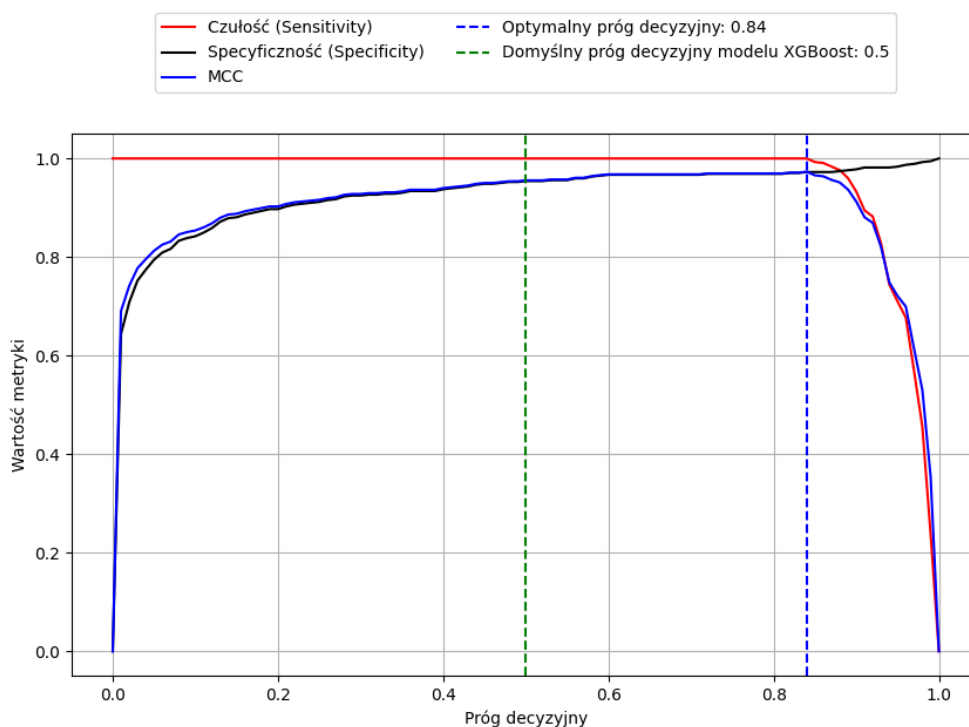


Rys. 9.9 Krzywa uczenia modelu

Jak widać na Rys. 9.9 dokładność klasyfikacji obliczona dla zbioru treningowego utrzymuje się na wysokim poziomie niezależnie od liczby obserwacji, co sugeruje, że model dobrze dopasowuje się do danych treningowych. Z kolei dokładność klasyfikacji obliczona dla zbioru testowego początkowo jest niska, ale znacząco wzrasta wraz ze zwiększaniem liczby obserwacji w zbiorze treningowym. Po osiągnięciu około 800 obserwacji, dokładność klasyfikacji stabilizuje się na poziomie około 90%, co wskazuje na optymalne możliwości klasyfikacyjne modelu. Stabilizacja ta sugeruje, że dalsze zwiększanie liczby danych może nie przynieść istotnej poprawy wyników.

Istotnym zagadnieniem w problemach klasyfikacji binarnej jest zdefiniowanie odpowiedniej wartości progu decyzyjnego [102]. Próg decyzyjny wyznacza granicę wartości wskazującej na przynależność do danej klasy. Na przykład, w rozpatrywanym przypadku klasyfikacji binarnej z klasami „0” i „1”, klasyfikator przypisuje obserwację do klasy „1” (pozytywnej), jeśli wartość wyjściowa modelu przekracza ustalony próg decyzyjny. Tradycyjnie próg ten ustala się na wartość 0.5 [102], jednakże w zależności od potrzeb co do czułości i specyficzności modelu próg ten może być zmieniany.

Na Rys. 9.10 przedstawiono wpływ wartości progu decyzyjnego na czułość modelu (czerwona krzywa), specyficzną (czarna krzywa) oraz korelację pomiędzy przewidywanymi a rzeczywistymi klasami (niebieska krzywa), mierzoną za pomocą współczynnika korelacji Matthews MCC (zob. p. 6.3.2). Obliczenia do wykresu zostały przeprowadzone z użyciem zbioru danych treningowych, a wartości poszczególnych metryk są średnimi uzyskanymi w procesie walidacji krzyżowej. Dodatkowo oznaczone są dwa progi decyzyjne: domyślny próg decyzyjny modelu XGBoost (0.5) oraz próg decyzyjny, przy którym wartość współczynnika MCC była najwyższa (0.84).



Rys. 9.10 Wpływ progu decyzyjnego na czułość, specyficzną i współczynnik korelacji Matthews (MCC)

Z perspektywy zastosowań przemysłowych kluczowa jest wysoka czułość modelu. Niewłaściwa klasyfikacja parametrów technologicznych jako niepowodujących przerwania ciągłości obróbki, gdy w rzeczywistości dojdzie przerwania, może prowadzić do istotnych strat produkcyjnych. Jak wskazuje wykres, model bardzo dobrze identyfikuje przypadki przerwania ciągłości obróbki. Czułość modelu pozostaje wysoka (blisko 1.0) dla większości progów, gwałtownie spadając dopiero powyżej progu 0.84. Specyficzność natomiast rośnie wraz z progiem, osiągając maksymalne wartości bliskie 1.0. MCC osiąga maksymalną wartość przy progu 0.84, co wskazuje na najlepszą równowagę między czułością a specyficzną.

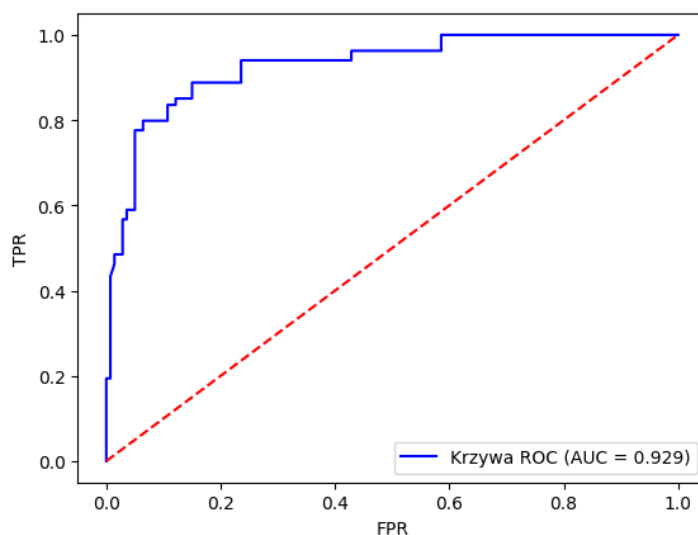
W celu pogłębienia analizy wpływu progów decyzyjnych na jakość modelu obliczono również odpowiednie wskaźniki jakości klasyfikacji dla zbioru danych testowych. Wyniki dla progu 0.5 i 0.84 przedstawiono w Tabeli 9.15.

Tabela 9.15 Porównanie wpływu progu decyzyjnego na jakość modelu

Próg decyzyjny modelu	ACC_{test}	SEN_{test}	SPE_{test}	PRE_{test}
0.5	0.989	1.000	0.979	0.978
0.84	0.993	1.000	0.986	0.985

Wyniki przedstawione w tabeli oraz wcześniejszy wykres jednoznacznie wskazują na przewagę progu decyzyjnego równego 0.84. W związku z tym, ten próg został przyjęty w dalszych badaniach.

Kolejnym krokiem analizy jakości modelu jest ocena jego zdolności do rozróżniania klas za pomocą krzywej ROC i wartości AUC (Rys. 9.11).



Rys. 9.11 Krzywa ROC modelu

Jak widać na powyższym wykresie, krzywa ROC modelu jest przesunięta blisko lewego górnego rogu, co wskazuje na wysoką czułość i swoistość. Wartość AUC wynosząca 0.929 wskazuje, że model skutecznie rozróżnia klasy pozytywne i negatywne. Taki wynik świadczy o wysokiej jakości modelu w kontekście klasyfikacji.

Podsumowując, model wykazuje wysoką skuteczność w przewidywaniu ciągłości obróbki, osiągając dokładność klasyfikacji na poziomie 99% na zbiorze testowym. Spośród 274 obserwacji w zbiorze testowym odnotowano jedynie 2 błędne klasyfikacje. Model niepoprawnie zaklasyfikował te przypadki jako powodujące utratę stabilności obróbki, mimo że w rzeczywistości stabilność została zachowana. Kluczowym wnioskiem

z perspektywy wdrożeniowej jest 100% skuteczność modelu w wykrywaniu rzeczywistych przypadków niestabilności obróbki.

9.4. Model Kw

9.4.1. Dobór metody modelowania

W Tabeli 9.12 przedstawiono wyniki analizy porównawczej zidentyfikowanych modeli Kw.

Tabela 9.16 Porównanie opracowanych modeli Kw

Model	\overline{AUC}_{cv}	σAUC_{cv}	\overline{ACC}_{cv}	σACC_{cv}	\overline{PRE}_{cv}	σPRE_{cv}	\overline{SEN}_{cv}	σSEN_{cv}
SVC	0.890	0.019	0.802	0.015	0.816	0.014	0.789	0.016
XGB	0.908	0.037	0.782	0.099	0.807	0.067	0.804	0.074
ANN	0.847	0.069	0.791	0.085	0.818	0.075	0.791	0.085
LG	0.873	0.024	0.761	0.103	0.751	0.099	0.757	0.090

Wyniki zamieszczone w Tabeli 9.16 wskazują,

- Najkorzystniejsze wartości w większości z ocenianych wskaźników uzyskano przy zastosowaniu modelu SVC. Wyróżnia się on szczególnie wysoką dokładnością klasyfikacji ($\overline{ACC}_{cv} = 0.802$) przy jednoczesnej stabilności wyników, co potwierdzają niskie wartości odchyłeń standardowych dla wszystkich metryk.
- Przy zastosowaniu modelu XGB uzyskano $\overline{AUC}_{cv} = 0.908$, co wskazuje na wysokie zdolności w rozróżnianiu klas. Jednakże większe odchylenia standardowe, zwłaszcza dla dokładności klasyfikacji i precyzji, wskazują na większą zmienność wyników w porównaniu do modelu SVC.
- Zastosowanie ANN przyniosło umiarkowane wyniki, uzyskując porównywalną precyzję ($\overline{PRE}_{cv} = 0.818$) do SVC, lecz niższe wartości \overline{AUC}_{cv} oraz większe odchylenia standardowe, co może wskazywać na mniejszą stabilność tego modelu.
- Najmniej korzystne wyniki uzyskano przy zastosowaniu modelu LG, który charakteryzuje się najniższą dokładnością klasyfikacji, precyzją oraz czułością

Na podstawie uzyskanych wyników stwierdzono, że model SVC jest najbardziej odpowiedni do dalszych zastosowań w badaniach.

W Tabeli 9.17 zestawiono wartości hiperparametrów zidentyfikowanego modelu SVC dostosowane w procesie walidacji krzyżowej.

Tabela 9.17 Hiperparametry zidentyfikowanego modelu SVC

Parametr	C	coef0	degree	gamma	kernel	max_iter	tol
Wartość	2.5	1.0	2	0.1	poly	1000000	0.001

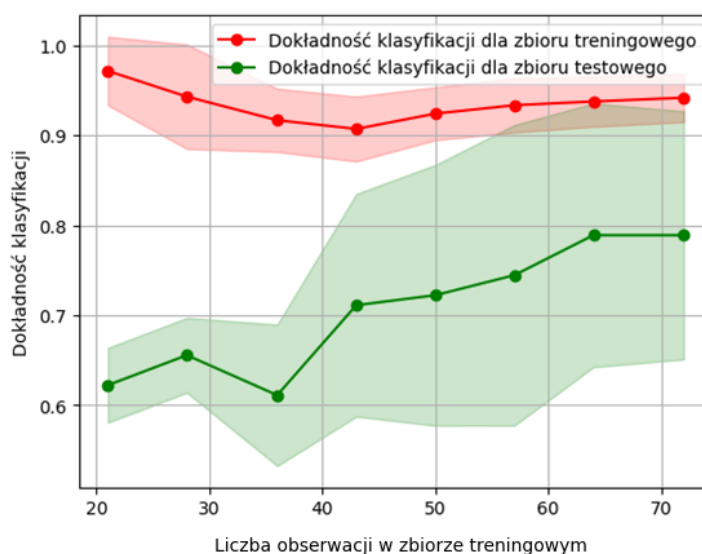
9.4.2. Dobór metody przygotowania danych

Rozkład danych pomiędzy klasami był niemal równomierny (Klasa 1 – 27 obserwacji, Klasa 2 – 27 obserwacji, Klasa 3 – 26 obserwacji). W związku z minimalnymi różnicami licznosci, nie zachodziła potrzeba stosowania metod balansowania danych.

Aby zapewnić jednolitą skalę zmiennych i poprawić efektywność algorytmów, zastosowano standaryzację danych przy użyciu standaryzacji Z.

9.4.3. Analiza jakości modelu

Rys. 9.12 przedstawia zależność dokładności klasyfikacji modelu od liczby obserwacji treningowych, osobno dla zbioru treningowego i testowego.

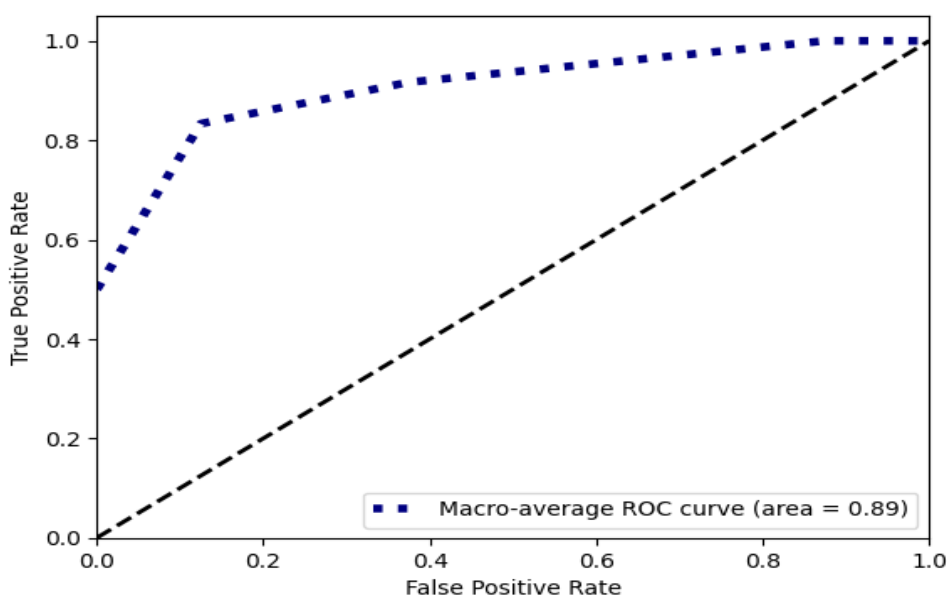


Rys. 9.12 Krzywa uczenia modelu Kw

Na przedstawionym powyżej wykresie można zauważyć, że dokładność klasyfikacji obliczona dla zbioru treningowego pozostaje na wysokim poziomie, niezależnie od liczby obserwacji. Z kolei dokładność klasyfikacji obliczona dla zbioru testowego, początkowo niższa, stopniowo wzrasta wraz ze zwiększaniem liczby danych. Taki trend sugeruje, że model zyskuje większą zdolność do generalizacji przy większej liczbie obserwacji. Mimo to, dokładność klasyfikacji obliczona dla zbioru testowego pozostaje niższa niż na zbiorze treningowym, co wskazuje na ryzyko przeuczenia modelu. Dodatkowo, należy zwrócić uwagę na znaczne odchylenie standardowe dokładności klasyfikacji dla zbioru testowego,

co sugeruje, że model jest niestabilny, a jego wyniki są bardzo wrażliwe na wahania w danych testowych. Może to prowadzić do niepewności w ocenie rzeczywistej zdolności modelu do generalizacji.

W celu przedstawienia zdolności modelu do rozróżniania klas wygenerowano uśrednioną makro krzywą ROC dla klasyfikacji wieloklasowej (Rys. 9.13).



Rys. 9.13 Krzywa ROC dla klasyfikacji wieloklasowej (uśrednianie makro)

Krzywa ROC jest wypukła i znajduje się powyżej linii referencyjnej, co świadczy o skuteczności modelu w rozróżnianiu klas. Gwałtowny wzrost wskaźnika prawdziwie pozytywnych klasyfikacji (TPR) przy niskim wskaźniku fałszywie pozytywnych klasyfikacji (FPR) wskazuje na wysoką jakość modelu. Potwierdza to również wartość AUC na poziomie 0.89.

W Tabeli 9.18 przedstawiono wyniki metryk oceniających jakość działania modelu na zbiorze testowym

Tabela 9.18 Ocena jakości modelu na zbiorze testowym

ACC_{test}	SEN_{test}	PRE_{test}
0.82	0.81	0.8

Podsumowując, wyniki klasyfikacji wieloklasowej na zbiorze testowym są ogólnie korzystne. Dokładność, czułość i precyzja wskazują na wysoką zdolność modelu do poprawnej klasyfikacji oraz efektywnego rozróżniania pomiędzy klasami. Jednakże, mimo tych pozytywnych wyników, należy zwrócić uwagę na niestabilność modelu, objawiającą

się znacznym odchyleniem standardowym wyników na zbiorze testowym. Wskazuje to na wrażliwość modelu na zmiany w danych.

9.5. Analiza wpływu parametrów technologicznych na proces WEDM

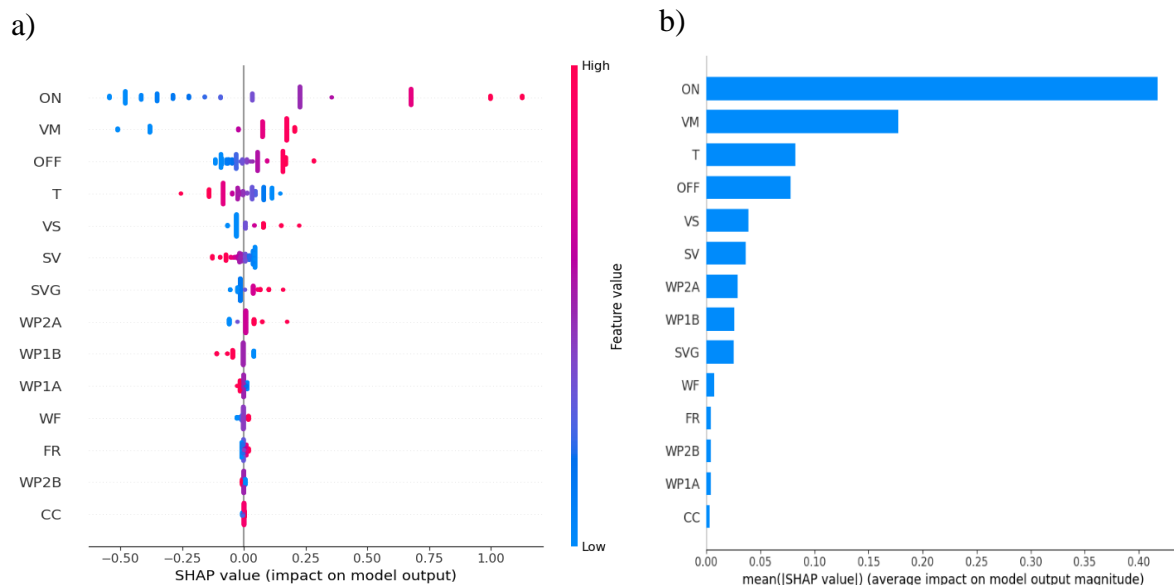
Analiza wpływu parametrów technologicznych została przeprowadzona dla każdej ze zmiennych zależnych według powtarzalnej struktury. W pierwszej części, dla danej zmiennej, przedstawiono wpływ parametrów, wykorzystując metodę interpretacji modeli SHAP (zob. p. 6.4). Zaprezentowano wykres punktowy oraz słupkowy ilustrujący wpływ parametrów na wyniki uzyskane z zastosowaniem danego modelu.

