

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. PŚ

Rozprawa doktorska bada wykorzystanie metod uczenia maszynowego i wiedzy dziedzinowej do rozwiązywania wybranych zagadnień związanych z prognozowaniem zapotrzebowania. Zrozumienie i dokładne prognozowanie zapotrzebowania na produkty 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 będzie poświęcona różnym aspektom związanym z tym tematem. Niektóre z poruszanych aspektów zostały zidentyfikowane jako luka badawcza w badaniach nad metodami uczenia maszynowego.

W pracy przedstawiono oryginalny podział zagadnień związanych z prognozowaniem zapotrzebowania. Każde zadanie zostało opisane wraz z konkretnymi problemami. Taki podział może pokazać szerszy obraz zagadnień związanych z prognozowaniem zapotrzebowania. Następnie zwrócono uwagę na wybrane zagadnienia 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 do tworzenia 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 nowatorskie podejście do analizy jakości wskaźników przy użyciu oryginalnej metody indukcji przeżyciowych reguł akcji.

Kolejnym poruszonym tematem była kwestia podziału prognoz wyższego rzędu na prognozy niższego rzędu. Jest to prognozowanie odgórne (*top-down forecasting*). Przeprowadzono porównanie wybranych algorytmów prognostycznych. Przedstawiono możliwość wykorzystania tablicy

product types dissimilarity table dla typów produktów, która pozwala na zmianę modeli produktów (wartości nominalnych) na odległości wykorzystywane w rozważanych podejściach.

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 szeregach czasowych. Dla obu badanych zestawów danych wykazano, że tworzenie modeli prognostycznych z wykorzystaniem wszystkich danych poprawia wyniki uzyskanych prognoz. Następnie przyjrano się zagadnieniu prognozowania, w którym dodano informacje o szeregach czasowych z tej samej lokalizacji. Podejście to samo w sobie było innowacyjne dla badanej dziedziny, a mianowicie prognozowania zużycia energii elektrycznej. Można je było zbadać dzięki stworzeniu cyfrowego modelu bliźniaczego budynku. Uzyskując dane o zużyciu energii elektrycznej z urządzeń gospodarstwa domowego i wprowadzając je do modelu prognozowania, 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. Przegląd literatury wskazał, że jest to potencjalna luka badawcza. Zgodnie z naszą najlepszą wiedzą, wcześniejsze badania nie dotyczyły przewidywania prawdopodobieństwa zwrotu określonego produktu. Początkowo zaproponowano niestandardowy algorytm łączący transakcję zwrotu z transakcją zakupu. Podany algorytm był niezbędny do stworzenia zestawu treningowego. Jak pokazały późniejsze wyniki, dodanie informacji o transakcji zakupu i jej charakterystyce znacznie poprawiło wyniki przewidywania. Uzyskano dobrą jakość klasyfikacji, potwierdzając skuteczność zaproponowanego procesu.

Przedstawiona rozprawa 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 przekrój wykorzystanych danych i poruszonych zagadnień, jest możliwość, że przedstawiona wiedza będzie cennym wkładem w badania nad uczeniem maszynowym i społeczności biznesowe.

Rozprawa doktorska była realizowana w ramach projektu InterPOWER finansowanego ze środków Unii Europejskiej (POWR.03.05.00-00-Z305).