

ABSTRACT

Stock price movement prediction is an important task in financial analysis to help investors, analysts, and portfolio managers make informed decisions. This study classifies weekly stock price movements into three categories: up, stagnant, and down, using the XGBoost algorithm. Ten feature configurations were tested, including a combination of historical data (Close), technical indicators (ARMA/ARIMA), and fundamental ratios such as Debt to Equity Ratio (DER), Price to Book Value (PBV), and Price to Earnings Ratio (PE). In addition, the use of PCA for dimensionality reduction was also evaluated. Correlation analysis using the Spearman method was performed to determine the most relevant fundamental features, which showed that PBV had the best correlation with the Close (Movement) classification. However, the experimental results showed that DER was closer to the Baseline performance than PBV. The results showed that the Baseline configuration (Close + ARMA/ARIMA) gave the best performance with an average accuracy of 75% and an F1-Score of 74%. The addition of fundamental features made a limited contribution, with DER approaching the Baseline performance, while PBV and PE did not have a significant impact. PCA proved to be less effective, with an accuracy of only 37%. The study set a threshold of 3% to capture weekly volatility, reduce noise, and focus on significant price movements. These findings suggest that historical and technical data are more powerful than fundamental data in predicting stock price movements, while also opening up opportunities for developing more complex models and exploring fundamental features to improve prediction accuracy.

Keywords: *classification, machine learning, XGBoost, stock market, finance*