Wykres punktowy wartości SHAP obrazuje indywidualny wpływ parametrów na wyniki uzyskane z zastosowaniem danego modelu, gdzie kolory wskazują wartości parametrów, a położenie punktów pokazuje, w jakim stopniu wpływają one na te wyniki. Z kolei wykres słupkowy prezentuje uśrednione wartości SHAP, które pozwalają zidentyfikować, które z parametrów mają największe znaczenie w kontekście wpływu na przewidywania danej zmiennej zależnej uzyskane podczas symulacji modelu.

W dalszej części zaprezentowano trójwymiarowe wykresy, ilustrujące wpływ wybranych par parametrów w znormalizowanej skali $[-1, 1]$ (6.5) na zmienną zależną, przy założeniu, że pozostałe parametry przyjmują wartości średnie. Na zakończenie analizy przedstawiono podsumowanie wpływu najważniejszych parametrów na daną zmienną zależną.

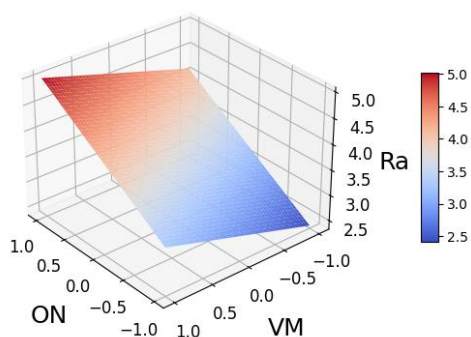
9.5.1. Wpływ parametrów technologicznych na chropowatość powierzchni

Poniżej przedstawiono wykres ukazujący rozkład wartości SHAP dla analizowanych parametrów technologicznych, ilustrujący ich wpływ na wynik modelu Ra (Rys. 9.14).

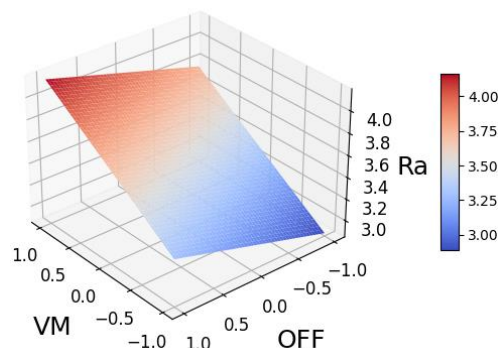


Rys. 9.14 Wykres punktowy wartości SHAP (a) oraz wykres uśrednionych wartości SHAP (b) dla poszczególnych zmiennych niezależnych modelu Ra

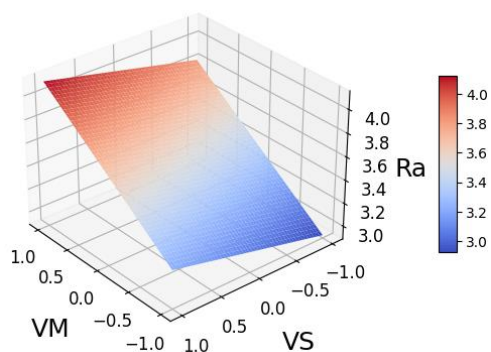
Aby pogłębić analizę wpływu poszczególnych parametrów, wykorzystano wykresy trójwymiarowe, które umożliwiają wizualizację zależności w badanym procesie. Wybrane wyniki przedstawiono na Rys. 9.15 - Rys. 9.18.



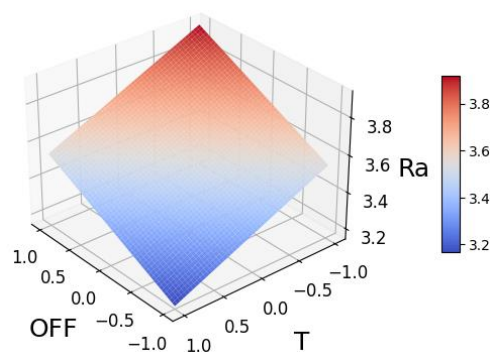
Rys. 9.15 Wpływ ON i VM na Ra



Rys. 9.16 Wpływ VM i OFF na Ra



Rys. 9.17 Wpływ VM i VS na Ra



Rys. 9.18 Wpływ OFF i T na Ra

Na podstawie analizy wyników SHAP oraz wizualnej oceny przedstawionych wykresów trójwymiarowych, a także przeglądu literatury (zob.p.4.1), sformułowano wnioski dotyczące wpływu najistotniejszych badanych parametrów technologicznych na wartość parametru chropowatości Ra:

ON: Zdecowanie najistotniejszym parametrem w kontekście wpływu na Ra jest parametr ON. Przy długich czasach ON zauważalny jest wzrost chropowatości powierzchni Ra, szczególnie w przypadku, gdy pozostałe parametry definiujące energię wyładowania, takie jak VM i VS, są ustawione na wysokich wartościach. Efekt ten, ilustrowany np. na Rys. 9.15, wynika z większej ilości energii dostarczanej do obszaru obróbki, co prowadzi do intensyfikacji procesów topnienia i odparowywania materiału. Konsekwencją tego zjawiska jest powstawanie głębszych kraterów erozyjnych na powierzchni obrabianej, co bezpośrednio przyczynia się do wzrostu Ra (efekt ten omówiono również w p.4.1.2).

- **VM:** Zwiększenie wartości VM prowadzi do wzrostu energii wyładowania, co w konsekwencji zwiększa chropowatość powierzchni Ra (Rys. 9.15).
- **OFF:** Analiza SHAP oraz zależność przedstawiona na Rys. 9.16 wskazują na ujemny wpływ parametru OFF na wartość Ra. Oznacza to, że wydłużenie czasu przerwy wyładowania skutkuje wzrostem chropowatości powierzchni. Wynik ten jest sprzeczny z doniesieniami literaturowymi, np. [26], gdzie zaobserwowano odwrotną zależność – krótszy czas przerwy prowadził do większej chropowatości powierzchni. Aby wyjaśnić tę rozbieżność, przeprowadzono dodatkową analizę korelacji między zmiennymi niezależnymi, która wykazała, że współczynnik korelacji Pearsona pomiędzy parametrami ON i OFF wyniósł 0.723. Stosunkowo wysoka korelacja (wynikająca z konieczności zapobiegania zerwaniu elektrody) oraz silny wpływ parametru ON na wartość Ra sugerują, że zjawisko to może być efektem interferencji tych dwóch parametrów.

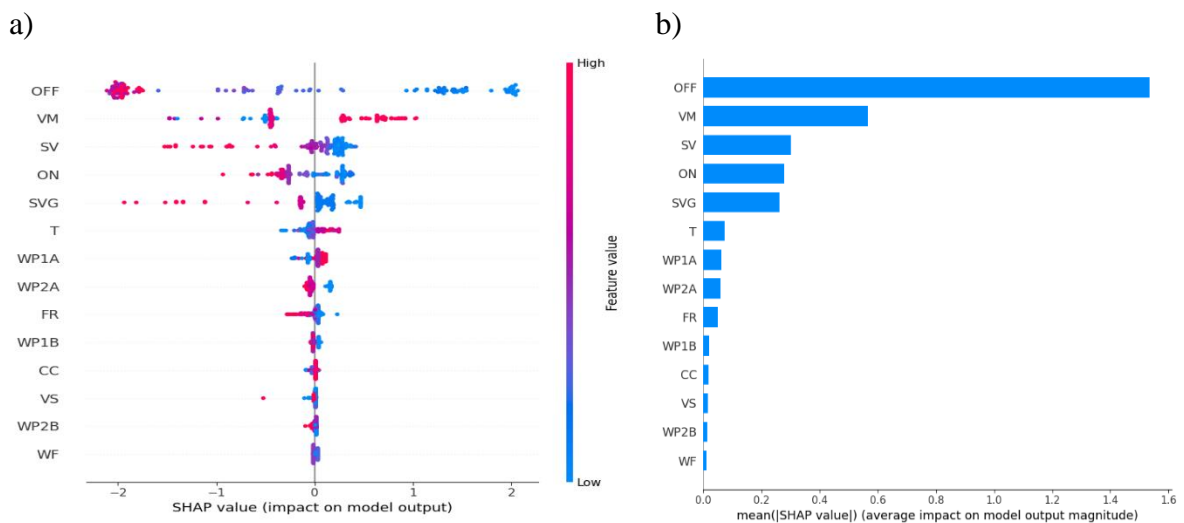
W świetle powyższych wyników, należy przyjąć, że wpływ parametru OFF na chropowatość powierzchni Ra jest zgodny z doniesieniami literaturowymi, gdzie to krótszy czas przerwy wyładowania prowadzi do większej chropowatości. Najprawdopodobniej błędna interpretacja wpływu parametru OFF, wynikająca z interferencji z parametrem ON, nie ma jednak istotnego wpływu na jakość całego modelu, co potwierdzają wcześniejsze analizy. W związku z tym, omawiana rozbieżność nie powinna negatywnie wpłynąć na dalsze badania optymalizacyjne,

ponieważ wpływ OFF jest w tym kontekście drugorzędny w porównaniu do dominującego wpływu parametru ON.

- **T:** Zwiększenie naciągu drutu powoduje zmniejszenie wartości Ra, co zostało zilustrowane na Rys. 9.18. Wynik ten jest zgodny z teoretycznymi przewidywaniami, ponieważ wyższy naciąg drutu prowadzi do redukcji amplitudy jego drgań. Zmniejszenie drgań drutu ogranicza niestabilności wyładowań elektrycznych oraz zmniejsza liczbę mikrouszkodzeń powstających na powierzchni obrabianego materiału, co w konsekwencji sprzyja uzyskaniu bardziej jednorodnego wykończenia powierzchni.
- **VS:** Zwiększenie wartości napięcia otwartej przerwy VS prowadzi do wzrostu chropowatości powierzchni Ra. Rola tego parametru jest ograniczona czasowo, ponieważ odnosi się głównie do momentu przebicia elektrycznego dielektryka. VS wpływa na początkową energię wyładowania, determinując szybkość inicjalizacji wyładowania, co z kolei kształtuje warunki początkowe procesu. Choć parametr ten ma wpływ na chropowatość powierzchni, jego oddziaływanie jest mniej znaczące w porównaniu do wpływu napięcia wyładowania VM, które w większym stopniu determinuje charakterystykę i intensywność wyładowań oraz ostateczną jakość powierzchni (Rys. 9.17).

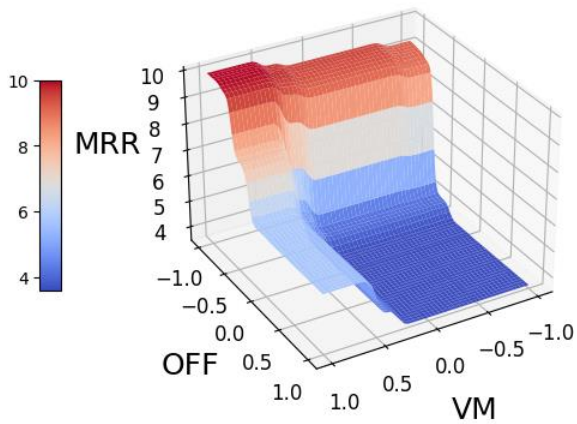
9.5.2. Wpływ parametrów technologicznych na prędkość usuwania materiału

Poniżej zaprezentowano wykres ilustrujący rozkład wartości SHAP dla badanych parametrów technologicznych, pokazując ich wpływ na wynik modelu MRR (Rys. 9.19)

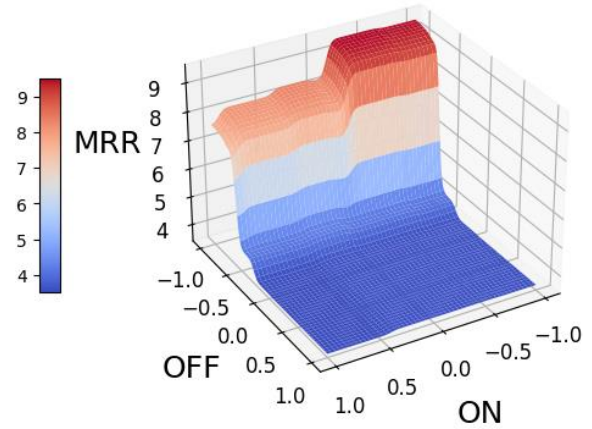


Rys. 9.19 Wykres punktowy wartości SHAP (a) oraz wykres uśrednionych wartości SHAP (b) dla poszczególnych zmiennych niezależnych modelu MRR

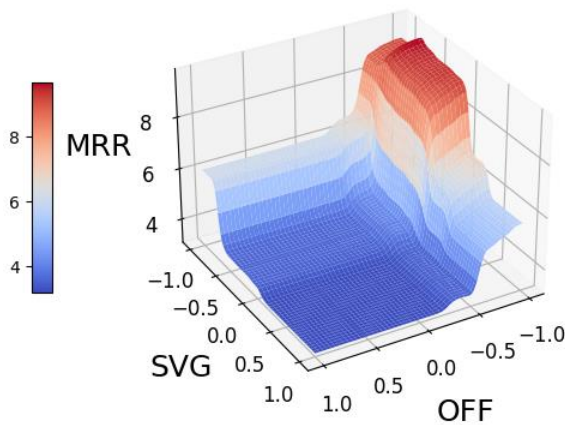
Dalsze pogłębienie analizy wpływu parametrów na wyniki procesu przeprowadzić można z wykorzystaniem poniższych wykresów trójwymiarowych (Rys. 9.20 - Rys. 9.25).



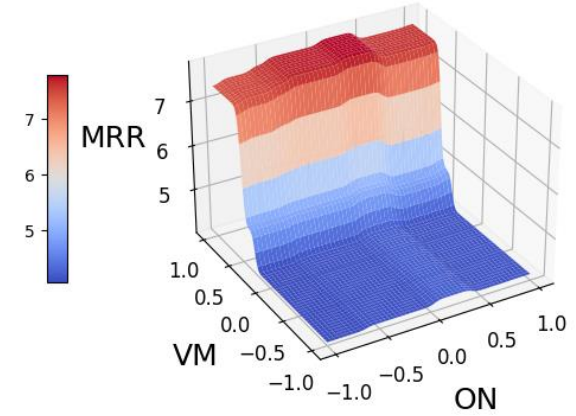
Rys. 9.20 Wpływ OFF i VM na MRR



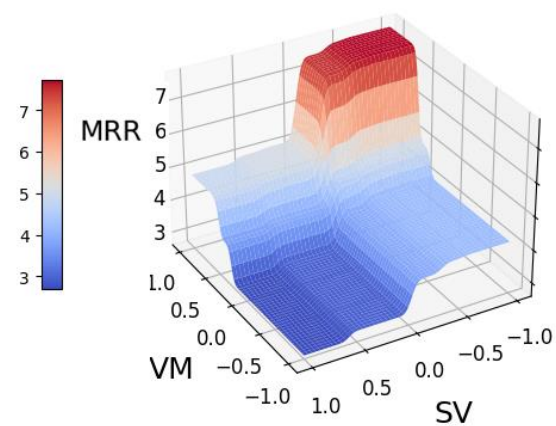
Rys. 9.21 Wpływ OFF i ON na MRR



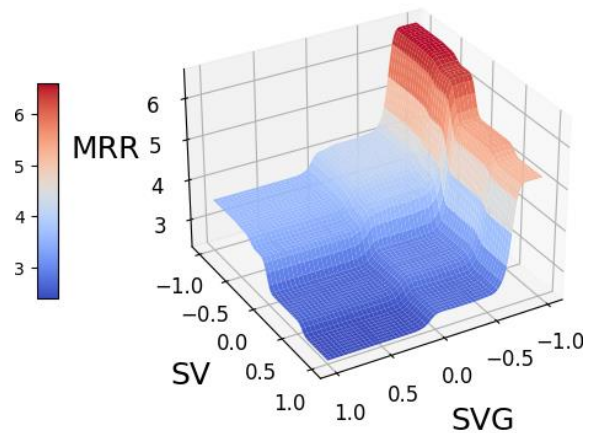
Rys. 9.22 Wpływ SVG i OFF na MRR



Rys. 9.23 Wpływ VM i ON na MRR



Rys. 9.24 Wpływ VM i SV na MRR



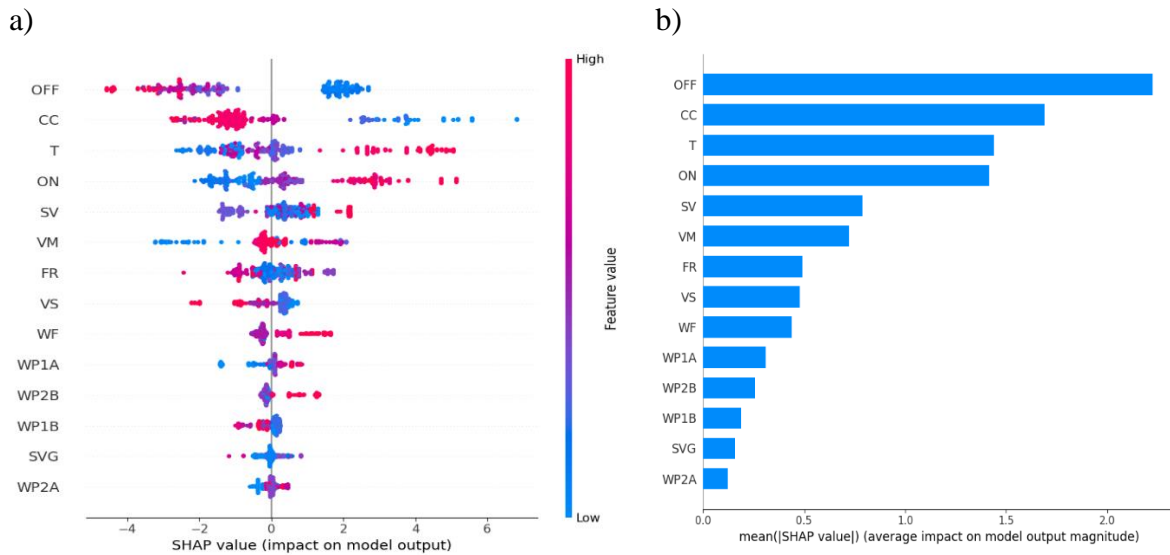
Rys. 9.25 Wpływ SV i SVG na MRR

Na podstawie powyższych wyników oraz analizy literatury sformułowano wnioski:

- **OFF:** Parametr OFF ma kluczowe znaczenie w kontekście wpływu na MRR. Dłuższy czas przerwy między impulsami prowadzi do zmniejszenia liczby wyładowań na jednostkę czasu, co ogranicza MRR. I odwrotnie, krótszy czas OFF znacznie zwiększa MRR (Rys. 9.20).
- **ON:** Wydłużenie czasu ON w pewnym zakresie nieznacznie zwiększa MRR poprzez dostarczenie większej energii na impuls, co intensyfikuje proces usuwania materiału (Rys. 9.23). Jednak przy bardzo długich impulsach dochodzi do obniżenia MRR. Efekt ten wynika prawdopodobnie z faktu przegrzewania elektrody, co wpływa na obniżenie efektywności usuwania materiału, szczególnie gdy czas OFF jest krótki (Rys. 9.21).
- **VM:** Zwiększenie VM prowadzi do wzrostu energii wyładowań, co zwiększa MRR (Rys. 9.20).
- **SV:** Niższe wartości SV sprzyjają stabilności procesu, co pozwala na bardziej efektywne usuwanie materiału. Wysokie wartości SV mogą powodować zbyt częste korekty odległości między elektrodą a materiałem, co zmniejsza MRR (Rys. 9.24).
- **SVG:** Zbyt wysokie wartości SVG mogą prowadzić do nadmiernych oscylacji i destabilizacji procesu, szczególnie przy wysokim VM i SV. Niższe wartości SVG mogą stabilizować proces, gdy OFF jest wydłużony, umożliwiając efektywniejsze wykorzystanie energii impulsów (Rys. 9.25).

