# Forecasting Indonesia's Youth Unemployment Rate


Imaji Kasih Ayunda Rizal

S5348943


Thesis Supervisor: Dr. Noman Haleem


BSc Data Science and Society

Campus Fryslân, University of Groningen

## Abstract

This study aims to efficiently forecast Indonesia's youth unemployment rate. Youth unemployment is a worldwide socioeconomic challenge. As one of the world's emerging economies, the distinct conditions, including skill mismatches, prevalence of informal jobs, and regional disparities contribute to youth unemployment. Despite its widespread implications, forecasting this specific demographic has been overlooked. This study addresses this gap by applying univariate time series forecasting models using Python's PyCaret library. Two modeling approaches were explored: a baseline and a feature-engineered setup incorporating autoregressive lags and seasonal indicators. Models were assessed primarily using MASE and MAPE. While feature engineering improved some models, results varied, with further parameter tuning sometimes leading to overfitting. The tuned Exponential Smoothing model performed best, with a MAPE of 7.11%. These findings can guide targeted interventions to mitigate youth unemployment and promote socioeconomic stability.



**Keywords:** Youth Unemployment, Time Series Forecasting, Indonesia, PyCaret

# Table of Contents

# Abbreviations

## Institutions and Related Keywords

| | |
|---|---|
| **ALMPs** | Active Labour Market Policies |
| **BPS** | Statistics Indonesia (Badan Pusat Statistik) |
| **ILO** | International Labour Organisation |
| **OECD** | Organisation for Economic Co-Operation and Development |
| **SAKERNAS** | National Labour Force Survey |

## Time Series Models and Methodologies

| | |
|---|---|
| **AR** | AutoRegressive |
| **ARIMA** | Auto-Regressive Integrated Moving Average |
| **ARIMAX** | Auto-Regressive Integrated Moving Average with eXogenous variables |
| **BVAR** | Bayesian Vector AutoRegressive |
| **ES** | Exponential Smoothing |
| **ETS** | Error, Trend, Seasonality |
| **GARCH** | Generalised Autoregressive Conditional Heteroscedasticity |
| **HES** | Holt's Exponential Smoothing |
| **MA** | Moving Average |
| **SARIMA** | Seasonal Autoregressive Integrated Moving Averages |
| **SES** | Single Exponential Smoothing |
| **SMA** | Simple Moving Average |
| **VEC** | Vector Error Correction |
| **WMA** | Weighted Moving Average |

| | |
|---|---|
| **ANN** | Artificial Neural Network |

## Evaluation Metrics

| | |
|---|---|
| **AIC** | Akaike Information Criterion |
| **MAE** | Mean Absolute Error |
| **MAD** | Mean Absolute Deviation |
| **MAPE** | Mean Absolute Percentage Error |
| **MASE** | Mean Absolute Scaled Error |
| **MSE** | Mean Squared Error |
| **$R^2$** | Coefficient of Determination |
| **RMSE** | Root Mean Squared Error |
| **RMSSE** | Root Mean Squared Scaled Error |
| **SBC** | Schwarz Bayesian Criterion |
| **SMAPE** | Symmetric Mean Absolute Percentage Error |

## Time Series Statistical Analysis Terms

| | |
|---|---|
| **ACF** | Autocorrelation Function |
| **ADF** | Augmented Dickey-Fuller |
| **EDA** | Exploratory Data Analysis |
| **KPSS** | Kwiatkowski-Phillips-Schmidt-Shin |
| **PACF** | Partial Autocorrelation Function |
| **STL** | Seasonal-Trend decomposition using Loess |

## Other Terms

| | |
|---|---|
| **GDP** | Gross Domestic Product |
| **NEET** | Not in Education, Employment, or Training |

