



AGH UNIVERSITY OF SCIENCE AND TECHNOLOGY  
FACULTY OF ELECTRICAL ENGINEERING, AUTOMATICS, COMPUTER SCIENCE AND BIOMEDICAL ENGINEERING  
INSTITUTE OF APPLIED COMPUTER SCIENCE

PDITT-npi. 24.07.2023

M. Skow

Kraków 14 lipca 2023

Dr. hab. inż. Jarosław Wąs, prof. AGH  
Wydział Elektrotechniki, Automatyki, Informatyki i Inżynierii Biomedycznej  
Akademia Górniczo Hutnicza im. Stanisław Staszica w Krakowie

### Recenzja rozprawy doktorskiej

Recenzję pracy doktorskiej mgr inż. Joanny Badury zatytułowanej „*Solutions for selected problems of demand forecasting based on machine learning methods and domain knowledge*”, opracowano na prośbę prof. Andrzeja Polańskiego Przewodniczącego Rady Dyscypliny Informatyka Techniczna i Telekomunikacja (na podstawie umowy UMC/1948/2023 z dnia 27 czerwca 2023).

#### 1a) Tematyka rozprawy i tytuł pracy

Tematyka rozprawy doktorskiej mgr inż. Joanny Badury pt. „*Solutions for selected problems of demand forecasting based on machine learning methods and domain knowledge*”, dotyczy problematyki prognozowania zapotrzebowania z wykorzystaniem algorytmów uczenia maszynowego.

Praca obejmuje takie zagadnienia jak: propozycję klasyfikacji prognozowania zapotrzebowania, konfigurację szeregu algorytmów uczenia maszynowego do zagadnień związanych z przewidywaniami zapotrzebowania i określenia parametrów wpływu promocji na sprzedaż produktów, wpływ konfiguracji szeregów czasowych na rozwiązanie postawionych problemów itd.

W mojej ocenie temat pracy poprawnie koresponduje z zawartymi treściami, gdyż praca poświęcona jest metodom sztucznej inteligencji w prognozowaniu zapotrzebowania.

## **1b) Charakterystyka rozprawy i ocena układu rozprawy**

Praca doktorska mgr inż. Joanny Badury składa się łącznie z siedmiu rozdziałów, obejmujących: wstęp, zasadniczą część pracy, podsumowanie oraz bibliografię. Całość pracy zajmuje 206 stron wraz z bibliografią.

Pierwszy rozdział pracy zatytułowany „Introduction” poświęcony jest określeniu celów ogólnych i szczegółowych pracy, wskazaniu problemu naukowego oraz określeniu metodyki pracy z uwzględnieniem: badań literaturowych, wypracowaniu generycznego podejścia i szeregu szczegółowych rozwiązań w modelowaniu predykcji.

Rozdział 2 zatytułowany „Demand forecasting – approaches and challenges” przedstawia wprowadzenie do tematu prognozowania zapotrzebowania z wykorzystaniem uczenia maszynowego i wiedzy dziedzinowej.

Z kolei w rozdziale 3 *Promotions* przedstawiono temat prognozowania efektywności promocji.

Rozdział 4 *Top-down forecasting* dotyczy problemu prognozowania w schemacie od ogółu do szczegółu (top-down forecasting).

Rozdział 5 *Incorporating additional time series related to demand forecasting* przedstawia problem użycia dodatkowych szeregów czasowych w celu poprawy prognozowania badanego zjawiska.

Rozdział 6 *Returns* koncentruje się na prognozowaniu prawdopodobieństwa zwrotu produktu.

Ostatni z rozdziałów 7 *Summary* stanowi podsumowanie odnoszące się do uzyskanych wyników. Odniesiono się do realizacji celów ogólnych i szczegółowych pracy, które zostały osiągnięte.

W mojej ocenie układ redakcyjny pracy jest poprawny. Autorka posługuje się poprawną terminologią w zakresie zagadnienia prognozowania i, patrząc szerzej, stosowanych metod sztucznej inteligencji.

Należy stwierdzić, że oceniana praca jest napisana poprawnym językiem. Bibliografia jest generalnie aktualna. Odwołania do pozycji literaturowych są prawidłowe.

Tytuły i kolejność rozdziałów jest również prawidłowa, treść rozdziałów odpowiada, w mojej ocenie, wymaganiom stawianym dla prac doktorskich.