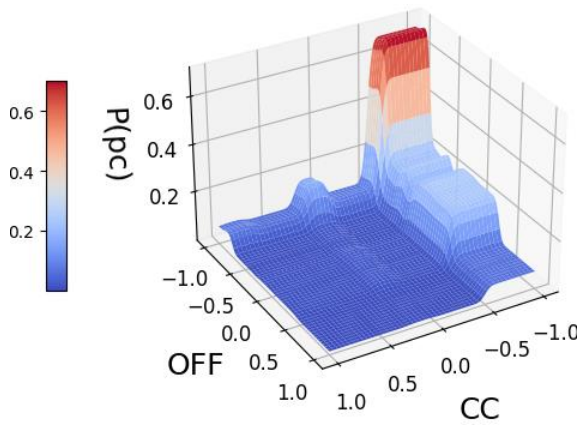
9.5.3. Wpływ parametrów technologicznych na ciągłość procesu

Wykres poniżej przedstawia wartości SHAP, obrazując wpływ poszczególnych parametrów technologicznych na wynik modelu C (Rys. 9.26).

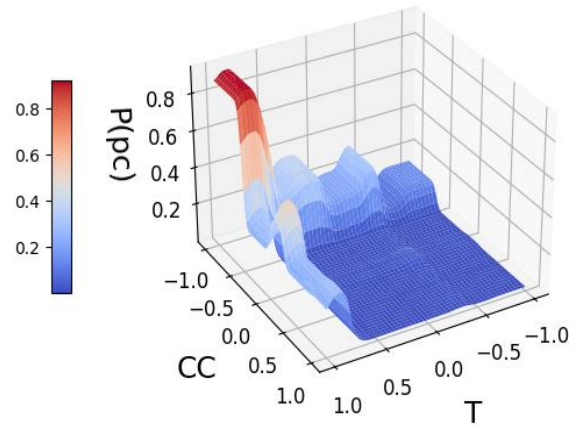


Rys. 9.26 Wykres punktowy wartości SHAP (a) oraz wykres uśrednionych wartości SHAP (b) dla poszczególnych zmiennych niezależnych modelu C

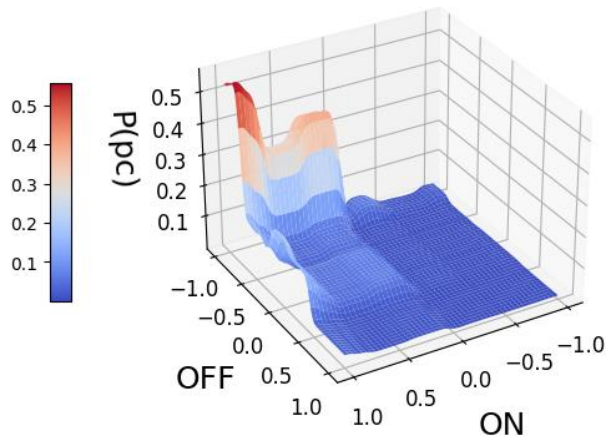
Aby móc przedstawić wpływ parametrów na ciągłość obróbki, wprowadzono wskaźnik $P_{(pc)}$, który reprezentuje prawdopodobieństwo wystąpienia przerwania ciągłości obróbki ($C = 1$). Prawdopodobieństwo to zostało wyznaczone przy użyciu wbudowanej funkcji „predict_proba” dostępnej w implementacji modelu w bibliotece scikit-learn. Poniższe wykresy trójwymiarowe przedstawiają zatem zależność między wybranymi parametrami a tym wskaźnikiem, umożliwiając wizualizację wpływu parametrów na ciągłość obróbki.



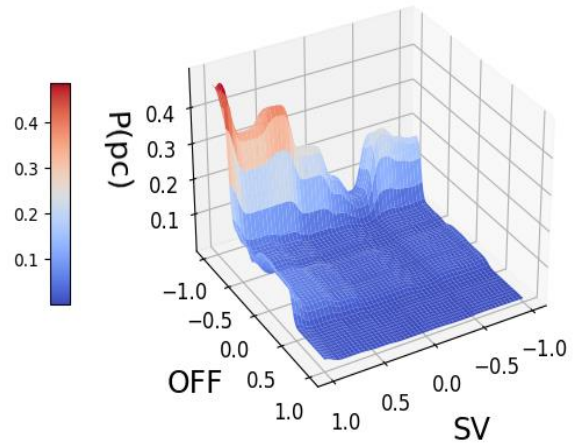
Rys. 9.27 Wpływ CC i OFF na $P_{(pc)}$



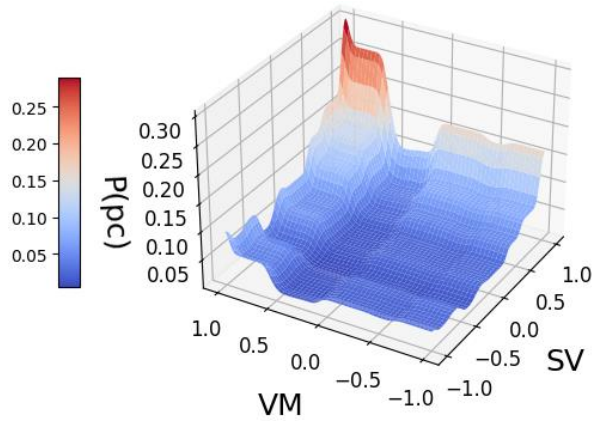
Rys. 9.28 Wpływ T i CC na $P_{(pc)}$



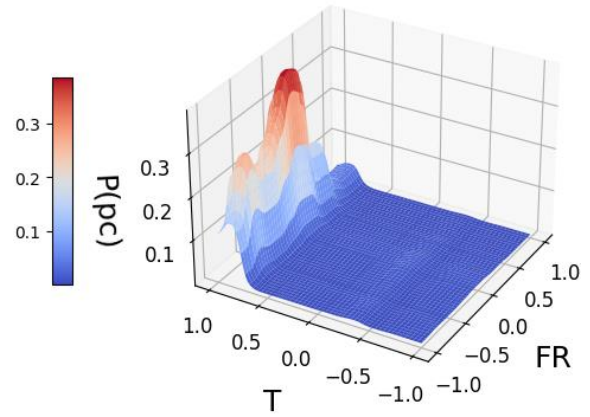
Rys. 9.29 Wpływ ON i OFF na $P_{(pc)}$



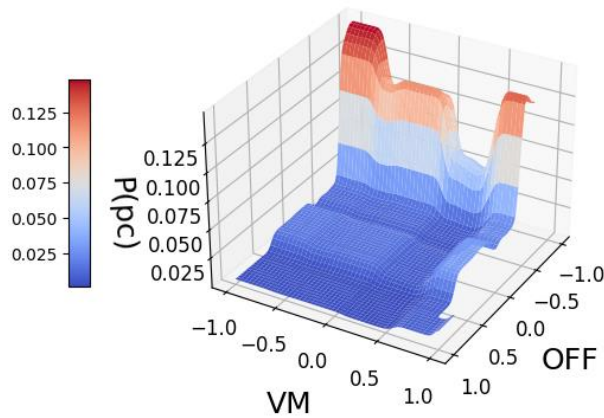
Rys. 9.30 Wpływ SV i OFF na $P_{(pc)}$



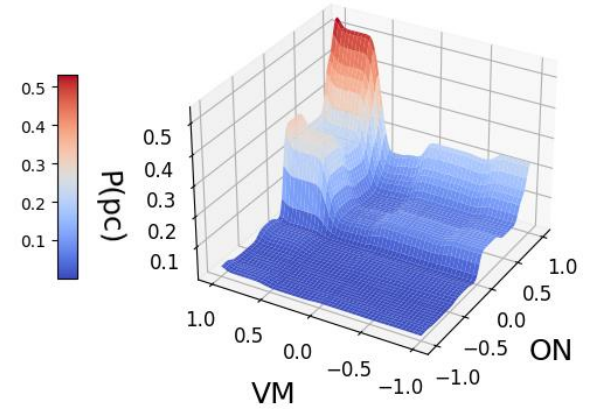
Rys. 9.31 Wpływ SV i VM na $P_{(pc)}$



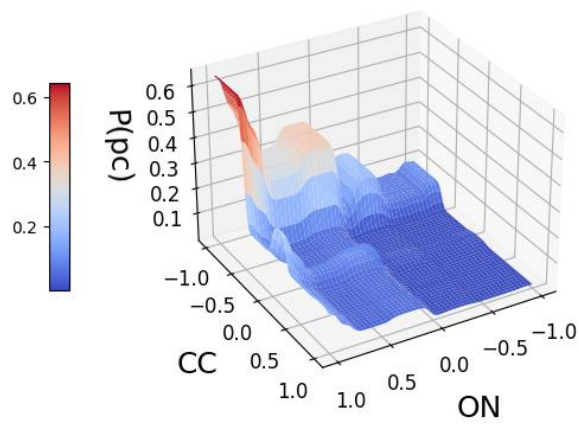
Rys. 9.32 Wpływ FR i T na $P_{(pc)}$



Rys. 9.33 Wpływ OFF i VM na $P_{(pc)}$



Rys. 9.34 Wpływ VM i ON na $P_{(pc)}$



Rys. 9.35 Wpływ ON i CC na $P_{(pc)}$

Na podstawie powyższych wyników oraz analizy literatury sformułowano wnioski:

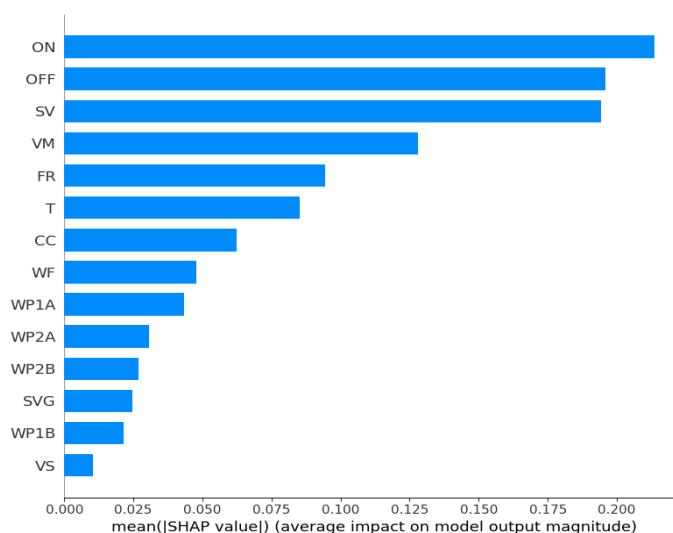
- **OFF:** Parametr ten jest kluczowy w kontekście wpływu na ciągłość obróbki. Dłuższy czas przerwy między impulsami pozwala na skuteczniejsze schłodzenie zarówno strefy obróbki, jak i elektrody, co przyczynia się do stabilizacji procesu. Z kolei skrócenie czasu OFF, skutkuje większą liczbą wyładowań na jednostkę czasu, co prowadzi do szybszego nagrzewania się drutu oraz materiału, co zwiększa ryzyko niestabilności procesu i potencjalnie prowadzi do zerwania elektrody (Rys. 9.29).
- **CC:** Analiza wykazała, że niższe wartości natężenia prądu (CC) są powiązane z wyższym prawdopodobieństwem wystąpienia zakłóceń ciągłości procesu obróbki. Prawdopodobnie wynika to z problemów z inicjacją wyładowań elektrycznych przy niskim natężeniu, co skutkuje niestabilnością procesu i zwiększa ryzyko wystąpienia zwarć lub innych zakłóceń. Z kolei wyższe wartości CC mogą zapewniać bardziej stabilne wyładowania, co zmniejsza prawdopodobieństwo przerwania ciągłości obróbki (Rys. 9.27).
- **T:** Wyższe wartości naciągu drutu (T) powodują wzrost $P_{(pc)}$. Prawdopodobnie wynika to z faktu, że zbyt wysoki naciąg powoduje nadmierne obciążenie drutu, co zwiększa ryzyko jego przerwania (Rys. 9.32).
- **ON:** Wyższe wartości tego parametru zwiększają prawdopodobieństwo zakłóceń ciągłości procesu $P_{(pc)}$. Zbyt długi czas trwania impulsu (ON) prowadzi do nadmiernego nagrzewania się strefy obróbki i elektrody, co może powodować niestabilność procesu, zwłaszcza gdy czas OFF jest zbyt krótki, aby zapewnić odpowiednie schłodzenie (Rys. 9.29).
- **SV:** Zarówno nadmiernie wysokie, jak i zbyt niskie wartości parametru SV zwiększają ryzyko przerwania ciągłości obróbki $P_{(pc)}$. W przypadku wysokich wartości SV, w połączeniu z wysokimi wartościami parametrów określających energię wyładowania, takich jak VM (Rys. 9.31), dochodzi do wzrostu energii wyładowań, co skutkuje przegrzewaniem elektrody, a w konsekwencji może prowadzić do destabilizacji procesu. Natomiast zbyt niskie wartości SV (Rys. 9.30) mogą powodować nadmierne wycofywanie się elektrody, co również zwiększa ryzyko destabilizacji obróbki.
- **VM:** Wyższe wartości napięcia wyładowania (VM) zwiększają $P_{(pc)}$. Zbyt wysokie napięcie może prowadzić do bardziej intensywnych wyładowań, co sprzyja

niestabilności procesu. Z kolei zbyt niskie wartości VM mogą utrudniać inicjację wyładowań, co również negatywnie wpływa na ciągłość obróbki (Rys. 9.31).

- **FR:** Zarówno zbyt niskie, jak i zbyt wysokie natężenie przepływu płynu dielektrycznego wpływa na ciągłość obróbki. Zbyt małe natężenie prowadzi do gromadzenia się produktów erozji w strefie obróbki, co zwiększa prawdopodobieństwo zakłóceń, takich jak zwarcia i może skutkować przerwaniem procesu. Z kolei zbyt duże natężenie przepływu może destabilizować wyładowania, co również negatywnie wpływa na ciągłość obróbki (Rys. 9.32).

9.5.4. Wpływ parametrów technologicznych na szerokość szczeliny erozyjnej

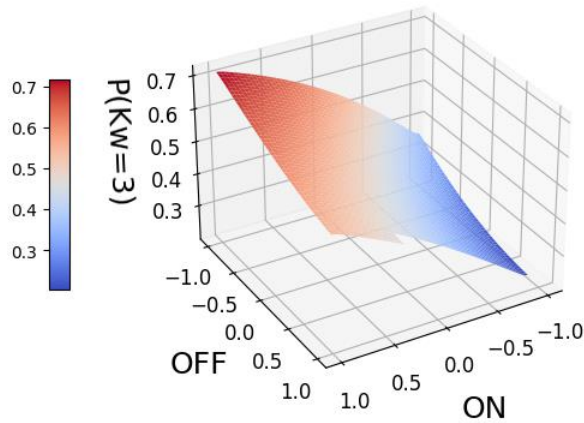
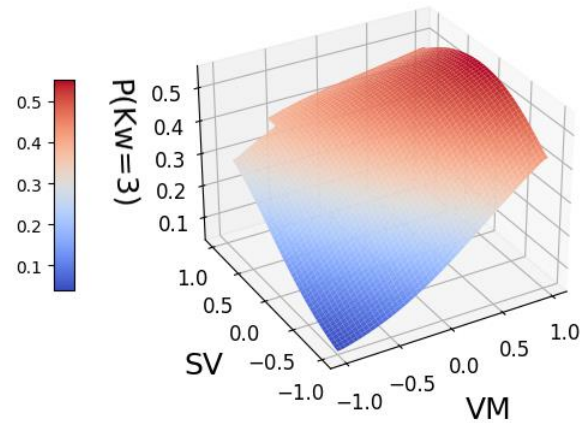
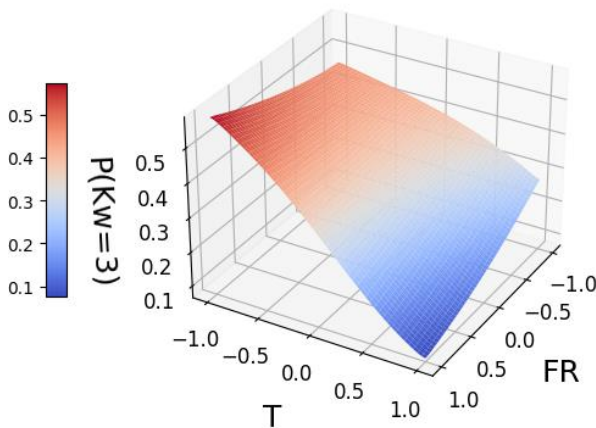
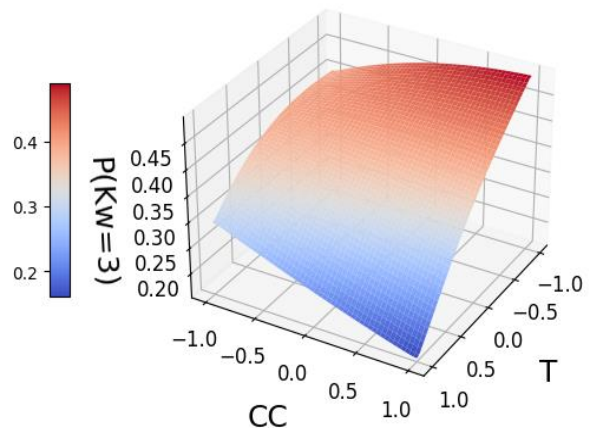
Ze względu na to, że model Kw jest klasyfikatorem wieloklasowym (3 klasy szerokości szczeliny erozyjnej), przedstawienie wykresu punktowego wartości SHAP dla każdej z klas byłoby problematyczne z powodu złożoności interpretacji wyników. Wartości SHAP dla klasyfikatorów wieloklasowych wymagają oddzielnego obliczenia wpływu dla każdej klasy, co prowadzi do trudności w jednoznacznym zaprezentowaniu wpływu parametrów w jednej przestrzeni wykresu. Dodatkowo, przy trzech klasach, wizualizacja wartości SHAP w formie punktowej mogłaby być nieczytelna i utrudniałaby dokładną interpretację wpływu poszczególnych zmiennych na różne klasy szerokości szczeliny erozyjnej. W związku z tym zdecydowano się na przedstawienie jedynie średnich wartości SHAP, które bardziej przystępnie obrazują globalny wpływ parametrów technologicznych na wynik modelu (Rys. 9.36).



