

SILESIAAN UNIVERSITY OF TECHNOLOGY
FACULTY OF AUTOMATIC CONTROL,
ELECTRONICS AND COMPUTER SCIENCE

DOCTORAL DISSERTATION

**Solutions for selected problems of
demand forecasting based on
machine learning methods and
domain knowledge**

Joanna Badura

Supervisor : dr hab. Marek Sikora, Associate Professor

Gliwice 2023

Rozwiązania wybranych problemów prognozowania zapotrzebowania oparte na metodach uczenia maszynowego i wiedzy dziedzinowej (Solutions for selected problems of demand forecasting based on machine learning methods and domain knowledge)

Abstrakt

Joanna Badura
Politechnika Śląska
Wydział Automatyki, Elektroniki i Informatyki

Promotor: dr hab. Marek Sikora, prof. Politechniki Śląskiej

Rozprawa doktorska bada wykorzystanie metod uczenia maszynowego i wiedzy dziedzinowej do rozwiązywania wybranych zagadnień związanych z prognozowaniem zapotrzebowania. Zrozumienie i dokładne prognozowanie zapotrzebowania ma kluczowe znaczenie dla firm, wpływając na zarządzanie zasobami, wydajność łańcucha dostaw i zadowolenie klientów.

Kwestie prognozowania zapotrzebowania nie ograniczają się jedynie do prognozowania sprzedaży na produkty w oparciu o istniejące historyczne dane dotyczące sprzedaży konkretnego produktu. Temat prognozowania zapotrzebowania jest znacznie szerszy i dana rozprawa doktorska jest poświęcona różnym aspektom związanym z tym tematem.

W pracy przedstawiono oryginalną taksonomię zagadnień dotyczących prognozowania zapotrzebowania. Każde zagadnienie zostało opisane oraz wskazano powiązane z nimi wyzwania. Następnie skupiono się na wybranych zagadnieniach i rozważono, w jaki sposób można zastosować metody uczenia maszynowego do rozwiązania danych problemów.

W pierwszej kolejności podjęto temat analizy i prognozowania efektywności promocji. Zaproponowano zestaw wskaźników, które można wykorzystać do oceny promocji. Pokazano proces tworzenia atrybutów warunkowych i zbioru treningowego w celu stworzenia modeli predykcyjnych dla każdego wskaźnika. Następnie pokazano, w jaki sposób proponowane modele predykcyjne mogą być wykorzystywane w systemie rekomendacji promocji i do odkrywania wiedzy na temat wrażliwości cenowej. Na koniec rozważono wykorzystanie oryginalnej metody indukcji przeżyciowych reguł akcji do analizy zmian cech promocji.

Kolejnym poruszonym tematem była kwestia podziału prognoz wyższego rzędu na prognozy niższego rzędu. Jest to prognozowanie od ogółu do szczegółu (*top-down forecasting*). Przeprowadzono porównanie wybranych algorytmów prognostycznych. Przedstawiono możliwość wykorzystania tablicy niepodobieństwa modeli produktów, która pozwala na zmianę modeli produktów (wartości nominalnych) na odległości wykorzystywane w rozważanych podejściach. Pokazano również zmianę atrybutu opisującego rozmiar ubrania na wartość numeryczną.

W kolejnym rozdziale przedstawiono możliwości i zalety wykorzystania dodatkowych szeregów czasowych w celu poprawy skuteczności prognozowania. Po pierwsze, pokazano możliwość wykorzystania analogicznych szeregów czasowych z innych lokalizacji. Przedstawiono różne możliwości grupowania i dodawania informacji z podobnych szeregów czasowych. Dla obu badanych zestawów danych wykazano, że tworzenie modeli prognostycznych z wykorzystaniem dodatkowych danych poprawia wyniki uzyskanych prognoz. Następnie zwrócono uwagę na możliwości wykorzystania informacji o szeregach czasowych z tej samej lokalizacji w celu poprawy jakości prognoz. Podejście to samo w sobie było innowacyjne dla problemu prognozowania zużycia energii elektrycznej. Takie podejście można było zbadać dzięki stworzeniu modelu cyfrowego bliźniaczego budynku. Uzyskując dane o zużyciu energii elektrycznej z urządzeń gospodarstwa domowego i wprowadzając je do modelu prognostycznego, osiągnięto poprawę w prognozowaniu ogólnego zużycia energii. Wniosek z tych badań jest taki, że dodanie danych z innych lokalizacji lub tej samej lokalizacji może znacznie poprawić jakość prognoz.

Ostatnim rozważanym przypadkiem był problem prognozowania zwrotów produktów. Początkowo zaproponowano własny algorytm łączący transakcję zwrotu z transakcją zakupu. Podany algorytm był niezbędny do stworzenia zbioru danych treningowych. Jak pokazały późniejsze wyniki, dodanie informacji o transakcji zakupu i jej charakterystyce znacznie poprawiło wyniki prognoz. Uzyskano dobrą jakość klasyfikacji, potwierdzając skuteczność zaproponowanego rozwiązania.

Przedstawiona rozprawa doktorska nie koncentruje się na typowym zagadnieniu prognozowania szeregów czasowych, a skupia się na specyficznych zagadnieniach dotyczących wybranych, ale bardzo ważnych, aspektów prognozowania zapotrzebowania. Ze względu na zakres wykorzystanych danych oraz omówionych kwestii, prezentowana wiedza może stanowić cenny wkład w badania nad wykorzystaniem metod uczenia maszynowego do rozwiązywania biznesowych problemów związanych z prognozowaniem zapotrzebowania. Rozprawa doktorska była realizowana w ramach projektu InterPOWER finansowanego ze środków Unii Europejskiej (POWR.03.05.00-00-Z305).