## **1c) Ocena zastosowanego piśmiennictwa**

Bibliografia składa się ze 190 pozycji związanych z tematyką pracy. Pozycje literaturowe zostały dobrane prawidłowo biorąc pod uwagę istotność i

aktualność poszczególnych pozycji bibliograficznych. Zostały one poprawnie przeanalizowane i wkomponowane w treść rozprawy.

#### **1d) wskazanie oraz ocena celu pracy**

Można wskazać dwa główne cele pracy: pierwszym celem było zaproponowanie taksonomii problemów badawczych związanych z prognozowaniem popytu, które mogą być rozwiązane metodami uczenia maszynowego.

Drugim zasadniczym celem było zaproponowanie wykorzystania i parametryzacji metod uczenia maszynowego (a także eksperckiej wiedzy) do rozwiązywania konkretnych zagadnień.

W mojej ocenie obydwie te cele zostały zrealizowane w prawidłowy sposób.

#### **1e) wskazanie oraz ocena zastosowanych metod badawczych**

W prognozowaniu zapotrzebowania Autorka zastosowała metodę analizy i na tej podstawie zaproponowała taksonomię zagadnień związanych z tworzeniem prognoz. Następnie zaproponowano parametryzację i poprawę szeregu metod uczenia maszynowego do wpływu skuteczności promocji i wpływu zwrotów. Uważam, że metody badawcze zostały zastosowane w prawidłowy sposób.

#### **1f) ocena części rozprawy doktorskiej dotyczącej omówienia wyników badań**

Wyniki badań zostały opisane w praktycznej części pracy z uwzględnieniem zaproponowanej przez Autorkę taksonomii, wpływu promocji na sprzedaż z uwzględnieniem metod uczenia maszynowego i oceną jakości rozwiązań, prognozowania od ogółu do szczegółu (począwszy od metody naiwnej, poprzez parametryczny model mieszany, skończywszy na metodzie XGBoost), włączenie dodatkowych szeregów czasowych związanych z prognozowaniem zapotrzebowania (włączanie danych z podobnych lokalizacji, użycie danych z tej samej lokalizacji), czy wreszcie algorytmika związana z prawdopodobieństwem zwrotów poszczególnych artykułów.

W mojej ocenie prace merytoryczne zostały przeprowadzone i opisane prawidłowo. Autorka szczegółowo opisała metody, w wielu przypadkach zostały one zilustrowane za pomocą obrazów i tabel. Podsumowując ten punkt oceny uważam, że wyniki badań zostały opisane w prawidłowy sposób.

#### **1g) informacje dotyczące praktycznego zastosowania uzyskanych wyników badań**

Badania zawarte w rozprawie mają wymiar praktyczny –zaś praca jest związana z następującymi projektami:

- “System wspomaganie decyzji i zarządzania wiedzą dla branży handlu detalicznego (DSS-RETAIL)” (POIR.01.01.01-00-0871/17-00)
- “System SupplAI automatyzujący i optymalizujący planowanie łańcucha dostaw, wykorzystujący metody inteligencji obliczeniowej – w tym prognozowania sprzedaży – do analizy online bardzo dużych zbiorów danych pozwalający na ograniczenie strat i zwiększenie efektywności działania firm branży handlu detalicznego” (POIR.01.01.01-00-0500/21) – realizacja z firmą 3Soft S.A.
- “System wspomaganie decyzji i zarządzania wiedzą operacyjną i procesową dla rynku dystrybutorów gazu LPG” współfinansowany przez Narodowe Centrum Badań i Rozwoju w ramach grantu 17 POIR.01.01.01-00-0104/17. Badania związane z prognozowaniem sprzedaży paliw na stacjach benzynowych zostały zrealizowane we współpracy AIUT Sp. z o.o.

**1h) informacje o ewentualnych nieprawidłowościach, które pojawiły się w ocenianej rozprawie doktorskiej**

Nie zidentyfikowano poważniejszych nieprawidłowości w rozprawie.

**1i) ocena czy rozprawa doktorska stanowi oryginalne rozwiązanie problemu naukowego**

W pracy rozwiązano szereg zagadnień związanych z prognozowaniem zapotrzebowania, które świadczą o oryginalności rozwiązania. Po pierwsze usystematyzowano podejście do uczenia poprzez zaproponowanie taksonomii, po drugie skonstruowano metody przewidywania efektu promocji za pomocą dedykowanych wskaźników i metod uczenia maszynowego, po trzecie zaproponowano metodę przygotowania cech przy użyciu tabeli niepodobieństwa modeli produktów i konwersji różnych rozmiarów produktów na wartość liczbową. Po czwarte wykazano, że wprowadzenie dodatkowych szeregów czasowych do prognozowania zapotrzebowania dla rozważanych przypadków/produktów, biorąc pod uwagę szeregi czasowe z podobnych lokalizacji i tej samej lokalizacji. Po piąte zaproponowano algorytm łączenia transakcji zwrotu z transakcją zakupu i analizę prognozowania zwrotu określonego produktu z określonej transakcji zakupu.