# 1 Introduction

Unemployment is a worldwide socioeconomic issue faced by countries all over the world. The term itself is defined as the proportion of the labour force that is without work but available and looking for work (*Labour Force Statistics (LFS, STLFS, RURBAN Databases) - ILOSTAT*, 2024). It is measured by the unemployment rate, which is calculated as the number of unemployed individuals divided by the total number of people in the labour force (*Labour Force Statistics (LFS, STLFS, RURBAN Databases) - ILOSTAT*, 2024). The unemployment rate is a critical macroeconomic indicator, playing a central role in forecasting future labor market performance and overall economic health (Huruta, 2024). While the overall country's unemployment rate gives an overview of the labour market, it does not show the disparities within demographic groups. A key demographic that needs particular consideration is the youth. Youth unemployment rate tends to exceed those of a higher age group, both in developing and developed countries. This emphasises systemic barriers the youth face when transitioning into the workforce (Görlich et al., 2013). Apart from the immediate economic impact, early unemployment can have lasting repercussions such as skill deterioration and harmful signalling to future employers, which may limit long-term employability and earnings potential (O'Higgins, 2007). Governments are increasingly aware that effectively integrating the youth into the labour force is critical not only for economic growth, but also for social cohesion and political stability (Görlich et al., 2013; Sachs & Smolny, 2015).

Indonesia is a developing country with the fourth largest population in the world. Facing a demographic bonus, a large share of the population is part of the working-age group (15-64 years old), accounting for over 70% of the total population (BPS, n.d.). Youth is defined as those of the ages between 15 and 29 (*Statistics on Youth - ILOSTAT*, 2025). Indonesia's youth makes up over a quarter–26.5%–of the total labour force (BPS, n.d.-a). Youth unemployment presents a particularly important challenge. If employed effectively, this workforce could drive economic growth. However, persistently high youth unemployment, combined with a predominance of informal, precarious jobs among the youth, risks converting this potential into a demographic burden (Yanindah, 2022). The long-term impact of youth unemployment is significant. When the youth have difficulties finding employment, this can lead to lower income and productivity, and increase the likelihood of poverty–not just for them as individuals, but also to the future generations. It also increases the strain on social support systems and hinders the country's progress (Yanindah, 2022). Addressing youth unemployment is thus vital, not only to promote social inclusion and reduce inequality, but also to protect Indonesia's long-term socioeconomic stability and effectively maximise on the demographic dividend before it declines.

Given the risks, it is essential to understand and anticipate labour market trends that affect youth unemployment. Existing studies have explored forecasting Indonesia's general unemployment rate. Syafwan et al. (2023) applied an ARIMA model projecting gradual decline in unemployment and demonstrating the model's effectiveness. Similarly, Isityani et al. (2023) used ARIMA to capture pandemic-related fluctuations. Huruta (2024) introduced demographic variables into the forecasting models and links the unemployment to changes in population structure. The existing literature has largely focused on national-level trends. There are also studies related to forecasting the unemployment rate on the provincial level. Gustriansyah et al. (2023) compared exponential smoothing approaches, focusing on the province of South Sumatra. Didiharyono and Syukri (2020) used the ARIMA model to examine unemployment rates in

South Sulawesi. There remains a large gap in forecasting efforts that specifically target youth unemployment.

Youth unemployment has a large effect on multiple aspects of society, yet forecasting models tailored to this demographic in Indonesia remain scarce. This study aims to fill that gap by answering the research question

**RQ How can the youth unemployment rate in Indonesia be effectively forecasted?**

To achieve this, the study utilises the PyCaret library in Python, comparing available forecasting models through a baseline setup and a feature-engineered setup. The best performing model will be finalised for future predictions, evaluated by a set of error metrics for forecasting. The findings hold relevance for policymakers in establishing employment programs to more efficiently address youth unemployment, ensuring Indonesia's demographic potential translates into sustained growth in the economy, and potentially discovering gaps between education and workforce demands. The private sector can use optimised projections to evaluate labour market resilience and plan workforce needs. More broadly, the research contributes to discussions about macroeconomic forecasting.

The subsequent sections begin by discussing the existing literature on youth unemployment studies, notably in Indonesia, youth unemployment forecasting methods globally, and research on unemployment forecasting in Indonesia. The data sources and methodology are then described along with its limitations and ethical considerations, followed by the forecasting model results, their respective comparisons, and future projections with the final model. The concluding parts explore the findings' implications, emphasising the strengths and limitations of the selected approach, as well as proposing recommendations for future research.