Rys. 9.36 Wykres uśrednionych wartości SHAP dla poszczególnych zmiennych niezależnych modelu Kw

Aby przedstawić wpływ parametrów na szerokość szczeliny erozyjnej, podobnie jak w przypadku ciągłości obróbki, wprowadzono wskaźnik $P_{(Kw=3)}$, który odzwierciedla

prawdopodobieństwo uzyskania szczeliny o klasie Kw=3. Wykresy trójwymiarowe Rys. 9.37 - Rys. 9.40 pokazują zależności między wybranymi parametrami a tym wskaźnikiem, umożliwiając wizualizację wpływu parametrów na szerokość szczeliny erozyjnej.

Rys. 9.37 Wpływ OFF i ON na $P_{(Kw=3)}$ Rys. 9.38 Wpływ SV i VM na $P_{(Kw=3)}$ Rys. 9.39 Wpływ FR i T na $P_{(Kw=3)}$ Rys. 9.40 Wpływ CC i T na $P_{(Kw=3)}$

Na podstawie powyższych wyników oraz analizy literatury sformułowano wnioski:

- **ON:** Parametr ON jest najważniejszym parametrem w kontekście wpływu na szerokość szczeliny erozyjnej. Wysoka wartość ON skutkuje poszerzeniem szczeliny. Dłuższy czas impulsu dostarcza więcej energii do przestrzeni roboczej, co intensyfikuje proces erozji i powoduje większe usuwanie materiału (Rys. 9.37).
- **OFF:** Niska wartość parametru OFF powoduje rozszerzenie szczeliny erozyjnej. Krótszy czas OFF skutkuje wyższą częstotliwością wyładowań, co prowadzi do efektu analogicznego jak przy wysokim ON, więcej energii dostarczonej do przestrzeni roboczej powoduje poszerzenie szczeliny erozyjnej (Rys. 9.37).

- **CC:** Ponownie wpływ tego parametru analogiczny do wpływu ON. Wyższe wartości natężenia prądu CC dostarczają więcej energii na każde wyładowanie, co prowadzi do bardziej intensywnej erozji i poszerzenia szczeliny erozyjnej (Rys. 9.40).
- **VM:** Podobnie jak CC, wyższe wartości napięcia wyładowania VM zwiększają energię dostarczaną podczas wyładowań, co intensyfikuje proces erozyjny i skutkuje szerszą szczeliną (Rys. 9.38).
- **T:** Większy naciąg drutu redukuje wibracje, co pozwala na bardziej kontrolowane wyładowania, prowadząc do węższej szczeliny. Mniejszy naciąg sprzyja większym wibracjom, które mogą powodować bardziej rozproszone wyładowania, co prowadzi do szerszej szczeliny (Rys. 9.39).
- **FR:** Niskie natężenie przepływu dielektryka sprzyja gromadzeniu się produktów erozji w strefie obróbki, co powoduje zakłócenia w procesie usuwania materiału i zwężenie szczeliny. Wyższe natężenie skuteczniej usuwa produkty erozji, co poszerza szerokość szczeliny (Rys. 9.39).
- **SV:** Zwiększenie wartości napięcia serwomechanizmu prowadzi do intensyfikacji wyładowań elektrycznych, co skutkuje efektywniejszym usuwaniem materiału oraz poszerzeniem szczeliny erozyjnej (Rys. 9.38).

9.6. Podsumowanie badań modelowych

Przeprowadzone badania modelowe procesu cięcia elektroerozyjnego pozwoliły zidentyfikować modele umożliwiające przewidywanie wartości zmiennych zależnych: chropowatości powierzchni (Ra), prędkości usuwania materiału (MRR), szerokości szczeliny erozyjnej (Kw) oraz ciągłości procesu obróbki (C).

Na podstawie przeprowadzonych analiz porównawczych ustalono, że najbardziej odpowiednim modelem do przewidywania chropowatości powierzchni (Ra) są maszyny wektorów nośnych (SVR) z liniową funkcją jądrową. W przypadku prędkości usuwania materiału (MRR) najlepsze wyniki uzyskano z zastosowaniem modelu XGBoost. Model SVC okazał się najskuteczniejszy w przewidywaniu szerokości szczeliny erozyjnej (Kw), natomiast w klasyfikacji ciągłości procesu (C) najwyższą skuteczność osiągnął model XGBoost.

Wyniki badań pozwoliły również na analizę wpływu parametrów technologicznych na wartości zmiennych zależnych. Dla chropowatości powierzchni (Ra) zidentyfikowano, że czas trwania impulsu (ON) był kluczowym czynnikiem wpływającym na wynik. W przypadku prędkości usuwania materiału (MRR) najistotniejszym czynnikiem okazał się czas przerwy między impulsami (OFF). Analiza ciągłości obróbki (C) wykazała największy wpływ czasu przerwy (OFF), natężenia prądu (CC) oraz naciąg drutu (T). W odniesieniu do szerokości szczeliny erozyjnej (Kw) zaobserwowano, że parametry ON, OFF oraz CC miały największy wpływ na jej wartość.

10. OPTIMALIZACJA PROCESU WEDM

Niniejszy rozdział poświęcony jest optymalizacji parametrów technologicznych obróbki elektroerozyjnej z wykorzystaniem różnych algorytmów optymalizacyjnych. W ramach tego rozdziału omówiono kolejne etapy optymalizacji, począwszy od wyboru kryteriów i funkcji celu, poprzez konfigurację algorytmów, aż po analizę uzyskanych wyników.

Rozdział rozpoczyna się od przeglądu kryteriów optymalizacji i funkcji celu (p.10.1), z naciskiem na elastyczność ich konfiguracji w odpowiedzi na zmieniające się warunki produkcyjne. W kolejnych częściach rozdziału (p.10.2 i p.10.3) przedstawiono konfigurację parametrów i przestrzeń poszukiwań dla rozpatrywanych algorytmów, tj. algorytm genetyczny, algorytm różnicowej ewolucji, optymalizacja roju cząstek oraz algorytm mrówkowy (opis teoretycznych tych algorytmów znajduje się p. 5.3).

W dalszej części rozdziału (p.10.4) przeprowadzono szczegółową analizę porównawczą wybranych algorytmów optymalizacyjnych. Analiza ta obejmuje ocenę efektywności algorytmów w kontekście czasu obliczeniowego, stabilności wyników oraz jakości uzyskanych rozwiązań.

Następnie, w p.10.5, przedstawiono wyniki optymalizacji przeprowadzonej z wykorzystaniem wybranego algorytmu, skoncentrowanej na różnych zadaniach optymalizacyjnych, takich jak minimalizacja chropowatości powierzchni, czy maksymalizacja prędkości usuwania materiału. Wyniki te poddano analizie pod kątem jakości uzyskanych rozwiązań oraz ich zbieżności do globalnego optimum. Przeprowadzono również interpretację najważniejszych parametrów technologicznych odpowiadających znalezionym rozwiązaniom.

Na zakończenie rozdziału (p.10.6) przedstawiono weryfikację eksperymentalną rozwiązań uznanych za optymalne oraz wnioski wynikające z przeprowadzonych badań.

10.1. Kryteria optymalizacji

Dotychczas rozpatrywane zmienne zależne zostały uwzględnione w kryteriach optymalizacji procesu. W celu zapewnienia odpowiedniej elastyczności tych kryteriów w odniesieniu do zmieniających się warunków produkcyjnych zaimplementowano funkcję, która umożliwi definiowanie celów optymalizacji dla każdego kryterium oraz regulację ich istotności. Regulacja ta odbywa się poprzez modyfikację wag przypisanych do poszczególnych kryteriów. Tabela 10.1 przedstawia zdefiniowane w ramach badań kryteria, możliwe cele optymalizacji oraz zakres wartości wag, które mogą być przypisane dla każdego kryterium.

Tabela 10.1 Kryteria optymalizacji procesu WEDM

Kryterium optymalizacji	Cel optymalizacji	Zakres możliwych wag
Chropowatość powierzchni obrabianej (Ra)	Minimalizacja różnicy między przewidywaną a docelową wartością chropowatości	[0, 1, 2, 3]
Wydajność obróbki (MRR)	Maksymalizacja	[0, 1, 2, 3]
Szerokość szczeliny erozyjnej (Kw)	Minimalizacja różnicy między przewidywaną a docelową wartością szerokości szczeliny	[0, 1, 2, 3]

Cele optymalizacyjne przedstawione w Tabeli 10.1, dotyczące parametru chropowatości powierzchni Ra oraz szerokości szczeliny erozyjnej (Kw), są ustalane na podstawie wartości docelowych. Dla parametru Ra można samodzielnie zdefiniować dowolną wartość docelową, a docelowa szerokość szczeliny erozyjnej może zostać wybrana spośród trzech dostępnych klas. W przypadku prędkości usuwania materiału (MRR) algorytm optymalizacyjny dąży do maksymalizacji tej wartości. Dla każdego z tych celów istnieje możliwość przypisania wagi, co pozwala na dostosowanie ich istotności, zgodnie z następującą skalą:

- 0 – cel nieistotny,
- 1 – cel mało istotny,
- 2 – cel o średniej istotności,
- 3 – cel o wysokiej istotności.

10.2. Funkcja celu

Kluczowym elementem optymalizacji jest odpowiednio zdefiniowana funkcja celu. W poprzednim punkcie przedstawiono kilka kryteriów optymalizacji, a zatem funkcja celu powinna spełniać wymogi optymalizacji wielokryterialnej. Spośród wielu metod optymalizacji wielokryterialnej, zdecydowano się na podejście oparte na formule ważonej (znane również jako podejście priori) [74]. Wybór ten był podyktowany prostotą implementacji tak sformułowanej funkcji celu oraz możliwością łatwej adaptacji wag dla różnych kryteriów optymalizacyjnych.

W badaniach optymalizacja polegała na minimalizacji funkcji celu sformułowanej jako kombinacja liniowa:

$$f(x) = \tilde{\omega}_{Ra} \cdot \Delta Ra(x) - \tilde{\omega}_{MRR} \cdot MRR(x) + \tilde{\omega}_{Kw} \cdot \Delta Kw(x) + 3 \cdot C(x), \quad (10.1)$$

gdzie:

- x – wektor zmiennych niezależnych (wartości parametrów technologicznych),
- $\tilde{\omega}_{Ra}$, $\tilde{\omega}_{MRR}$, $\tilde{\omega}_{Kw}$ - znormalizowane wagi odpowiednio dla kryterium Ra , MRR oraz Kw . Normalizacja wag polegała na przekształceniu pierwotnych wag ω_{Ra} , ω_{MRR} , ω_{Kw} tak, aby ich suma była równa 1 (10.2 (10.5).

$$\tilde{\omega}_{Ra} = \frac{\omega_{Ra}}{\omega_{Ra} + \omega_{MRR} + \omega_{Kw}} \quad (10.2)$$

$$\tilde{\omega}_{MRR} = \frac{\omega_{MRR}}{\omega_{Ra} + \omega_{MRR} + \omega_{Kw}} \quad (10.3)$$

$$\tilde{\omega}_{Kw} = \frac{\omega_{Kw}}{\omega_{Ra} + \omega_{MRR} + \omega_{Kw}} \quad (10.4)$$

$$\tilde{\omega}_{Ra} + \tilde{\omega}_{MRR} + \tilde{\omega}_{Kw} = 1 \quad (10.5)$$

- $\Delta Ra(x)$ - różnica pomiędzy przewidywaną wartością $Ra(x)$ a wartością docelową T_{Ra} :

$$\Delta Ra(x) = |Ra(x) - T_{Ra}| \quad (10.6)$$

- $\Delta Kw(x)$ - różnica pomiędzy przewidywaną klasą $Kw(x)$ a klasą docelową T_{Kw} :

$$\Delta Kw(x) = |Kw(x) - T_{Kw}| \quad (10.7)$$

Wszystkie zmienne niezależne przewidywane przez modele oraz wartości docelowe tych zmiennych zostały znormalizowane przy użyciu metody min-max (6.4).

$$Ra(x) = \frac{Ra(x) - 0}{7 - 0}, \quad (10.8)$$

$$MRR(x) = \frac{MRR(x) - 0}{11 - 0}, \quad (10.9)$$

$$KW(x) = \frac{KW(x) - 1}{3 - 1}. \quad (10.10)$$

Poszczególne wartości maksymalne zmiennych Ra i MRR zostały dobrane na podstawie wstępnych badań optymalizacyjnych, które uwzględniały zarówno przeprowadzone i opisane wcześniej badania doświadczalne, jak i doniesienia literaturowe.

Dodatkowo, w celu uwzględnienia ciągłości obróbki, wprowadzono warunek, zgodnie z którym, jeśli przewidywana wartość wskaźnika ciągłości obróbki C przyjmuje wartość równą 1, co wskazuje na brak ciągłości procesu, funkcja celu przyjmuje wartość co najmniej 3 (maksymalna możliwa wartość funkcji celu) co eliminuje to rozwiązanie z dalszej optymalizacji.

10.3. Przestrzeń poszukiwań

Przestrzeń poszukiwań w algorytmach optymalizacyjnych definiuje zbiór wszystkich możliwych rozwiązań, które mogą być rozważane podczas procesu optymalizacji. W rozpatrywanym przypadku optymalizacji parametrów technologicznych, przestrzeń ta została ograniczona przez dopuszczalne wartości parametrów, wynikające z technicznych specyfikacji maszyny. W Tabeli 10.2 przedstawiono przedziały wartości dla poszczególnych parametrów.

Tabela 10.2 Zakresy wartości dla parametrów technologicznych

Zmienna	Jednostka	Minimalna wartość	Maksymalna wartość
VS	V	1	10
CC	A	1	13
VM	V	4	32
ON	μs	1	30
OFF	μs	6	200
SV	V	2	255
SVG	-	1	2000
WP1A	-	1	38
WP1B	-	1	30
WP2A	-	1	10
WP2B	-	1	18
T	gf	10	2550
WF	m/min	1	15
FR	l/min	1	18

10.4. Analiza porównawcza algorytmów optymalizacji

10.4.1. Kryteria porównania

W analizie porównawczej przyjęto zestaw kryteriów, które umożliwiają ocenę efektywności algorytmów optymalizacyjnych w procesie optymalizacji obróbki elektroerozyjnej:

- **Czas obliczeniowy t_c :** Czas potrzebny do uzyskania zbieżności algorytmu lub osiągnięcia określonego progu jakości rozwiązania. Jest to istotne kryterium z perspektywy praktycznego zastosowania algorytmów w rzeczywistych procesach produkcyjnych.
- **Wartość funkcji celu FC :** Ponieważ w każdym z analizowanych algorytmów wykorzystano tę samą postać funkcji celu, jakość rozwiązania oceniana jest na podstawie wartości tej funkcji. Dzięki temu można bezpośrednio porównać efektywność poszczególnych algorytmów pod względem realizacji założonych celów optymalizacyjnych – niższa wartość funkcji celu oznacza lepszą jakość uzyskanego rozwiązania zgodnie z przyjętymi kryteriami.
- **Stabilność wyników:** Zdolność algorytmu do uzyskiwania podobnych wyników przy wielokrotnym uruchamianiu procesu optymalizacji z różnymi inicjalizacjami. W pracy stabilność jest mierzona poprzez odchylenia standardowe przewidywanych wartości zmiennych niezależnych oraz wartości funkcji celu.

10.4.2. Konfiguracja algorytmów

Wybór parametrów dla każdego z algorytmów został dokonany na podstawie iteracyjnego podejścia prób i błędów, w którym oceniano stabilność rozwiązań oraz ich jakość. W Tabeli 10.3 przedstawiono konfigurację parametrów algorytmu genetycznego (GA), gdzie kluczowe znaczenie miały parametry takie jak rozmiar populacji, liczba generacji, prawdopodobieństwo krzyżowania oraz mutacji. Tabela 10.4 zawiera konfigurację algorytmu różnicowej ewolucji (DE), w której określono strategię krzyżowania oraz inne parametry kontrolujące proces ewolucji. Kolejne tabele, Tabela 10.5 i Tabela 10.6, prezentują odpowiednio parametry dla algorytmu optymalizacji roju cząstek (PSO) oraz algorytmu mrówkowego (ACO), gdzie kluczowymi zmiennymi były rozmiar roju, liczba iteracji oraz współczynniki wpływu i rozpadu feromonów.

Tabela 10.3 Zastosowana konfiguracja parametrów algorytmu genetycznego GA

Rozmiar populacji	500
Liczba generacji	200
Prawdopodobieństwo krzyżowania	0.5
Prawdopodobieństwo mutacji	0.2
Metoda selekcji	Selekcja turniejowa (ang. tournament selection) z rozmiarem turnieju 5
Metoda krzyżowania	Krzyżowanie jednolite (ang. uniform crossover)
Metoda mutacji	Mutacja jednolita (ang. uniform mutation)

Tabela 10.4 Zastosowana konfiguracja parametrów algorytmu ewolucji różnicowej DE

Strategia krzyżowania	Krzyżowanie binarne (ang. best1bin strategy)
Maksymalna liczba iteracji	130
Wielkość populacji	10
Tolerancja	0.01

Tabela 10.5 Zastosowana konfiguracja parametrów algorytmu optymalizacji roju cząstek PSO

Rozmiar roju	300
Maksymalna liczba iteracji	100
Minimalny krok	1e-8
Minimalna wartość funkcji	1e-8
Inercja	0.4
Współczynnik przyciągania do najlepszej pozycji cząstki	0.7

Tabela 10.6 Zastosowana konfiguracja parametrów algorytmu mrówkowego ACO

Liczba mrówek	350
Liczba iteracji	100
Współczynnik rozpadu feromonów	1.7
Współczynnik wpływu feromonów	Początkowo 0.5, dynamicznie zmniejszany do 0.1
Współczynnik wpływu heurystyki	Początkowo 1.0, dynamicznie zwiększany do 2.0
Inicjalizacja feromonów	Wszystkie feromony na poziomie 1.0
Aktualizacja feromonów	Po każdej iteracji na podstawie najlepszego rozwiązania
Dodatkowa losowość	Dodawana co 5 iteracji, aby zwiększyć różnorodność

10.4.3. Porównanie wyników optymalizacji

Wyniki przedstawione w Tabeli 10.7 zostały uzyskane w wyniku optymalizacji z zastosowaniem algorytmów GA, PSO, DE oraz ACO. Każdy z tych algorytmów był testowany w analogicznych warunkach początkowych, z zastosowaniem tych samych parametrów technologicznych, kryteriów optymalizacji oraz funkcji celu w każdej próbie.