Biorąc pod uwagę wszystkie składowe uważam, że oceniana rozprawa zawiera oryginalne rozwiązanie problemu naukowego.

**1j) ocena czy rozprawa doktorska prezentuje ogólną wiedzę teoretyczną kandydata w dyscyplinie albo dyscyplinach oraz umiejętność samodzielnego prowadzenia pracy naukowej lub artystycznej**

W mojej ocenie praca bardzo dobrze wpisuje się w dyscyplinę informatyka techniczna i telekomunikacja i napisana jest poprawnie z punktu widzenia metodyki. Autorka wykazała się wiedzą teoretyczną i praktycznymi umiejętnościami w zakresie analizy danych i tworzenia algorytmiki.

**Uwagi:**

Jak to zostało napisane wcześniej, praca została napisana poprawnym językiem i złożona starannie ze strony edytorskiej.

Poniżej przedstawiono uwagi polemiczne dotyczące strony merytorycznej pracy:

- W sekcji 1.1. Autorka proponuje tezę pracy i suplement w której opisuje rolę nowych atrybutów i rekordów danych. W mojej opinii brakuje w tym miejscu dokładniejszego opisu: jak rozumiana jest w niniejszej pracy wiedza domenowa i precyzyjnego opisu jej reprezentacji i zastosowania. Owszem pojęcie to jest przywoływane zarówno w tytule i w wielu miejscach pracy (w sekcji końcowej pracy Autorka odwołuje się do niego - bodajże dziesięciokrotnie), ale brakuje mi jasnej definicji na początku pracy.
- W niektórych fragmentach rozprawa doktorska przypomina bardziej dokumentację techniczną niż rozprawę doktorską. Przykładowo tytuły rozdziałów 3 – *Promotions* czy 6 - *Returns* powinny być, w mojej opinii, sformułowane w bardziej precyzyjny sposób, dotyczą bowiem analizy danego scenariusza i stworzenia efektywnej algorytmiki predykcyjnej dla tych przypadków. Aktualne brzmienie tych tytułów sekcji ma mnemotechniczny charakter.
- Istotnym problemem w praktycznych zastosowaniach metod ML są problemy niekonsystencji danych. Wzmiankę o niekonsystencji znalazłem jedynie w rozdziale 6 dotyczącym analizy zwrotów. Przydałoby się opisać to zagadnienie, gdyż w praktyce bardzo często ono rzutuje na poszczególne etapy rozwiązywania zadań z wykorzystaniem metod ML.
- Zaproponowana taksonomia przedstawiona w formie rysunku na stronie 12 pracy oraz skomentowana w dalszych sekcjach rozdziału 2, jest niewątpliwie ciekawym podejściem. Przydałoby się jednak dokładniejsza wykładnia dotycząca parametryzacji ze względu na czas (który jest tu

pojęciem względnym ze względu na różnorodność procesów) i pozostałe czynniki

Opisane powyżej uwagi nie wpływają jednak negatywnie na mój bardzo pozytywny odbiór ocenianej pracy doktorskiej.

### **Wniosek końcowy**

Rozprawa stanowi oryginalne rozwiązanie problemu naukowego i wskazuje na wysoki poziom wiedzy teoretycznej i praktycznej Kandydatki w dyscyplinie naukowej Informatyka Techniczna i Telekomunikacja. Przedstawiona praca doktorska spełnia w pełni warunki określone w art. 187 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (j.t. Dz.U. z 2020 r. poz. 85 z późn. zm.). W szczególności zgodnie z wymaganiami ww. ustawy oceniana rozprawa pozytywnie prezentuje wiedzę teoretyczną Kandydatki oraz umiejętność samodzielnego prowadzenia pracy naukowej, przedmiotem rozprawy jest oryginalne rozwiązanie problemu naukowego. W opinii recenzenta praca może więc być dopuszczona do publicznej obrony.

Dr hab. inż. Jarosław Wąs Prof. AGH  
jaroslaw.was@agh.edu.pl