# 2 Literature Review

This study aims to explore how youth unemployment rate in Indonesia can be forecasted with the available data. Unemployment forecasting in Indonesia has been more widely researched, particularly using time series models. They have, however, paid little attention to age-specific trends, emphasising instead on national or provincial unemployment rates. Simultaneously, youth unemployment has become a major issue in both developed and developing countries. Some research has looked at the causes and long-term effects of youth unemployment, particularly in emerging economies where labour markets are generally more unpredictable and informal. This literature review thus places the current research within three main areas: theoretical and empirical studies on youth unemployment, approaches to youth unemployment forecasting globally, and forecasting efforts particular to the Indonesian labour market. In doing so, it reveals existing gaps and establishes the rationale for forecasting youth unemployment in Indonesia.

## 2.1 Youth Unemployment

Youth unemployment has garnered traction in the research literature due to its persistent nature and long-term consequences. The International Labour Organisation (ILO) and the Organisation for Economic Co-operation and Development (OECD) typically considers youth to be those between the ages of 15 and 24, especially in the context of employment and labor market analysis (ILO, 2022; OECD, n.d.). Youth unemployment is then the percentage of unemployed youth within the age group (15-24). Other organisations, such as the European Commission, define youth as those between the ages of 15 and 29 years old (European Commission: Directorate-General for Education, Youth, Sport and Culture, 2021). This captures a wider segment of youth transitioning into the labor market. The exact definition of youth differs depending on the organisation, research objective, and context. For Indonesia, Statistics Indonesia (BPS) defines youth as those between the ages of 16 and 30, according to the Law No. 40 of 2009 (Yanindah, 2022).

There are several drivers of youth unemployment. Macroeconomic aspects such as economic growth and shifts in the business cycle have a substantial influence on youth labour market outcomes, with declines leading to greater rises in youth unemployment than those of other age groups (Kokotović, 2016; Kang, 2021). Prospects of youth unemployment are also heavily influenced by labour market regulations, such as employment protection legislation and minimum wage policies (Destefanis & Mastromatteo, 2010; Kang, 2021). Furthermore, the employability of young workers is also impacted by the quality and structure of human resources, particularly regarding commitments in skill development and vocational training (Kang, 2021). In this context, education systems, particularly dual education and vocational programs, are emphasised as critical avenues to more efficiently transition from school-to-work and reduce youth unemployment (Kokotović, 2016; Alfonsi et al., 2020; Kang, 2021). Demographic aspects, such as the ratio of youth to elderly, also influence youth unemployment rates, which frequently magnify labour market issues (Kang, 2021). Another important aspect is the broader economic structure of the country, particularly the sectoral composition of work opportunities, for example, the share of service sector occupations (Kang, 2021). From the aforementioned factors to youth unemployment, Gomez-Salvador and Leiner-Killinger (2008) identified the conditions associated with lower youth employment rates. Their findings suggest that youth unemployment tends to be lower in countries with a smaller youth population relative to the total population, stronger economic growth, less stringent employment protection, more extensive active labour market policies (ALMPs), a large share of employment in the service sector, and lower youth labour force participation.

However, there are differences in the main drivers to youth unemployment in developed and emerging markets. This is due to differences in economic structures, labor market institutions, and the level of human capital development. Developed markets are characterised by their advanced economies with high income, stable institutions, and mature financial systems (Glen & Singh, 2004; Shankar & Narang, 2020). Labour market segmentation, stringent employment protection, and rapid technological development all hinder youth employment, limiting entry-level prospects during downturns. Although ALMPs and robust institutions help to mitigate these consequences, persistent disparities in education and labour market needs remain a significant concern (Kang, 2021). Emerging markets are countries in transition, experiencing rapid growth but with lower income and less mature institutions. Emerging markets often face more fragmented and incomplete labor markets than their developed counterparts (Glen & Singh, 2004). Structurally, these economies tend to deal with market fragmentation, governance issues, limited resources, unregulated competition, and poor infrastructure (Sheth, 2011). The difficulties include skill mismatches, insufficient vocational training, and institutional shortcomings in labour markets (Sharma, 2022). These restrict the effective integration of youth into the workforce.