Eksperymenty przeprowadzono dla siedmiu wybranych zadań optymalizacyjnych, które obejmowały różne cele optymalizacji oraz odpowiadające im wartości wag przypisanych do poszczególnych kryteriów.

Dla każdego zadania wykonano pięć powtórzeń testów z różnymi wartościami początkowymi zmiennych niezależnych. Na podstawie wyznaczonych zestawów parametrów technologicznych dokonano przewidywania wartości zmiennych zależnych przy użyciu opracowanych modeli procesu, a następnie obliczono ich średnie wartości $\overline{Ra(x)}$ i $\overline{MRR(x)}$, oraz odchylenia standardowe $\sigma Ra(x)$ i $\sigma MRR(x)$, a także średni błąd kwadratowy $MSE_{Kw(x)}$, obliczony na podstawie różnic pomiędzy przewidywanymi wartościami Kw , a wartościami docelowymi T_{Kw} . Obliczono również średnią wartość funkcji celu \overline{FC} dla każdego algorytmu w danym zadaniu optymalizacyjnym, odchylenie standardowe funkcji celu σFC oraz średni czas obliczeń \bar{t}_c . Algorytmy, które osiągnęły najmniejszą wartość \overline{FC} dla danego zadania optymalizacyjnego, zostały wyróżnione kolorem zielonym.

Tabela 10.7 Porównanie wyników optymalizacji z wykorzystaniem algorytmów GA, PSO, DE oraz ACO

Kryteria	Wagi	Algorytm	Przewidywane wartości zmiennych zależnych					\overline{FC}	σFC	\bar{t}_c [s]
			$\overline{Ra(x)}$	$\sigma Ra(x)$	$\overline{MRR(x)}$	$\sigma MRR(x)$	$MSE_{Kw(x)}$			
$T_{Ra}: 2.5$ MRR: max $T_{Kw}: 1$	$\omega_{Ra}: 3$	GA	2.499	0	2.17	0.631	2.8	0	0	90.2
		DE	2.499	0	2.425	0.409	2.2	0	0	80.2
	$\omega_{MRR}: 0$	PSO	2.0	0	3.377	1.759	2.8	0	0	70
		ACO	1.999	0	2.566	0.461	3.4	0	0	74
$T_{Ra}: 2.5$ MRR: max $T_{Kw}: 1$	$\omega_{Ra}: 0$	GA	3.808	0.036	10.79	0	4	-0.98	0	94.7
		DE	3.79	0.098	10.764	0.017	4	-0.978	0.001	87.9
	$\omega_{MRR}: 3$	PSO	3.634	0.86	9.254	1.059	1.8	-0.841	0.096	73.1
		ACO	3.571	0.35	9.525	0.331	1.8	-0.865	0.030	75.2
$T_{Ra}: 5.2$ MRR: max	$\omega_{MRR}: 0$	GA	3.926	0.399	3.446	0.84	0	0	0	94.7
		DE	4.103	0.608	3.654	0.305	0	0	0	86.7
		PSO	3.651	1.172	3.425	1.171	0	0	0	74

$T_{KW}: 3$	$\omega_{KW}: 3$	ACO	3.803	1.061	3.942	0.937	0	0	0	80
$T_{Ra}: 2.7$ MRR: max $T_{KW}: 2$	$\omega_{Ra}: 3$	GA	2.699	0	10.268	0.118	1	-0.466	0.005	93.1
		DE	2.7	0	10.268	0.115	1	-0.466	0.005	90
	$\omega_{MRR}: 3$	PSO	2.7	0	8.159	1.172	0.8	-0.37	0.053	77
		ACO	3.1	0.409	9.252	0.365	1	-0.386	0.008	77
$T_{Ra}: 2.0$ MRR: max $T_{KW}: 1$	$\omega_{Ra}: 1$	GA	2.481	0.43	9.415	1.126	0	-0.357	0.085	92.4
		DE	2.422	0.041	10.37	0.028	0	-0.395	0	89.1
	$\omega_{MRR}: 3$	PSO	2.485	0.35	7.279	2.297	0	-0.273	0.09	75.6
		ACO	2.762	0.481	8.349	0.862	0	-0.309	0.028	76.4
$T_{Ra}: 6.0$ MRR: max $T_{KW}: 1$	$\omega_{Ra}: 3$	GA	5.771	0.511	6.877	1.913	0	-0.197	0.033	92.4
		DE	5.922	0.107	7.228	1.625	0	-0.215	0.044	90
	$\omega_{MRR}: 3$	PSO	5.999	0	5.718	0.515	0	-0.173	0.015	85.8
		ACO	4.311	0.983	8.027	2.195	0	-0.162	0.027	80
$T_{Ra}: 5.2$ MRR: max $T_{KW}: 1$	$\omega_{Ra}: 3$	GA	5.062	0.188	9.038	1.771	0	-0.264	0.046	92
		DE	5.199	0	6.047	0.024	0	-0.183	0	92.2
	$\omega_{MRR}: 2$	PSO	5.19	0	7.2	0.763	0.8	-0.184	0.004	63.8
		ACO	4.927	0.455	6.295	1.208	0	-0.168	0.024	81

Na podstawie przedstawionych wyżej wyników obliczono statystyki podsumowujące wyniki analizy porównawczej algorytmów optymalizacji. W Tabeli 10.8 zestawiono sumę średnich wartości funkcji celu uzyskanych z wykorzystaniem danego algorytmu optymalizacji $\Sigma\overline{FC}$, która służy jako wskaźnik ogólnej efektywności algorytmu w rozpatrywanych zadaniach. Zestawiono również średnie odchylenie standardowe funkcji celu $\overline{\sigma FC}$, średni czas ze średnich czasów obliczeń \overline{t}_c , a także odsetek zadań optymalizacyjnych, dla których dany algorytm uzyskał najlepsze rozwiązanie w porównaniu do pozostałych algorytmów P_{opt} .

Tabela 10.8 Podsumowanie analizy porównawczej algorytmów optymalizacji

Algorytm	$\Sigma\overline{FC}$	$\overline{\sigma FC}$	$\overline{t}_c [s]$	$P_{opt} [\%]$
GA	-2.264	0.024	92.78	71.4
DE	-2.237	0.007	88.01	71.4
PSO	-1.841	0.036	74.18	28.5
ACO	-1.89	0.016	77.65	28.5

Powyższe zestawienie wskazuje na zbliżoną efektywność algorytmu genetycznego (GA) i algorytmu ewolucji różnicowej (DE). Algorytmy te wykazują jedynie nieznaczne różnice w zakresie minimalizacji funkcji celu, zbliżoną częstotliwość znajdowania rozwiązań uznanych za optymalne, wysoką stabilność wyników oraz porównywalny czas obliczeń. Niemniej jednak, algorytm GA osiąga nieznacznie lepsze wyniki, co znajduje odzwierciedlenie w niższej sumarycznej wartości funkcji celu. Zastosowanie algorytmów rojowych, takich jak PSO i ACO, również prowadzi do znaczącej minimalizacji funkcji celu, jednak w mniejszym stopniu niż GA i DE.

W związku z tym, do dalszych badań zdecydowano się na wybór algorytmu genetycznego jako najefektywniejszego spośród rozważanych metod.

10.5. Analiza wybranych wyników optymalizacji

W celu pogłębionej weryfikacji przeprowadzonej optymalizacji, dokonano analizy trzech wybranych wyników optymalizacji uzyskanych przy zastosowaniu GA. Kryterium selekcji wyników było ściśle powiązane z wymaganiami wdrożeniowymi, co pozwoliło na wybór rezultatów o największym potencjale aplikacyjnym z perspektywy potrzeb przedsiębiorstwa.

Poniżej przedstawiono wybrane kryteria optymalizacji, zastosowane wartości wag poszczególnych kryteriów, przewidywane wartości zmiennych zależnych na podstawie parametrów technologicznych wyznaczonych przez algorytm, a także wartości parametrów wyznaczone jako optymalne. Dodatkowo, zaprezentowano wykres ilustrujący zmienność wartości funkcji celu w zależności od kolejnych generacji algorytmu genetycznego.

Tabela 10.9 przedstawia wyniki optymalizacji ukierunkowanej na minimalizację chropowatości powierzchni obrabianej. Cel ten zrealizowano poprzez określenie minimalnej wartości docelowej parametru chropowatości R_a (T_{Ra}), przypisaniu temu kryterium największej wagi oraz wykluczeniu innych kryteriów przypisując im zerową wagę.

Tabela 10.9 Wyniki optymalizacji ukierunkowanej na minimalizację chropowatości powierzchni obrabianej

Kryteria optymalizacji	Wartości wag	Przewidywane wartości											
$T_{Ra}: 2.5$ MRR: max $T_{KW}: 0$	$\omega_{Ra}: 3$ $\omega_{MRR}: 0$ $\omega_{KW}: 0$	$Ra(x): 2.5$ MRR(x): 2.31 $Kw(x): 1$											
Wartości parametrów wyznaczone jako optymalne													
VS	CC	VM	ON	OFF	SV	SVG	WP1A	WP1B	WP2A	WP2B	T	WF	FR
4	5	13	3	146.8	24	223	26	26	9	17	1077	11	17

Analizując wykres funkcji celu z Tabela 10.9 można zauważyć, że po początkowych fluktuacjach, od około 22. generacji następuje gwałtowny spadek wartości funkcji celu, co wskazuje na szybkie zbliżanie się algorytmu do optymalnego rozwiązania. Po tym etapie, wartość funkcji celu stabilizuje się, a dalsze zmiany są minimalne. W końcowych generacjach, między 125. a 200., funkcja celu osiąga niemal zerową wartość. Stabilność tej wartości wskazuje na skuteczność algorytmu w zbieżności do najlepszego rozwiązania i sugeruje, że algorytm znalazł optimum globalne.

Wartość przewidywana $Ra(x)$ jest równa wartości docelowej, co potwierdza skuteczność zastosowanego algorytmu optymalizacyjnego. W parametrach technologicznych wyznaczonych przez algorytm szczególną uwagę należy zwrócić na wyjątkowo niską wartość parametru ON. Jak wykazano w punkcie 9.5.1 podczas badań modelowych, parametr ten odgrywa kluczową rolę w kształtowaniu chropowatości powierzchni. Algorytm skutecznie zidentyfikował tę zależność, co znajduje odzwierciedlenie w minimalnej wartości ON. W przypadku tak niskiej wartości ON, pozostałe parametry mają relatywnie mniejsze znaczenie. Prawdopodobnie algorytm dążył do znalezienia rozwiązania, które zapewni równocześnie odpowiednią ciągłość procesu obróbki, co potwierdza wysoka wartość parametru OFF, mającego kluczowe znaczenie w kontekście wpływu na ciągłość obróbki, jak opisano w p.9.5.3.

W Tabela 10.10 przedstawiono wyniki optymalizacji ukierunkowanej na maksymalizację MRR. Wykres funkcji celu w tym przypadku wskazuje, że po początkowych wahaniach, od około 5. generacji następuje szybki spadek wartości funkcji celu, stabilizując się na poziomie 25. generacji. W dalszych generacjach, od 50. do 200.,

funkcja celu pozostaje stabilna, co sugeruje skuteczne osiągnięcie globalnego optimum przez algorytm.

Tabela 10.10 Wyniki optymalizacji ukierunkowanej na maksymalizację prędkości usuwania materiału

Kryteria optymalizacji	Wartości wag	Przewidywane wartości											
$T_{Ra}: 2.0$	$\omega_{Ra}: 0$	$Ra(x): 3.68$											
$T_{MRR}: \max$	$\omega_{MRR}: 3$	$MRR(x): 10.8$											
$T_{KW}: 0$	$\omega_{KW}: 0$	$Kw(x): 2$											
Wartości parametrów wyznaczone jako optymalne													
VS	CC	VM	ON	OFF	SV	SVG	WP1A	WP1B	WP2A	WP2B	T	WF	FR
3	12	32	12	18.9	19	306	2	4	2	9	1390	7	11

Wartości parametrów technologicznych wyznaczone przez algorytm wskazują na zdecydowane ukierunkowanie na maksymalizację prędkości usuwania materiału. Wyznaczona maksymalna możliwa wartość VM i bardzo niska wartość OFF jednoznacznie wpływają na zwiększenie MRR, co potwierdzono w punkcie 9.5.2. Jednocześnie, niemal maksymalna wartość CC ogranicza prawdopodobieństwo przzerwania ciągłości obróbki – jest to drugi w kolejności parametr pod względem istotności wpływu na zachowanie ciągłości obróbki.

Kolejnym zadaniem optymalizacyjnym było znalezienie parametrów, które zapewnią maksymalną prędkość usuwania materiału, przy jednoczesnym zachowaniu chropowatości powierzchni na ustalonym, niskim poziomie. Wyniki zaprezentowano w Tabela 10.10.

Tabela 10.11 Wyniki optymalizacji ukierunkowanej na maksymalizację prędkości usuwania materiału, przy jednoczesnym zachowaniu chropowatości powierzchni na ustalonym poziomie

Kryteria optymalizacji	Wartości wag	Przewidywane wartości											
$T_{Ra}: 2.5$	$\omega_{Ra}: 2$	$Ra(x): 3.5$											
$T_{MRR}: \max$	$\omega_{MRR}: 3$	$MRR(x): 8.4$											
$T_{KW}: 0$	$\omega_{KW}: 0$	$Kw(x): 2$											
Wartości parametrów wyznaczone jako optymalne													
VS	CC	VM	ON	OFF	SV	SVG	WP1A	WP1B	WP2A	WP2B	T	WF	FR
3	12	18	1	11.5	13	291	11	6	3	3	1475	7	12

Na wykresie funkcji celu można zauważyć, że proces optymalizacji przebiegł efektywnie. Algorytm znalazł rozwiązanie już na etapie ok 25. generacji. Dalsza stabilizacja wartości funkcji celu sugeruje znalezienie globalnego optimum.

Przewidywana wartość Ra wynosząca 3,5 μm , w porównaniu do założonej wartości docelowej 2,5 μm , istotnie odbiega od założonego celu. Jednakże, biorąc pod uwagę niższą przypisaną wagę dla tego kryterium, wynik ten jest uzasadniony. Najwyższą wagę przypisano maksymalizacji MRR – w tym przypadku przewidywana wartość wynosząca 8,41 odbiega od wcześniej wyznaczonej wartości maksymalnej 10.8, jednak biorąc pod uwagę sprzeczne cele optymalizacji (minimalizacja chropowatości i maksymalizacja prędkości usuwania materiału) wynik ten również jest uzasadniony. Wyniki te potwierdzają prawidłową reakcję algorytmu na przyjęte wagi kryteriów.

Podobnie jak w poprzednich analizowanych zadaniach optymalizacyjnych, algorytm preferuje niskie wartości parametrów ON i VM w celu minimalizacji chropowatości powierzchni, a także niską wartość parametru OFF, co sprzyja zapewnieniu wysokiego wskaźnika MRR. Z kolei wysoka wartość parametru CC zapewnia mniejsze prawdopodobieństwo przerwania ciągłości procesu obróbki.

10.6. Weryfikacja eksperymentalna wyników optymalizacji

W celu ostatecznej weryfikacji opracowanego rozwiązania zdefiniowano trzy odrębne zadania optymalizacyjne, oparte na wybranych kryteriach i wagach. Należy podkreślić, że weryfikacja poszczególnych modeli oraz algorytmów optymalizacji została szeroko opisana w poprzednich rozdziałach. Stąd też niniejszy podrozdział stanowi próbę podsumowania i uogólnienia przeprowadzanych badań.

Algorytm optymalizacyjny zostały zastosowany w celu wyznaczenia optymalnych zestawów parametrów technologicznych, które miały zapewnić uzyskanie zadanych wartości zmiennych zależnych. Na podstawie tych parametrów wycięto próbki testowe, a następnie dokonano pomiarów chropowatości powierzchni (Ra), prędkości usuwania materiału (MRR) oraz szerokości szczeliny erozyjnej (Kw). Otrzymane wyniki pomiarów porównano z przewidywaniami modeli. Wyniki weryfikacji przedstawiono w Tabeli 10.12.

Tabela 10.12 Wyniki weryfikacji eksperymentalnej przeprowadzonej optymalizacji

Kryteria		Wartości wag		Przewidywane wartości zmiennych zależnych			Zmierzone wartości zmiennych zależnych			Błąd przewidywania			
Zadanie nr 1													
$T_{Ra}: 2.0$ $MRR: \max$ $T_{Kw}: 1$		$\omega_{Ra}: 3$ $\omega_{MRR}: 3$ $\omega_{Kw}: 0$		\widehat{Ra}	\widehat{MRR}	\widehat{KW}	Ra	MRR	KW	$MAPE_{Ra}$		$MAPE_{MRR}$	
				2.155	6.4	1	2.28	7.1	1	5.48%		9.86%	
Wartości parametrów wyznaczone jako optymalne													
VS	CC	VM	ON	OFF	SV	SVG	WP1A	WP1B	WP2A	WP2B	T	WF	FR
3	12	18	1	11.5	13	291	11	6	3	3	1475	7	12
Zadanie nr 2													
$T_{Ra}: 5.2$ $MRR: \max$ $T_{Kw}: 2$		$\omega_{Ra}: 3$ $\omega_{MRR}: 2$ $\omega_{Kw}: 1$		\widehat{Ra}	\widehat{MRR}	\widehat{KW}	Ra	MRR	KW	$MAPE_{Ra}$		$MAPE_{MRR}$	
				4.051	9.57	2	3.879	10.2	2	4.43%		6.18%	
Wartości parametrów wyznaczone jako optymalne													
VS	CC	VM	ON	OFF	SV	SVG	WP1A	WP1B	WP2A	WP2B	T	WF	FR
4	10	16	25	8.8	20	416	3	10	10	13	1002	9	16
Zadanie nr 3													
$T_{Ra}: 2.0$ $MRR: \max$ $T_{Kw}: 1$		$\omega_{Ra}: 1$ $\omega_{MRR}: 3$ $\omega_{Kw}: 0$		\widehat{Ra}	\widehat{MRR}	\widehat{KW}	Ra	MRR	KW	$MAPE_{Ra}$		$MAPE_{MRR}$	
				3.46	10.2	1	4.23	11.29	1	18.2%		9.65%	
Wartości parametrów wyznaczone jako optymalne													
VS	CC	VM	ON	OFF	SV	SVG	WP1A	WP1B	WP2A	WP2B	T	WF	FR
3	8	14	10	7.4	11	287	4	4	3	15	1070	8	8

Analiza porównawcza wyników pomiarowych i modelowych umożliwiła sformułowanie ostatecznych wniosków na temat skuteczności zastosowanej metody optymalizacji:

- Zadanie optymalizacji nr 1:

Celem optymalizacji w tym zadaniu było wyznaczenie parametrów technologicznych, które zapewnią chropowatość powierzchni (Ra) na poziomie $2.0\ \mu\text{m}$, przy jednoczesnej maksymalizacji prędkości usuwania materiału (MRR). Obu kryteriom przypisano najwyższą wagę. Algorytm optymalizacyjny zidentyfikował zestaw parametrów, które zgodnie z przewidywaniami modeli powinny zapewnić chropowatość powierzchni (Ra) na poziomie $2.155\ \mu\text{m}$ oraz prędkość usuwania materiału (MRR) równą $6.4\ \text{mm}^3/\text{min}$. Przewidywana wartość Ra jest zatem nieznacznie wyższa niż wartość docelowa, co można uznać za uzasadnione ze względu na sprzeczność kryteriów optymalizacyjnych, wynikającą z jednoczesnej maksymalizacji MRR.

