

Predictive Analytics in Business Intelligence for Sales Forecasting

Kshitij Dixit

Independent Researcher

kshitijdixit102@gmail.com

Abstract: *The accurate competitive advantage of businesses depends on accurate sales forecasting. The increased availability of retail sales data and the emergence of machine learning (ML) algorithms offer a chance to improve prediction accuracy. The purpose of the proposed research is to create an efficient ML model to predict weekly retail sales based on the Walmart dataset. To extract meaningful temporal and economic features out of the data, the study uses data preprocessing, exploratory data analysis, and feature engineering. Random Forest and CatBoost as ensemble learning models are used to forecast the pattern of sales every week. The models are evaluated using various measures like R^2 , MAE, RMSE, and MSE. Both models are quite predictive of the experimental results. Random Forest has the best R^2 0.9690 and MAE of 0.369 and CatBoost has the best R^2 0.9842 and MAE of 0.317 with low forecasting error and high prediction accuracy. It is also noted that the proposed models are better in the comparison with the base models such as XGB, LSTM, Extra Trees, SARIMA, and Prophet. Overall, the present research work presents a potent ML model, which integrates both temporal engineering of features and ensemble models to make the correct decisions in terms of retail sales. The findings highlight the importance of historical sales trends and extraneous factors in forecasting, and provide realistic insights into how demand planning and inventory use can be improved in the retailing business.*

Keywords: Retail Analytics, Sales Forecasting, Business Analytics, Product Demand, Machine Learning, Walmart Dataset