As Southeast Asia's largest economy, Indonesia has demonstrated strong economic performance and expansion, strengthening its position as one of the world's most dynamic emerging markets (World Bank, 2023). Similar to other emerging market economies, Indonesia faces substantial structural and developmental challenges related to youth unemployment. In particular, according to BPS, a considerable number of the unemployed youth have low educational backgrounds or have not attended certified training programs (as cited in Yanindah, 2022). Furthermore, the majority of jobs held by young people in Indonesia are informal and require few specialized skills, often translating to unstable employment and poor prospects for career development (Allen, 2016). Paradoxically, despite education and skill development being recognized as crucial for mitigating youth unemployment in Indonesia, observations show that highly educated young people frequently experience joblessness. This situation points to a substantial mismatch between the skills supply and labor market demand, compounded by fierce competition and a deficit in relevant training (Yanindah, 2022). Easier access to employment often comes with higher education, yet the shortage of practical experience and specialized training continues to hinder job placement, thereby stressing the importance of developing programs that better integrate skills with current market demands (Sitompul & Athoillah, 2023). The situation is further complicated by existing gender disparities and regional variations. Specifically, young men face a 0.9% lower likelihood of unemployment compared to young women, and youth residing in Java are 1.9% more susceptible to unemployment than their counterparts outside Java (Yanindah, 2022). Despite the multitude of open positions, there is more competition and higher skill requirements (Muladi et al., 2018). The concentration of economic activity in urban areas, particularly on Java Island, produces excess labour (Sitompul & Athoillah, 2023). These factors not only shape the youth labour market in Indonesia but also contribute to the broader, often intergenerational, impacts of youth unemployment.

Youth unemployment affects multiple aspects of society. The economic impact of youth unemployment includes reduced productivity and stifled innovation and growth (Kang, 2021). Socially, it amplifies the danger of unrest and creates a risk of long-term detachment from employment, which could lead to a 'lost generation' trapped in cycles of inequality and exclusion (Marelli et al., 2013). At a personal level, enduring youth unemployment can result in skill deterioration, diminished self-worth, interrupted career trajectories, and a greater propensity for mental health issues (O’Higgins, 2007; Görlich et al., 2013; Başol et al., 2023). These individual struggles can translate into a higher probability of persistent unemployment in later life, ultimately eroding both economic progress and social cohesion. Therefore, it is important to accurately forecast youth unemployment rate as this can inform targeted policy interventions that could mitigate the long-term socioeconomic risks.

## 2.2 Youth Unemployment Forecasting Methods

Multiple studies have forecasted youth unemployment rates through various methods including time series models, econometric models, machine learning approaches, and hybrid models. For instance, Simionescu and Cifuentes-Faura (2022) used Google Trends data and employed Bayesian vector autoregressive (BVAR), vector error correction (VEC), Bayesian panel data, and fixed-effect models to forecast national and regional youth unemployment rate in Spain. Sharma and Soni (2021) focused on predicting youth unemployment rates in India using various time series models including Auto-Regressive Integrated Moving Average (ARIMA), Exponential smoothing methods including Simple Exponential Smoothing (SES), Holt's linear trend, and Error, Trend, Seasonality (ETS). Atanasova-Pacemska et al. (2015) employed linear and exponential trend models to forecast youth unemployment in the Republic of Macedonia. Jung (2018) evaluated the efficiency of incorporating web search query information from Naver and Google into ARIMA models to forecast Korea's youth unemployment rate. Meanwhile, Fenga and Son-Turan (2022) focused on forecasting youth (Not in Education, Employment, or Training (NEET)) unemployment in Italy during and after the COVID-19 pandemic by developing a counterfactual scenario using an Artificial Neural Network (ANN) model, that also incorporates Google Trends data as an exogenous variable.