Zmierzona wartość Ra dla wyciętej próbki wyniosła $2.28\ \mu\text{m}$, co stanowi niewielkie odchylenie od wartości przewidywanej. Różnica ta może być wyjaśniona błędem przewidywania modelu, określonym w badaniach modelowych jako $\text{MAE} = 0.170\ \mu\text{m}$. Z kolei zmierzona wartość MRR wykazała większe odchylenie od wartości przewidywanej (MAPE 9.86%), jednakże ze względu na cel maksymalizacji MRR, wyższa wartość zmierzona ($7.1\ \text{mm}^3/\text{min}$) jest korzystnym wynikiem.

- Zadanie optymalizacji nr 2:

Celem optymalizacji w zadaniu nr 2 było uzyskanie chropowatości powierzchni (Ra) na poziomie $5.2\ \mu\text{m}$, maksymalizacji prędkości usuwania materiału (MRR), oraz zachowanie szerokości szczeliny erozyjnej (Kw) na poziomie 2. Priorytet przypisano Ra, co znajduje odzwierciedlenie w wyższej wadze nadanej temu kryterium.

Wartość Ra przewidywana przez model wyniosła $4.051\ \mu\text{m}$, MRR wyniosło $9.57\ \text{mm}^3/\text{min}$, a szerokość szczeliny erozyjnej (Kw) była przewidywana na poziomie 2. Zmierzone wartości wyniosły odpowiednio $3.879\ \mu\text{m}$ dla Ra, $10.2\ \text{mm}^3/\text{min}$ dla MRR oraz 2 dla Kw. Odchylenie w przypadku Ra było niewielkie (MAPE 4.43%), natomiast dla MRR nieco większe (MAPE 6.18%). Zmierzona wartość Ra okazała się niższa niż przewidywana, a także niższa od wartości docelowej optymalizacji, ustalonej na poziomie $5.2\ \mu\text{m}$. Różnica ta może wynikać z większego błędu przewidywania modelu Ra dla wyższych wartości, jak wspomniano w p. 9.1.3.

W przypadku szerokości szczeliny erozyjnej, zarówno przewidywana, jak i zmierzona wartość wyniosły 2, co wskazuje na dobrą jakość modelu. Wyższa niż przewidywana wartość MRR, mimo odchylenia, pozostaje w zgodzie z celem maksymalizacji prędkości usuwania materiału, co świadczy o skuteczności optymalizacji pod kątem tego kryterium.

- Zadanie optymalizacji nr 3:

Celem optymalizacji w zadaniu nr 3 było osiągnięcie chropowatości powierzchni (Ra) na poziomie 2.0 μm , przy jednoczesnym maksymalizowaniu prędkości usuwania materiału (MRR). Najwyższy priorytet przyznano MRR, natomiast Ra miało znacznie niższe znaczenie.

Model przewidywał Ra na poziomie 3.46 μm oraz MRR wynoszące 10.2 mm^3/min . Zmierzone wartości wyniosły odpowiednio 4.23 μm dla Ra i 11.29 mm^3/min dla MRR. W przypadku Ra odchylenie było istotne (MAPE 18.2%), natomiast dla MRR różnica była mniejsza (MAPE 9.65%). Zmierzona wartość Ra była wyższa zarówno od przewidywanej, jak i od założonej wartości optymalizacyjnej (2.0 μm). Tak znaczne odchylenie można tłumaczyć niższą wagą przypisaną Ra w tym zadaniu, co wpłynęło na mniejszą dokładność modelu w przewidywaniu tej zmiennej. Z kolei wyższa od przewidywanej wartość MRR, mimo odchylenia, pozostaje zgodna z celem maksymalizacji prędkości usuwania materiału i jest zgodna z założeniami optymalizacyjnymi.

10.7. Podsumowanie optymalizacji procesu WEDM

Zastosowanie algorytmu genetycznego (GA) okazało się najskuteczniejsze w wyznaczaniu parametrów technologicznych uznawanych za optymalne. W porównaniu do innych algorytmów, takich jak algorytm ewolucji różnicowej (DE) czy algorytmy rojowe (PSO, ACO), GA osiągnął najlepsze wyniki w minimalizacji funkcji celu, wykazując stabilność i porównywalny czas obliczeń.

Algorytm skutecznie wyznaczył parametry zapewniające osiągnięcie zadanej chropowatości powierzchni (Ra), maksymalizację prędkości usuwania materiału (MRR) oraz utrzymanie szerokości szczeliny erozyjnej (Kw) na ustalonym poziomie. Niemniej jednak, zauważono pewne odchylenia. Przewidywania dla Ra były mniej dokładne, szczególnie dla wyższych wartości (np. MAPE 18.2% w zadaniu nr 3). Wartości MRR były zaniżone w porównaniu do wyników rzeczywistych, co z jednej strony było korzystne, ponieważ rzeczywista prędkość usuwania materiału była wyższa. Z drugiej strony, stanowiło to błąd modelu, wskazujący na jego ograniczenia.

11. OPRACOWANIE I WDROŻENIE OPROGRAMOWANIA OPTYMALIZUJĄCEGO

W niniejszym rozdziale przedstawiono zagadnienia związane z opracowaniem i wdrożeniem oprogramowania służącego do optymalizacji procesu WEDM. Opracowanie to stanowi bezpośrednią kontynuację oraz praktyczne zastosowanie wyników uzyskanych w badaniach modelowych i optymalizacyjnych opisanych we wcześniejszych rozdziałach. Wyniki tych badań pozwoliły na zidentyfikowanie modeli przewidujących wartości zmiennych zależnych, które stanowiły podstawę do projektowania funkcji celu w algorytmie optymalizacji. Algorytm ten zostały następnie zaimplementowany w oprogramowaniu, umożliwiającym optymalizację parametrów procesu WEDM zgodnie z określonymi kryteriami technologicznymi.

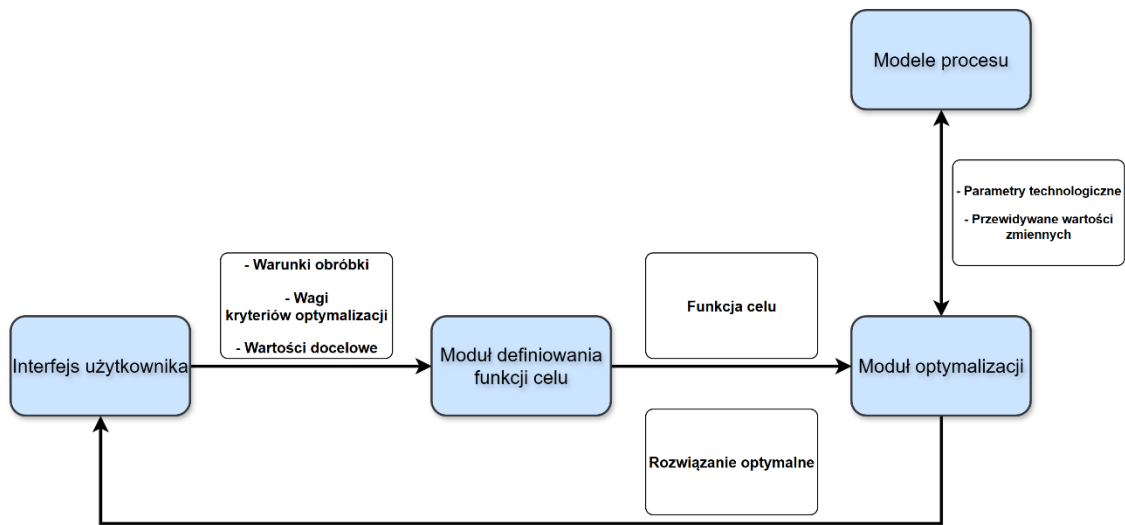
Rozdział podzielono na trzy główne sekcje, obejmujące architekturę oprogramowania (p.11.1), zastosowane technologie (p.11.2) oraz interfejs użytkownika (p.11.3).

11.1. Architektura oprogramowania

Architektura opracowanego oprogramowania optymalizacyjnego została zaprojektowana w sposób modułowy, co umożliwia jego elastyczność i skalowalność, a także łatwą integrację z istniejącymi systemami informatycznymi w przedsiębiorstwie. Jak przedstawiono na schemacie Rys. 11.1 oprogramowanie składa się z następujących modułów:

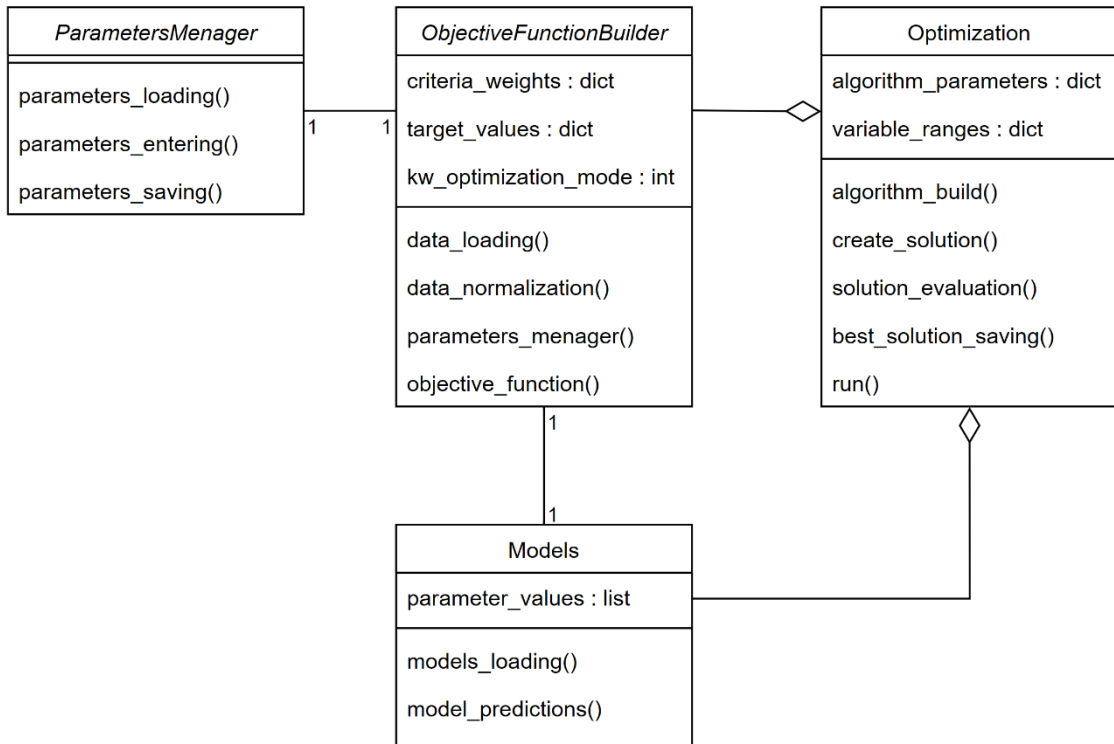
- **Interfejs użytkownika:** punkt wejścia do systemu, gdzie użytkownik może wprowadzać warunki obróbki, definiować wagi kryteriów optymalizacji, określać wartości docelowe dla zmiennych oraz monitorować wyniki optymalizacji.
- **Moduł definiowania funkcji celu:** odpowiedzialny za przetwarzanie danych wejściowych dostarczonych przez użytkownika. Moduł ten generuje funkcję celu na podstawie zdefiniowanych wag kryteriów, warunków obróbki oraz wartości docelowych. Funkcja celu jest następnie przekazywana do modułu optymalizacyjnego.
- **Moduł optymalizacji:** centralny element systemu, który odpowiada za realizację procesu optymalizacji. Wykorzystuje funkcję celu i algorytm optymalizacji, aby wygenerować optymalne rozwiązanie. Moduł ten współpracuje z modelami procesu, które dostarczają przewidywane wartości zmiennych zależnych na podstawie zadanych przez moduł optymalizacji parametrów technologicznych.
- **Modele procesu:** modele procesu przewidują wartości zmiennych zależnych na podstawie dostarczonych parametrów technologicznych. Przewidywane wartości

zmiennych są wykorzystywane przez moduł optymalizacji do oceny jakości rozwiązania i jego dalszego udoskonalania.

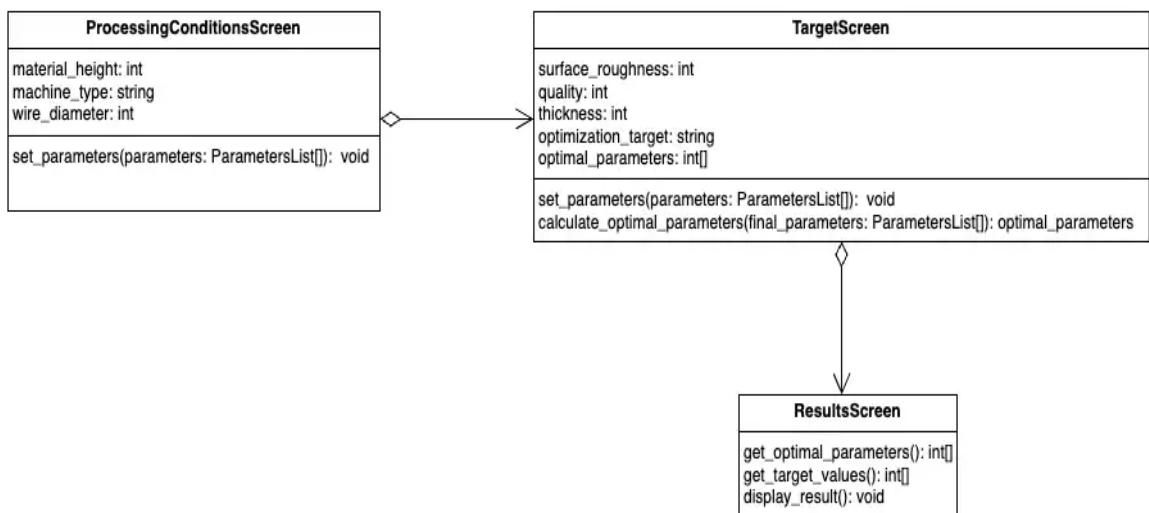


Rys. 11.1 Uproszczony schemat architektury oprogramowania optymalizującego

Na poziomie implementacyjnym architektura oprogramowania została zaprojektowana zgodnie z aktualnymi standardami w dziedzinie inżynierii oprogramowania [103], w wyniku czego wyodrębniono dwa odrębne serwisy: front-end (odpowiedzialny za graficzną reprezentację oprogramowania) w postaci aplikacji mobilnej oraz back-end (zajmujący się przetwarzaniem i logiką obliczeniową) zainstalowanym na fizycznym serwerze lokalnym. Oba serwisy zostały opracowane zgodnie z paradygmatem programowania obiektowego, co zostało przedstawione na diagramach klas serwisu back-endowego (Rys. 11.2) oraz front-endowego oprogramowania (Rys. 11.3).



Rys. 11.2 Diagram klas serwisu back-endowego oprogramowania



Rys. 11.3 Diagram klas serwisu front-endowego oprogramowania

11.2. Wykorzystana technologia

W celu opracowania oprogramowania wykorzystano szereg technologii. Ich wybór został podyktowany koniecznością zapewnienia wysokiej wydajności, elastyczności oraz bezpieczeństwa systemu. Poniżej przedstawiono kluczowe technologie, które zostały wykorzystane na różnych etapach tworzenia oprogramowania, zarówno w części mobilnej, jak i back-endowej:

1. Aplikacja mobilna:

- React Native: Aplikacja mobilna została opracowana przy użyciu frameworka React Native, który umożliwia tworzenie aplikacji natywnych na platformy iOS i Android z wykorzystaniem jednego wspólnego kodu źródłowego. To rozwiązanie pozwoliło na zredukowanie czasu związanego z tworzeniem aplikacji przy jednoczesnym zapewnieniu wysokiej wydajności i responsywności interfejsu użytkownika.

