

ABSTRACT

Unemployment is a strategic issue that significantly impacts Indonesia's social stability, economy, and public welfare. Accurate forecasting of unemployment rates is crucial to support the formulation of effective policies. However, the complexity of economic factors such as economic growth, inflation, labor force participation rate (LFPR), and population often hinders reliable data-driven forecasting.

Common forecasting methods like ARIMA have limitations as they rely solely on historical data and fail to capture the influence of external factors. This creates a gap that must be addressed to generate more comprehensive and accurate unemployment forecasts in Indonesia.

This study compares two time series modeling approaches, namely ARIMA and ARIMAX. ARIMAX integrates external variables such as GDP and LFPR to enhance prediction accuracy. The modeling process was conducted carefully using grid search to determine optimal parameters for both models, based on 28 annual data points sourced from BPS, the World Bank, and Bank Indonesia. The models were evaluated using RMSE, AIC, and MAPE metrics, and an interactive web-based forecasting interface was implemented.

The results show that the best ARIMA model is ARIMA(0,2,2) with an AIC value of 605.57, RMSE of 656,103, and MAPE of 6.30%. Meanwhile, the best ARIMAX model is achieved by combining GDP and LFPR as exogenous variables with the ARIMAX(4,0,5) configuration, producing an RMSE of 246,237 and a MAPE of 2.74%. These findings indicate that incorporating exogenous variables significantly improves forecasting accuracy.

Keywords: *unemployment, forecasting, ARIMA, ARIMAX, exogenous variables, model evaluation*