Despite variations in geographic scope and methodological detail, these studies have a number of core characteristics that provide useful insights for forecasting youth unemployment. A significant pattern observed is the frequent application of time series methods, especially ARIMA and its variations, which consistently appear in research from India (Sharma & Soni, 2021), Korea (Jung, 2018), and Macedonia (Atanasova-Pacemska et al., 2015), highlighting its dependability as a fundamental forecasting model. Several studies then use trend-based or exponential smoothing strategies to compare model accuracy and suitability. However, the reliance on historical data limits explanatory power and responsiveness to sudden shocks. One of the ways to address this, is that recent studies show the increasing use of new (real-time) data sources like Google Trends. Studies by Jung (2018), Simionescu & Cifuentes-Faura (2022), and Fenga & Son-Turan (2022) show that digital data like this can help make forecasts more accurate and up-to-date. These strategies are especially useful when official labour statistics are typically lagging and youth labour market data is incomplete or delayed. For example, such data helped anticipate the 2008 crisis (Fenga & Son-Turan, 2022) and outperformed baseline models (Jung, 2018; Simionescu & Cifuentes-Faura, 2022). Third, the literature emphasises the integration of exogenous variables and hybrid or machine learning models, such as Fenga & Son-Turan's (2022) ANN approach, to capture nonlinear effects and economic shocks. However, there are other limitations that remain. Some models omit broader socioeconomic variables, faced constraints from short forecast horizons or reliance on low-frequency data, and struggled with instability when training and prediction periods were discontinuous. Improvements could include the use of richer explanatory variables, higher-frequency data, and the application of hybrid or machine learning models.

To assess the accuracy and efficiency of youth unemployment forecasts, standard evaluation metrics are utilized. The academic literature frequently relies on a combination of error-based and goodness-of-fit statistics, with RMSE, MAE, and MAPE serving as primary indicators of forecast precision in both traditional time series and advanced econometric models (Jung, 2018; Davidescu et al., 2021; Simionescu & Cifuentes-Faura, 2022). Commonly, mean squared error (MSE) is also employed to assess forecasting model performance (Sharma & Soni, 2021), and was notably used as a cost function in an ANN by Fenga and Son-Turan (2022). The coefficient of determination ($R^2$) frequently accompanies these metrics, providing insight into a model's goodness-of-fit and explanatory power (Jung, 2018). Advanced evaluations might incorporate metrics like Theil's U1 and U2 coefficients, as utilised by Simionescu and Cifuentes-Faura (2022) in BVAR models, or formal hypothesis tests such as the Diebold-Mariano test to confirm significance of model performance variations, as observed in Davidescu et al.'s (2021) study on Romanian unemployment. Prior to modeling, assessing correlations between potential explanatory variables, such as Google Trends data, and unemployment rates can serve as an initial check for variable relevance and selection (Jung, 2018).

## 2.3 Unemployment Forecasting in Indonesia

While studies on youth unemployment forecasting globally have utilised a variety of models ranging from traditional time series to more advanced machine learning approaches, they provide methodological insights that can inform similar initiatives in Indonesia. However, the varying contexts must also be taken into account to determine which model would be most suitable for Indonesia's setting. Therefore, this subsection discusses existing research on unemployment rate forecasting in Indonesia, covering model choices, data constraints, and applicability for youth-focused use cases.