2. Backend serwerowy:

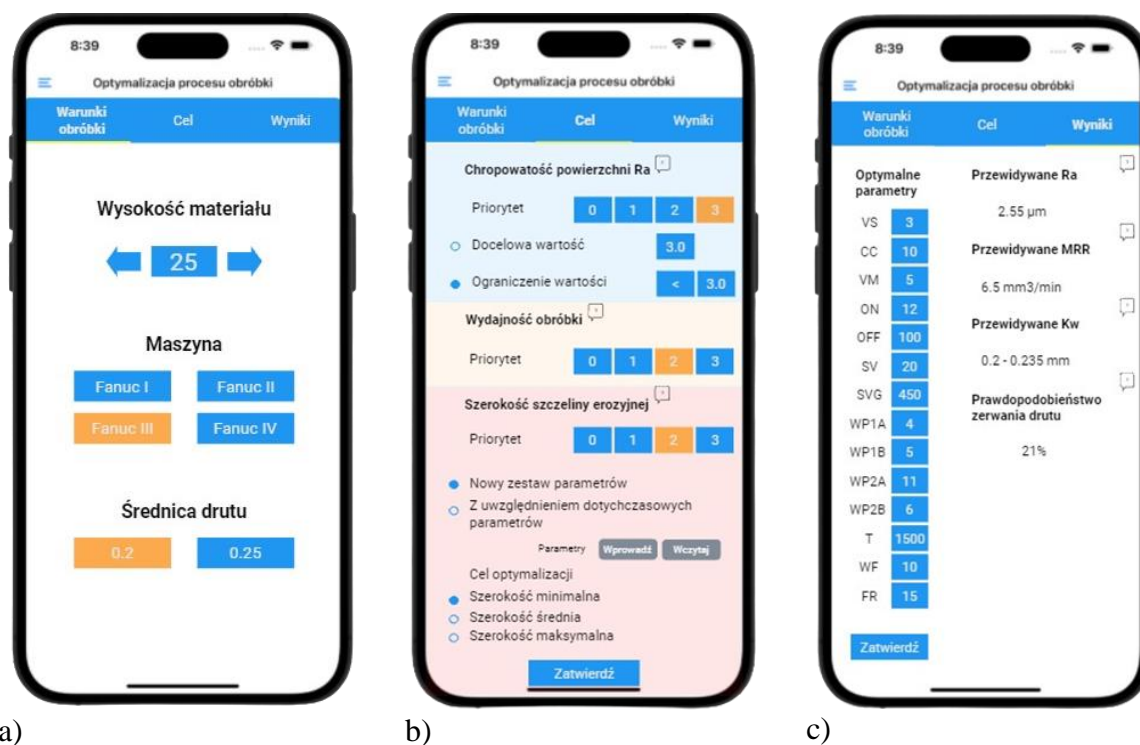
- Django: Back-end aplikacji został opracowany przy użyciu frameworka Django, który jest jednym z najbardziej zaawansowanych frameworków webowych opartych na języku Python. Django został wybrany ze względu na swoje zalety, takie jak szybkość tworzenia aplikacji, bezpieczeństwo, a także wbudowany ORM (Object-Relational Mapping), który ułatwia zarządzanie bazami danych.
- Django REST Framework: W celu umożliwienia komunikacji między aplikacją mobilną a serwerem lokalnym, wykorzystano Django REST Framework, który pozwala na łatwe budowanie architektury RESTful API ((Representational State Transfer Application Programming Interface). Architektura ta umożliwia efektywną wymianę danych między klientem (aplikacją mobilną) a serwerem, zachowując jednocześnie wysoką wydajność i skalowalność.

3. Baza danych:

- PostgreSQL: Do przechowywania i zarządzania danymi procesowymi wykorzystano relacyjną bazę danych PostgreSQL. Tego typu baza danych została wybrana ze względu na swoje zaawansowane możliwości obsługi skomplikowanych zapytań, integralność danych oraz skalowalność, co jest niezbędne w kontekście przechowywania dużych ilości danych produkcyjnych.

11.3. Interfejs użytkownika

Interfejs użytkownika został zaprojektowany zgodnie z zasadami projektowania zorientowanego na użytkownika [104], przy uwzględnieniu specyfiki procesu optymalizacji parametrów obróbki elektroerozyjnej. Główne założenia przyjęte przy projektowaniu interfejsu koncentrowały się na zapewnieniu maksymalnej funkcjonalności, ergonomii oraz intuicyjności obsługi.



Rys. 11.4 Widok interfejsu użytkownika, gdzie: a) ekran definiowania warunków obróbki, b) ekran określenia celu optymalizacji, c) ekran wyników

Interfejs został zorganizowany w trójdzielnej strukturze, odzwierciedlającej poszczególne etapy procesu optymalizacji:

1. Warunki obróbki (Rys. 11.4a):

W tej sekcji użytkownik wprowadza kluczowe parametry wejściowe, takie jak wysokość materiału, typ wykorzystywanej maszyny oraz średnicę drutu.

2. Cel (Rys. 11.4b):

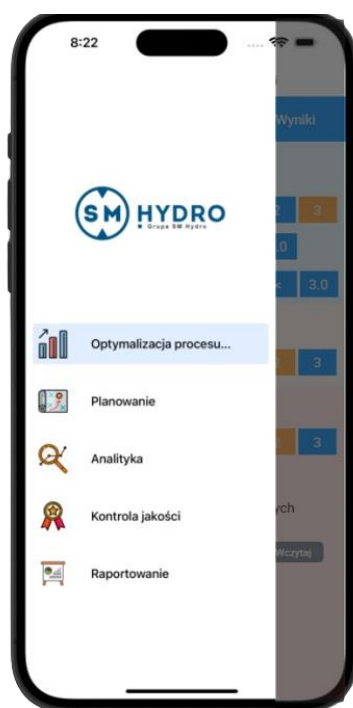
Sekcja ta umożliwia określenie parametrów docelowych, które służą do zdefiniowania kryteriów optymalizacji. Użytkownik może ustawić również priorytety (wagi) poszczególnych składników funkcji celu optymalizacji.

3. Wyniki (Rys. 11.4c):

Po zakończeniu etapu definiowania celu optymalizacji i zatwierdzeniu wprowadzonych danych, aplikacja generuje wartości parametrów technologicznych wyznaczonych jako optymalne, które są prezentowane na ekranie wyników. Na tym etapie użytkownik ma możliwość ręcznej edycji wygenerowanych wartości parametrów. Po zatwierdzeniu tych parametrów aplikacja oblicza i wyświetla przewidywane wartości chropowości powierzchni (Ra), szybkości usuwania materiału (MRR), oraz szerokości szczeliny erozyjnej. Taka funkcjonalność jest możliwa dzięki asynchronicznej realizacji procesów obliczeniowych, co umożliwia dynamiczną i iteracyjną optymalizację parametrów

technologicznych w czasie rzeczywistym. Dodatkowo prezentowane jest prawdopodobieństwo przerwania ciągłości obróbki.

Opracowano również ekran startowy aplikacji (Rys. 11.5), który umożliwia użytkownikowi dostęp do szeregu funkcjonalności, które są planowane do wdrożenia w przyszłości. Oprócz głównego modułu optymalizacji procesu, zaplanowano dodanie takich sekcji jak planowanie, analityka, kontrola jakości oraz raportowanie. Funkcjonalności te nie zostały jeszcze zaimplementowane i stanowią potencjalne obszary przyszłych badań i rozwoju. W miarę ich realizacji, będą mogły one zostać zintegrowane z systemem optymalizacji, co umożliwi zwiększenie stopnia automatyzacji oraz bardziej kompleksową analizę danych, co w efekcie może poprawić efektywność zarządzania procesami produkcyjnymi w przedsiębiorstwie.



Rys. 11.5 Ekran startowy aplikacji

11.4. Podsumowanie wdrożenia

Opracowane oprogramowanie zostało pomyślnie wdrożone w przedsiębiorstwie SMHydro. Implementacja została przeprowadzona poprzez podstawową integrację z istniejącą infrastrukturą informatyczną firmy.

Część serwerowa (backend serwerowy), odpowiedzialna za przetwarzanie danych, realizację algorytmów optymalizacyjnych oraz wykorzystanie modeli procesu, została zainstalowana na lokalnym serwerze przedsiębiorstwa. Wykorzystano bazę danych

PostgreSQL do przechowywania i zarządzania danymi procesowymi. Taka architektura umożliwiła centralizację obliczeń i danych.

Interfejs użytkownika, w postaci aplikacji mobilnej, został wdrożony na jednym urządzeniu mobilnym. Aplikacja ta komunikuje się z częścią serwerową i wyświetla wyniki optymalizacji oraz rekomendacje dotyczące parametrów technologicznych. Komunikacja pomiędzy serwerem a urządzeniem mobilnym odbywa się za pośrednictwem przewodu USB.

Obecna forma wdrożenia ma charakter testowy, co umożliwiło ocenę funkcjonalności oprogramowania w warunkach przemysłowych. Wdrożone oprogramowanie spełnia swoją rolę, pozwalając na optymalizację procesu. Jednak, aby mogło być szerzej stosowane w codziennym cyklu produkcyjnym, konieczne są dalsze usprawnienia, takie jak poprawa komunikacji oraz zwiększenie dostępności na wielu urządzeniach.

Plany na przyszłość związane z wdrożeniem obejmują:

- Rozszerzenie komunikacji na technologie bezprzewodowe (np. Wi-Fi, Bluetooth), co zwiększy mobilność i dostępność aplikacji mobilnej dla wielu użytkowników jednocześnie.
- Implementację dodatkowych modułów w aplikacji mobilnej, takich jak planowanie, analityka, kontrola jakości oraz raportowanie.
- Integrację z systemami sterowania obrabiarek CNC, umożliwiając automatyczne dostosowywanie parametrów technologicznych w czasie rzeczywistym i zwiększając stopień automatyzacji procesu produkcyjnego.
- Rozbudowę infrastruktury serwerowej, aby zapewnić większą skalowalność, wydajność i bezpieczeństwo systemu oraz umożliwić integrację z innymi systemami informatycznymi przedsiębiorstwa.

12. PODSUMOWANIE I WNIOSKI

W pracy przeprowadzono badania mające na celu opracowanie i wdrożenie oprogramowania do dynamicznej, wielokryterialnej optymalizacji procesu cięcia elektroerozyjnego, dostosowanego do wymagań przedsiębiorstwa SMHydro.

Podstawą realizacji tego zadania była identyfikacja modeli procesu WEDM, umożliwiających przewidywanie wartości wybranych zmiennych zależnych, tj. chropowatości powierzchni (Ra), prędkości usuwania materiału (MRR), szerokości szczeliny erozyjnej (Kw) oraz ciągłości procesu (C). W toku badań przeprowadzono analizę porównawczą wybranych metod modelowania pod kątem ich jakości, co pozwoliło wybrać najbardziej efektywne podejścia, w tym metody sztucznej inteligencji.

Opracowane modele umożliwiły przewidywanie wartości wymienionych wielkości na podstawie parametrów technologicznych, co z kolei pozwoliło na analizę zależności między tymi zmiennymi. Wpływ najistotniejszych parametrów zilustrowano za pomocą trójwymiarowych wykresów oraz przeprowadzono interpretację wyników.

Kolejnym etapem badań była optymalizacja procesu WEDM z wykorzystaniem opracowanych modeli oraz algorytmów optymalizacyjnych. Kryteria optymalizacji obejmowały analizowane zmienne zależne, przy czym cele polegały na osiągnięciu ustalonych wartości docelowych tych zmiennych, maksymalizacji MRR oraz zachowaniu ciągłości obróbki. Dodatkowo wprowadzono mechanizm pozwalający na regulację istotności poszczególnych kryteriów poprzez modyfikację przypisanych im wag. Dzięki takiemu podejściu możliwe było dynamiczne dostosowywanie funkcji celu do aktualnych priorytetów przedsiębiorstwa.

Na podstawie porównania wybranych algorytmów optymalizacyjnych inspirowanych biologicznie, wybrano ten, który okazał się najskuteczniejszy w poszukiwaniu rozwiązania uznanego za optymalne. Ostateczne wyniki badań potwierdziły, że zastosowanie wybranego algorytmu w połączeniu z opracowanymi modelami pozwala uzyskać pożądaną jakość obróbki oraz maksymalizację wydajności procesu cięcia elektroerozyjnego w warunkach przemysłowych. W efekcie opracowano oprogramowanie, które wdrożono w przedsiębiorstwie SMHydro, umożliwiające bieżące dostosowywanie parametrów technologicznych do zmieniających się wymagań.

12.1. Wnioski końcowe

Do najistotniejszych wniosków końcowych, które sformułowano na podstawie wyników przeprowadzonych badań, należą:

- Opracowanie modelu procesu WEDM uwzględniającego szeroki zakres zmiennych niezależnych, który jednocześnie charakteryzuje się wysoką

dokładnością przewidywania wartości zmiennej zależnej oraz zdolnością do generalizacji, wymaga zgromadzenia znacznej liczby danych treningowych.

- Zastosowanie metod sztucznej inteligencji w modelowaniu procesu WEDM zapewnia istotną przewagę nad podejściami statystycznymi. W szczególności zastosowanie modeli takich jak SVR (Support Vector Regression) i XGBoost umożliwiło osiągnięcie stosunkowo wysokiej dokładności w przewidywaniu wartości zmiennych zależnych.
- Zastosowanie algorytmu genetycznego (GA) w optymalizacji parametrów technologicznych okazało się najbardziej efektywne spośród testowanych algorytmów optymalizacyjnych.
- Wdrożone oprogramowanie do dynamicznej optymalizacji procesu WEDM stanowi skuteczne narzędzie wspomagające produkcję, pozwalając na elastyczne dostosowanie parametrów technologicznych do aktualnych potrzeb, co z kolei potencjalnie przyczynia się do zwiększenia wydajności i dokładności procesu.

12.2. Plan przyszłych badań

Na podstawie przeprowadzonych badań i uzyskanych wyników zidentyfikowano kilka obszarów, które mogą stanowić kierunki przyszłych badań:

- Przeprowadzenie badań z wykorzystaniem innych materiałów obrabianych oraz zastosowanie innych typów drutów w procesie WEDM. Umożliwi to uogólnienie opracowanych modeli i zwiększenie ich uniwersalności, co z kolei pozwoli na zastosowanie ich w szerszym spektrum aplikacji przemysłowych.
- Badanie możliwości zastosowania innych metod sztucznej inteligencji, takich jak głębokie sieci neuronowe, algorytmy uczenia ze wzmocnieniem czy modele hybrydowe. Celem jest zwiększenie dokładności przewidywań oraz efektywności optymalizacji procesu WEDM.
- Dostosowanie opracowanych metod i algorytmów optymalizacyjnych do innych procesów obróbki materiałów, takich jak frezowanie, toczenie czy szlifowanie.

- Rozszerzenie modeli o dodatkowe zmienne niezależne, takie jak warunki środowiskowe (np. temperatura) czy parametry związane z dielektrykiem. Może to przyczynić się do zwiększenia dokładności modeli i lepszego zrozumienia wpływu tych czynników na proces WEDM.
- Przeprowadzenie analizy ekonomicznej zastosowania opracowanych metod, w tym ocena kosztów i korzyści wynikających z optymalizacji procesu. Dodatkowo, badanie efektywności energetycznej procesu WEDM i możliwości jej poprawy poprzez optymalizację parametrów technologicznych.

BIBLIOGRAFIA

- [1] Pacana A., Czerwińska K.: Indicator analysis of the technological position of a manufacturing company. W: Production Engineering Archives. [online]. 2023, t. 29, nr. 2, s. 162–167.
- [2] Deisenroth M. P., Faisal A. A., Ong C. S.: Mathematics for machine learning. Cambridge University Press, Cambridge, United Kingdom, New York, NY 2020, ISBN 9781108679930.
- [3] Glodek Ł., Bysko S., Nocoń W.: Model Quality Assessment Method Based on Support Vector Machine. W: Pomiar Automatyka Robotyka. [online]. 2021, t. 25, nr. 1, s. 35–39.
- [4] Tadeusiewicz R., Szaleniec M.: Leksykon sieci neuronowych. Projekt Nauka, Wrocław 2015, ISBN 9788363270100.
- [5] Cichosz P.: Systemy uczące się. Wydawnictwa Naukowo-Techniczne, Warszawa 2007, ISBN 9788320433104.
- [6] Schumacher B. M., Krampitz R., Kruth J.-P.: Historical Phases of EDM Development Driven by the Dual Influence of “Market Pull” and “Science Push”. W: Procedia CIRP. [online]. 2013, t. 6, s. 5–12.
- [7] Lazarenko B.R: To Invert the effect of wear on electric power contacts. Związkowy Instytut Elektro-Techniczny, Moskwa 1943.
- [8] B. N. Solotych: Physikalische Grundlagen der Elektrofunkkenbearbeitung von Metallen. Verlag Technik, Berlin 1955.
- [9] Pramanik A., Basak A. K., Littlefair G., Debnath S., Prakash C., Singh M. A., Marla D., Singh R. K.: Methods and variables in Electrical discharge machining of titanium alloy - A review. W: Heliyon. [online]. 2020, t. 6, nr. 12.
- [10] Puri A., Bhattacharyya B.: An analysis and optimisation of the geometrical inaccuracy due to wire lag phenomenon in WEDM. W: International Journal of Machine Tools and Manufacture. [online]. 2003, t. 43, nr. 2, s. 151–159.
- [11] Tsai H., Yan B., Huang F.: EDM performance of Cr/Cu-based composite electrodes. W: International Journal of Machine Tools and Manufacture. [online]. 2003, t. 43, nr. 3, s. 245–252.
- [12] Sathiyaraj S., Venkatesan S., Ashokkumar S., senthilkumar A.: Wire electrical discharge machining (WEDM) analysis into MRR and SR on copper alloy. W: Materials Today: Proceedings. [online]. 2020, t. 33, s. 1079–1084.

