zeT 1

Abstract

Title: Forecasting prices for road freight transport services using machine learning"

This doctoral dissertation focuses on developing a new method for forecasting prices in road freight transport, based on machine learning tools and comprehensive data analysis. The adopted approach involved collecting, transforming, and verifying information derived from transport offers, which was then integrated with external databases (including fuel prices, road tolls, the number of registered trailers, and macroeconomic data). The analysis showed that distance remains the primary factor influencing price, yet historical data on specific transport relations, additional costs (e.g., ferry or tunnel fees), and selected operational characteristics—such as vehicle type or trailer type—also play a significant role.

The algorithms used achieved a significantly lower forecast error (MAPE) compared to traditional procedures and expert estimates (9.28% vs. 11.86%). Incorporating external databases (fuel prices, road tolls, and the number of trailers) further reduced the prediction error (to as low as 7.90% for the best model). The study encompassed over 45,000 records, and the stability of these results was confirmed through various validation methods (including cross-validation and time-based splitting). The methodology—encompassing detailed data cleaning and normalization, feature engineering, and model testing—can be effectively transferred to other markets and transport segments. Additionally, options were proposed for implementing the findings in logistics/forwarding companies and in training future specialists.

During the research, a new machine learning-based method for forecasting road freight transport prices was developed, whose effectiveness was confirmed by achieving a model forecast error (MAPE) of 9.28%, compared to 11.86% obtained by experts. Various machine learning algorithms were compared, revealing that six of them maintain an average error below 10%. Solutions were proposed for incorporating external databases (fuel, road tolls, trailer count, GDP), which significantly improved model performance—especially when fuel prices, road tolls, and trailer counts were considered simultaneously. In the best scenario, the error was reduced from 8.11% to 7.90%. The developed procedures can be applied both in the operations of different types of transport companies and in the educational process. The findings show that implementing machine learning in pricing road freight transport services not only enhances forecasting accuracy but also leads to significant operational savings.