

Abstract

This research introduces a comprehensive machine learning framework aimed at dynamic stock portfolio diversification, which incorporates clustering, time-series forecasting, classification, and optimization methodologies. The dataset consists of stocks from the Indonesia Stock Exchange, augmented with technical indicators including EMA, RSI, OBV, and beta. Stocks were grouped quarterly utilizing K-Means, DBSCAN, and Agglomerative Clustering based on metrics of return and risk. Representative stocks from each cluster were chosen based on their predictive Sharpe Ratio accuracy, which was forecasted through a multivariate Long Short-Term Memory (LSTM) model that utilized MACD and inflation data. Additionally, a Support Vector Machine (SVM) was trained on beta values to categorize market regimes into upward, downward, and sideways trends. These regime labels were integrated into a regime-aware portfolio optimization strategy employing a mean-variance approach. Experimental findings reveal that K-Means clustering consistently surpasses other methods in producing balanced portfolios with advantageous Sharpe Ratios. Optimization based on market regimes yielded the most favorable outcomes during sideways market conditions. The results illustrate that the integration of unsupervised learning, predictive modeling, and regime classification can significantly improve portfolio construction in fluctuating market environments.

Keywords: Stock Portfolio, Clustering, LSTM, SVM, Market Regime, Portfolio Optimization, Sharpe Ratio