- [13] Ho K., Newman S., Rahimifard S., Allen R.: State of the art in wire electrical discharge machining (WEDM). W: International Journal of Machine Tools and Manufacture. [online]. 2004, t. 44, 12-13, s. 1247–1259.
- [14] Oniszczyk-Świercz D., Świercz R.: Influence of processing parameters on the state of the surface layer after WEDM process. W: Mechanik. [online]. 2015, nr. 1, 54/14-54/17.
- [15] Almeida S., Mo J., Bil C., Ding S., Wang X.: Comprehensive servo control strategies for flexible and high-efficient wire electric discharge machining. A systematic review. W: Precision Engineering. [online]. 2021, t. 71, s. 7–28.
- [16] König W., Klocke F.: Fertigungsverfahren: Abtragen, Generieren Lasermaterialbearbeitung, SpringerLink Bücher. Springer, Berlin 1997-, ISBN 9783540489542.
- [17] FANUC Corporation: FANUC Robocut α -CiC series α -C400iC/C600iC: Maintenance Manual. FANUC Corporation 2021.
- [18] Slătineanu L., Dodun O., Coteață M., Nagiț G., Băncescu I. B., Hrițuc A.: Wire Electrical Discharge Machining—A Review. W: Machines. [online]. 2020, t. 8, nr. 4, s. 69.
- [19] Straka L., Piteľ J., Čorný I.: Influence of the main technological parameters and material properties of the workpiece on the geometrical accuracy of the machined surface at wedm. W: The International Journal of Advanced Manufacturing Technology. [online]. 2021, t. 115, 9-10, s. 3065–3087.
- [20] Burek J., Babiarczyk R., Buk J., Sułkiewicz P., Krupa K.: The Accuracy of Finishing WEDM of Inconel 718 Turbine Disc Fir Tree Slots. W: Materials. [online]. 2021, t. 14, nr. 3.
- [21] Wang C.-C., Chow H.-M., Yang L.-D., Lu C.-T.: Recast layer removal after electrical discharge machining via Taguchi analysis: A feasibility study. W: Journal of Materials Processing Technology. [online]. 2009, t. 209, nr. 8, s. 4134–4140.
- [22] Chalisgaonkar R., Kumar J., Pant P.: Prediction of machining characteristics of finish cut WEDM process for pure titanium using feed forward back propagation neural network. W: Materials Today: Proceedings. [online]. 2020, t. 25, s. 592–601.
- [23] Ragavendran U., Ghadai R. K., Bhoi A. K., Ramachandran M., Kalita K.: Sensitivity analysis and optimization of EDM process parameters. W: Transactions of the Canadian Society for Mechanical Engineering. [online]. 2019, t. 43, nr. 1, s. 13–25.
- [24] Devarasiddappa D., Chandrasekaran M., Arunachalam R.: Experimental investigation and parametric optimization for minimizing surface roughness during WEDM of

- Ti6Al4V alloy using modified TLBO algorithm. W: Journal of the Brazilian Society of Mechanical Sciences and Engineering. [online]. 2020, t. 42, nr. 3, s. 1–18.
- [25] Bose S., Nandi T.: A novel optimization algorithm on surface roughness of WEDM on titanium hybrid composite. W: Sādhanā. [online]. 2020, t. 45, nr. 1.
- [26] Nawaz Y., Maqsood S., Naeem K., Nawaz R., Omair M., Habib T.: Parametric optimization of material removal rate, surface roughness, and kerf width in high-speed wire electric discharge machining (HS-WEDM) of DC53 die steel. W: The International Journal of Advanced Manufacturing Technology. [online]. 2020, t. 107, 7-8, s. 3231–3245.
- [27] Gaitonde V. N., Manjaiah M., Maradi S., Karnik S. R., Petkar P. M., Paulo Davim J.: Multiresponse optimization in wire electric discharge machining (WEDM) of HCHCr steel by integrating response surface methodology (RSM) with differential evolution (DE). W: : Computational Methods and Production Engineering, Elsevier 2017, s. 199–221.
- [28] Doreswamy D., Bongale A. M., Piekarski M., Bongale A., Kumar S., Pimenov D. Y., Giasin K., Nadolny K.: Optimization and Modeling of Material Removal Rate in Wire-EDM of Silicon Particle Reinforced Al6061 Composite. W: Materials. [online]. 2021, t. 14, nr. 21, s. 6420.
- [29] Bergs T., Tombul U., Herrig T., Olivier M., Klink A., Klocke F.: Analysis of Characteristic Process Parameters to Identify Unstable Process Conditions during Wire EDM. W: Procedia Manufacturing. [online]. 2018, t. 18, s. 138–145.
- [30] Myers R. H., Montgomery D. C., Anderson-Cook C. M.: Response surface methodology: Process and product optimization using designed experiments, Wiley series in probability and statistics. Wiley, Hoboken 2016, ISBN 9781118916018.
- [31] Paturi U. M. R., Cheruku S., Pasunuri V. P. K., Salike S., Reddy N. S., Cheruku S.: Machine learning and statistical approach in modeling and optimization of surface roughness in wire electrical discharge machining. W: 2666-8270. [online]. 2021, t. 6, s. 100099.
- [32] Manoj I. V., Soni H., Narendranath S., Mashinini P. M., Kara F.: Examination of Machining Parameters and Prediction of Cutting Velocity and Surface Roughness Using RSM and ANN Using WEDM of Altemp HX. W: Advances in Materials Science and Engineering. [online]. 2022, t. 2022, s. 1–9.
- [33] Sagbas A., Gürtuna F., Polat U.: Comparison of ANN and RSM modeling approaches for WEDM process optimization. W: Materials Testing. [online]. 2021, t. 63, nr. 4, s. 386–392.

- [34] Lalwani V., Sharma P., Pruncu C. I., Unune D. R.: Response Surface Methodology and Artificial Neural Network-Based Models for Predicting Performance of Wire Electrical Discharge Machining of Inconel 718 Alloy. W: *Journal of Manufacturing and Materials Processing*. [online]. 2020, t. 4, nr. 2, s. 44.
- [35] Raj A., Misra J. P., Singh R. P., Singh G., Sharma S., Eldin S. M.: Performance analysis of WEDM during the machining of Inconel 690 miniature gear using RSM and ANN modeling approaches. W: *REVIEWS ON ADVANCED MATERIALS SCIENCE*. [online]. 2023, t. 62, nr. 1.
- [36] Oniszczyk-Świercz D., Świercz R., Michna Š.: Evaluation of Prediction Models of the Microwire EDM Process of Inconel 718 Using ANN and RSM Methods. W: *Materials*. [online]. 2022, t. 15, nr. 23.
- [37] Naresh C., Bose P. S. C., Rao C. S. P.: Artificial neural networks and adaptive neuro-fuzzy models for predicting WEDM machining responses of Nitinol alloy: comparative study. W: *SN Applied Sciences*. [online]. 2020, t. 2, nr. 2, s. 1–23.
- [38] Ghosh G., Mandal P., Mondal S. C.: Modeling and optimization of surface roughness in keyway milling using ANN, genetic algorithm, and particle swarm optimization. W: *The International Journal of Advanced Manufacturing Technology*. [online]. 2019, t. 100, 5-8, s. 1223–1242.
- [39] Hewidy M., Salem O.: Integrating experimental modeling techniques with the Pareto search algorithm for multiobjective optimization in the WEDM of Inconel 718. W: *The International Journal of Advanced Manufacturing Technology*. [online]. 2023, t. 129, 1-2, s. 299–319.
- [40] Patel J., Gianchandani R.: ANFIS control for robotic manipulators: Adaptive neuro fuzzy inference systems for intelligent control. Lambert Academic Publishing, Saarbrücken, Germany 2011, ISBN 9783846591710.
- [41] Voloşencu C.: Adaptive Neuro-Fuzzy Inference System as a Universal Estimator. IntechOpen 2024, ISBN 978-0-85466-152-7.
- [42] Natarajan M., Pasupuleti T., Abdullah M. M. S., Mohammad F., Giri J., Chadge R., Sunheriya N., Mahatme C., Giri P., Soleiman A. A.: Assessment of Machining of Hastelloy Using WEDM by a Multi-Objective Approach. W: *Sustainability*. [online]. 2023, t. 15, nr. 13, s. 10105.
- [43] Singh Nain S., Sai R., Sihag P., Vambol S., Vambol V.: Use of machine learning algorithm for the better prediction of SR peculiarities of WEDM of Nimonic-90 superalloy. W: *Archives of Materials Science and Engineering*. [online]. 2019, t. 1, nr. 95, s. 12–19.

- [44] Upadhyay V., Misra J. P., Singh B.: Wire-breakage prediction during WEDM of Ni-based superalloy using machine learning-based classifier approaches. W: International Journal on Interactive Design and Manufacturing (IJIDeM). [online]. 2023.
- [45] Nain S. S., Garg D., Kumar S.: Performance evaluation of the WEDM process of aeronautics super alloy. W: Materials and Manufacturing Processes. [online]. 2018, t. 33, nr. 16, s. 1793–1808.
- [46] Ukey K., Rameshchandra Sahu A., Sheshrao Gajghate S., Kumar Behera A., Limbadri C., Majumder H.: Wire electrical discharge machining (WEDM) review on current optimization research trends. W: Materials Today: Proceedings. [online]. 2023.
- [47] Kumar N., Kumari S., Abhishek K., Nandi G., Ghosh N.: Study on various parameters of WEDM using different optimization techniques: A review. W: Materials Today: Proceedings. [online]. 2022, t. 62, s. 4018–4024.
- [48] Ullah S., Niazi H. U. K., Tanveer A.: Techniques Used for Process Optimization of Wire Electrical Discharge Machining: A Review. W: Engineering Proceedings. [online]. 2023, t. 56, nr. 1, s. 118.
- [49] Majumder H., Maity K.: Multi-Response Optimization of WEDM Process Parameters Using Taguchi Based Desirability Function Analysis. W: IOP Conference Series: Materials Science and Engineering. [online]. 2018, t. 338, s. 12004.
- [50] Reddy V. C., Deepthi N., Jayakrishna N.: Multiple Response Optimization of Wire EDM on Aluminium HE30 by using Grey Relational Analysis. W: Materials Today: Proceedings. [online]. 2015, t. 2, 4-5, s. 2548–2554.
- [51] Tzeng Y., Chen F.: Multi-objective optimisation of high-speed electrical discharge machining process using a Taguchi fuzzy-based approach. W: Materials & Design. [online]. 2007, t. 28, nr. 4, s. 1159–1168.
- [52] Kumar A., Grover N., Manna A., Kumar R., Chohan J. S., Singh S., Singh S., Pruncu C. I.: Multi-Objective Optimization of WEDM of Aluminum Hybrid Composites Using AHP and Genetic Algorithm. W: Arabian Journal for Science and Engineering. [online]. 2022, t. 47, nr. 7, s. 8031–8043.
- [53] Kalita K., Ghadai R. K., Chakraborty S.: A comparative study on multi-objective pareto optimization of WEDM process using nature-inspired metaheuristic algorithms. W: International Journal on Interactive Design and Manufacturing (IJIDeM). [online]. 2023, t. 17, nr. 2, s. 499–516.
- [54] Zolpakar N. A., Yasak M. F., Pathak S.: A review: use of evolutionary algorithm for optimisation of machining parameters. W: The International Journal of Advanced Manufacturing Technology. [online]. 2021, t. 115, 1-2, s. 31–47.

- [55] Lenin N., Sivakumar M., Selvakumar G., Rajamani D., Sivalingam V., Gupta M. K., Mikolajczyk T., Pimenov D. Y.: Optimization of Process Control Parameters for WEDM of Al-LM25/Fly Ash/B4C Hybrid Composites Using Evolutionary Algorithms: A Comparative Study. W: Metals. [online]. 2021, t. 11, nr. 7, s. 1105.
- [56] Chen Y., Hu S., Li A., Cao Y., Zhao Y., Ming W.: Parameters Optimization of Electrical Discharge Machining Process Using Swarm Intelligence: A Review. W: Metals. [online]. 2023, t. 13, nr. 5, s. 839.
- [57] Padhi P. C., Mahapatra S. S., Yadav S. N., Tripathy D. K.: Multi-Objective Optimization of Wire Electrical Discharge Machining (WEDM) Process Parameters Using Weighted Sum Genetic Algorithm Approach. W: Journal of Advanced Manufacturing Systems. [online]. 2016, t. 15, nr. 02, s. 85–100.
- [58] Hardik A. Patel, Dr. Md. Najibullah, Prof. B. C. Khatri, S. P. Patel: A review of parametric optimization of wire electric discharge machining, nr. 3. International Journal for Scientific Research & Development, Ahmedabad, Indie 2015.
- [59] Król-Nowak A., Kotarba K.: Podstawy uczenia maszynowego. Wydawnictwa AGH, Kraków 2022, ISBN 978-83-67427-05-0.
- [60] Foster J. J., Barkus E., Yavorsky C.: Understanding and using advanced statistics. SAGE Publications, London, Thousand Oaks, Calif. 2006, ISBN 1412900131.
- [61] Boslaugh S.: Statistics in a nutshell: A desktop quick reference, In a nutshell. O'Reilly Media, Sebastopol, CA 2013, ISBN 9781449361143.
- [62] Weisberg S.: Applied linear regression, Wiley Series in Probability and Statistics Ser. Wiley, Hoboken, New Jersey 2014, ISBN 9781118789551.
- [63] Harrell F. E.: Regression modeling strategies: With applications to linear models, logistic and ordinal regression, and survival analysis, Springer Series in Statistics. Springer, Cham 2015, ISBN 978-3-319-19424-0.
- [64] Friedman J.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer Series in Statistics. Springer-Verlag New York, New York, NY 2009, ISBN 9780387848587.
- [65] Greenwell B. M.: Tree-based methods for statistical learning in R, Chapman & Hall/CRC data science series. CRC PRESS, Boca Raton, FL 2022, ISBN 9780367532468.
- [66] Müller A. C., Guido S.: Introduction to machine learning with Python: A guide for data scientists. O'Reilly Media, Sebastopol, CA 2017.
- [67] Genuer R., Poggi J.-M.: Random Forests with R, Use R! Springer International Publishing; Springer, Cham 2020, ISBN 9783030564858.

- [68] Shalev-Shwartz S., Ben-David S.: Understanding machine learning: From theory to algorithms. Cambridge University Press, Cambridge 2022, ISBN 9781107057135.
- [69] Chen T., Guestrin C.: XGBoost. W: B. Krishnapuram, M. Shah, A. Smola, C. Aggarwal, D. Shen, R. Rastogi (eds.): KDD '16, Association for Computing Machinery, New York 2016, s. 785–794.
- [70] Rad J. A., Parand K., Chakraverty S.: Learning with Fractional Orthogonal Kernel Classifiers in Support Vector Machines: Theory, Algorithms and Applications. Springer, Singapore 2023, ISBN 978-981-19-6552-4.
- [71] Hagan M. T., Demuth H. B., Beale M. H., Jesús O. de: Neural network design, Stillwater 2016, ISBN 9780971732117.
- [72] Badar A. Q. H.: Evolutionary optimization algorithms. CRC PRESS, Boca Raton, FL 2022, ISBN 9781000462142.
- [73] Kramer O.: Genetic Algorithm Essentials, Studies in computational intelligence nr. 679. Springer International Publishing, Cham, Switzerland 2017, ISBN 978-3-319-52155-8.
- [74] Cortez P.: Modern optimization with R, Use R! Springer, Cham 2014, ISBN 9783319082639.
- [75] Kumar B. V.: Differential evolution: From theory to practice, Studies in computational intelligence volume 1009. Springer, Singapore 2022, ISBN 9789811680823.
- [76] Dorigo M.: Ant colony optimization. MIT Press, Cambridge, Massachusetts 2004, ISBN 9780262256032.
- [77] Solnon C.: Ant Colony Optimization and Constraint Programming, ISTE. Wiley, London 2013, ISBN 1118557565.
- [78] Troyanskaya O., Cantor M., Sherlock G., Brown P., Hastie T., Tibshirani R., Botstein D., Altman R. B.: Missing value estimation methods for DNA microarrays. W: Bioinformatics (Oxford, England). [online]. 2001, t. 17, nr. 6, s. 520–525.
- [79] van Buuren S., Groothuis-Oudshoorn K.: mice : Multivariate Imputation by Chained Equations in R. W: Journal of Statistical Software. [online]. 2011, t. 45, nr. 3.
- [80] Wakefield J.: Bayesian and frequentist regression methods, Springer Series in Statistics. Springer, New York, NY 2013, ISBN 978-1-4419-0924-4.
- [81] Zhou X., Reiter J. P.: A Note on Bayesian Inference After Multiple Imputation. W: The American Statistician. [online]. 2010, t. 64, nr. 2, s. 159–163.

-
- [82] Kuhn M., Johnson K.: Feature Engineering and Selection: A Practical Approach for Predictive Models: A practical approach for predictive models, Chapman & Hall/CRC data science series. Chapman and Hall, Boca Raton, London, New York 2019, ISBN 9781138079229.
- [83] Mukhametzyanov I. Z.: Normalization of multidimensional data for multi-criteria decision making problems: Inversion, displacement, asymmetry, International series in operations research and management science series volume 348. Springer, Cham, Switzerland 2023, ISBN 9783031338373.
- [84] Duboue P.: The art of feature engineering: Essentials for machine learning. Cambridge University Press, Cambridge 2020, ISBN 9781108671682.
- [85] He H.: Imbalanced learning: Foundations, algorithms, and applications. John Wiley & Sons, Inc, Hoboken, New Jersey 2013, ISBN 9781118646205.
- [86] Bartz E.: Hyperparameter Tuning for Machine and Deep Learning with R: A Practical Guide. TH Koln, Institute for Data Science, Engineering, and Analytics, Singapore 2023, ISBN 9789811951701.
- [87] Durbin J., Watson G. S.: Testing for Serial Correlation in Least Squares Regression: I. W: Biometrika. [online]. 1950, t. 37, 3/4, s. 409.
- [88] Jarque C. M.: Jarque-Bera Test. W: M. Lovric (ed.): International encyclopedia of statistical science, Springer reference, Springer, Berlin, Heidelberg 2011, s. 701–702.
- [89] Shah M.: Evaluating learning algorithms - a classification perspective: A classification perspective. Cambridge University Press, New York 2013, ISBN 9781107653115.
- [90] Chicco D., Tötsch N., Jurman G.: The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation. W: BioData mining. [online]. 2021, t. 14, nr. 1, s. 13.
- [91] Suhaimi N. S., Othman Z., Yaakub M. R.: Comparative Analysis Between Macro and Micro-Accuracy in Imbalance Dataset for Movie Review Classification. W: : Yang Red. Proceedings of Seventh, 2022, s. 83–93.
- [92] Bhattacharya A.: Applied Machine Learning Explainability Techniques: Make ML Models Explainable and Trustworthy for Practical Applications Using LIME, SHAP, and More. Packt Publishing, Limited, Birmingham 2022, ISBN 9781803234168.
- [93] Nandi A., Pal A. K.: Interpreting Machine Learning Models: Learn model interpretability and explainability methods. Apress, New York 2022, ISBN 9781484278024.

- [94] Viering T., Loog M.: The Shape of Learning Curves: A Review. W: IEEE transactions on pattern analysis and machine intelligence. [online]. 2023, t. 45, nr. 6, s. 7799–7819.
- [95] Świercz R.: Modelowanie i optymalizacja obróbki elektroerozyjnej materiałów trudnoobrabialnych. Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa 2019, ISBN 9788378149217.
- [96] Comprehensive Guide to Factorial Two-Level Experimentation, SpringerLink Bücher. Springer-Verlag New York, New York, NY 2009, ISBN 9780387891033.
- [97] Korzyński M.: Metodyka eksperymentu: Planowanie, realizacja i statystyczne opracowanie wyników eksperymentów technologicznych. Wydawnictwo Naukowe PWN, Warszawa 2017, ISBN 9788301193188.
- [98] Ledolter J., Swersey A. J.: Testing 1 - 2 - 3: Experimental Design with Applications in Marketing and Service Operations. Stanford University Press, Stanford, CA 2007, ISBN 9781503626515.
- [99] Ferreira S. L. C., Bruns R. E., Ferreira H. S., Matos G. D., David J. M., Brandão G. C., Da Silva E. G. P., Portugal L. A., dos Reis P. S., Souza A. S., dos Santos W. N. L.: Box-Behnken design: an alternative for the optimization of analytical methods. W: Analytica chimica acta. [online]. 2007, t. 597, nr. 2, s. 179–186.
- [100] Hair J. F.: A primer on partial least squares structural equations modeling (PLS-SEM). SAGE Publications, Thousand Oaks 2014, ISBN 9781452217444.
- [101] Cabello-Solorzano K., Ortigosa de Araujo I., Peña M., Correia L., J. Tallón-Ballesteros A.: The Impact of Data Normalization on the Accuracy of Machine Learning Algorithms: A Comparative Analysis. W: : García Bringas (Red.) 18th International Conference on Soft, s. 344–353.
- [102] Freeman E. A., Moisen G. G.: A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. W: Ecological Modelling. [online]. 2008, t. 217, 1-2, s. 48–58.
- [103] Richards M., Ford N.: Fundamentals of software architecture: An engineering approach. O'Reilly, Beijing, Boston, Farnham, Sebastopol, Tokyo 2020, ISBN 9781492043409.
- [104] Dix A.: Human-computer interaction. Prentice Hall, Harlow 2003, ISBN 9780130461094.