Machine Learning Forecasting of Indonesian Stock Return: Firm Characteristics and Macroeconomic Risk
Fauzan Alifian Subekti, Gigih Fitrianto, S.E., M.Sc., Ph.D.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

ABSTRACT

Standard asset pricing models often face limitations in handling high-dimensional variables, multicollinearity, and unstable return behaviour, particularly in emerging markets such as Indonesia. In response to these limitations, this study applies a regularized machine learning framework to improve stock return predictability using broader firm-level characteristic and macroeconomic information, specifically the Ridge, Lasso, and Elastic Net models. This study analyzes the out-of-sample prediction of Indonesian stock returns using panel data of firm and macroeconomic variables from January 2006 to December 2025. The forecasting framework compares three variable sets, which are defined as firm-level set, firm-level and macroeconomic set, as well as firm–macro interaction set. The model performance is evaluated using Root-Mean-Squared Error, Mean-Absolute Error, and out-of-sample R-squared over the 2020–2025 testing period, with STL-ARIMA model used as a univariate time-series benchmark. Moreover, the results show that regularized machine learning models generate positive out-of-sample performance with firm-level set as the best performing specification, particularly the expanding-window Lasso and Ridge models. Adding macroeconomic variables provides selective insights during specific market regimes, especially during crisis and uncertainty periods. The benchmark comparison further shows that the machine learning models achieve lower RMSE, lower MAE, and higher out-of-sample R-squared than STL-ARIMA. This outperformance suggests that stock return predictability is not captured solely by past time-series patterns, but also by how firm characteristics and macroeconomic conditions reflect future uncertainty. Furthermore, variable-importance and selection results indicate that book-to-market, operating profitability, investment, size, market index return, and BI rate change are among the most relevant predictors. Lastly, robustness checks using bootstrap resampling, Monte Carlo simulation, 10-fold month-clustered cross-validation, and heteroskedasticity-adjusted targets confirm that the main findings are stable. Overall, this study shows that penalized machine learning provides a robust and interpretable forecasting framework for Indonesian stock returns, with firm-level variables serving as the most reliable source of stock return prediction, followed by macroeconomic variables during market stress and high-uncertainty periods.

Keywords: *Asset Pricing, Stock Return Forecasting, Machine Learning, Macroeconomic Factors, Indonesia Stock Exchange.*