Mahmudah (2017) applied an ARIMA model to annual unemployment rate data from 1986 to 2015. Similarly, Istiyani et al. (2023) forecasted national unemployment rate from 2005 to 2022 with an ARIMA model. Huruta (2024) tested six alternative ARIMA specifications on national data from 1990 to 2022 in the context of Indonesia's demographic dividend. Orisa and Faisol (2024) employed  Simple Moving Average (SMA) model to forecast unemployment, segmented by education-level with data from 1986 to 2022, while Syafwan et al. (2023) used a Weighted Moving Average (WMA) approach on annual unemployment rate data from 2000 to 2022. The WMA approach assigns different weights to historical data, giving more importance to recent values to better capture trends. There are also studies that perform forecasting on provincial-level unemployment rate. Didiharyono and Syukri (2020) used ARIMA for South Sulawesi's unemployment rates, combining annual and semi-annual data from 1986 to 2018. Gustriansyah et al. (2023) focused on South Sumatra, comparing exponential smoothing methods: Single Exponential Smoothing (SES), Brown's Exponential Smoothing (BES), and Holt's Exponential Smoothing (HES) on biannual data from 2008 to 2020. Aside from that, Ng et al. (2023) included Indonesia in a comparative study between five Southeast Asian countries using ARIMA, Seasonal Autoregressive Integrated Moving Average (SARIMA), and Generalised Autoregressive Conditional Heteroscedasticity (GARCH) models based on monthly unemployment data from 2010 to 2021. Notably, Fajar et al. (2020) incorporated Google Trends as an exogenous variable to forecast the unemployment rate. The data used was from 1986 to 2020 interpolated into monthly data from annual and biannual data, and consequently applied an AutoRegressive Integrated Moving Average with eXogenous variables (ARIMAX) model.

These studies collectively reveal a reliance on univariate time series forecasting approaches, with ARIMA models appearing as the most commonly utilised. ARIMA's popularity comes from its efficiency in handling non-stationary time series through differencing, use of past values for accurate short-term forecasting, its simplicity and wide acceptance in economic studies (Mahmudah, 2017; Istiyani et al., 2023; Ng et al., 2023). Simple and weighted moving average approaches are additionally employed in research that prioritise ease of implementation, while exponential smoothing has appeared in regional applications. Orisa and Faisol (2024), begin to tackle segmentation, specifically by education level, implying further modelling nuance. However, a recurring drawback in most studies is the use of historical unemployment rates without accounting for exogenous socioeconomic or policy variables, reducing their ability to account for unexpected shocks or structural changes. Fajar et al. (2020) used Google Trends data and the term 'phk'—a local acronym for layoffs—as an exogenous variable to produce a near-real-time representation of public interest in layoffs, which can serve as an early signal ahead of official unemployment data. The study however was limited to ARIMAX. Huruta (2024) and Mahmudah (2017) demonstrated ARIMA's short-term strengths, but both emphasise the importance of integrating demographic-specific and regionally disaggregated data, as well as advanced alternative methods such as machine learning models to handle randomness and structural changes over longer forecast horizons. The majority also rely on coarse temporal data (annual or biannual) given that this is what Statistics Indonesia or *Badan Pusat Statistik* (BPS) makes publicly available, with only Fajar et al. (2020) applying interpolation to existing data to produce monthly-level data and Ng et al. (2023) using monthly datasets from Bank Indonesia. In conclusion, a recurring theme in these studies is the need for models that can better handle complex, nonlinear unemployment dynamics, demographic variability, and longer forecast horizons. Some literature recommends the inclusion of varied datasets and more specific unemployment data (i.e., by location, age, and gender).

Assessing the accuracy and performance of unemployment rate forecasts in Indonesia commonly involves several evaluation metrics. MAPE is the predominant choice in much of the existing literature (Fajar et al., 2020; Syafwan et al., 2023; Huruta, 2024; Orisa & Faisol, 2024). Alongside MAPE, MAE, MSE, and RMSE are also widely applied in studies utilizing methods like WMA, Exponential Smoothing, and various ARIMA models (Gustriansyah et al., 2023; Ng et al., 2023; Syafwan et al., 2023). Other relevant metrics include the Mean Absolute Deviation (MAD), which calculates the simple average of absolute forecast errors (Syafwan et al., 2023), and Mean Absolute Scaled Error (MASE), which scales errors against a baseline method's average error (Hyndman, 2006; Mahmudah, 2017). Beyond the aforementioned error measures, evaluation in unemployment forecasting sometimes incorporates other metrics like Theil's Inequality Coefficient and Symmetric Mean Absolute Percentage Error (SMAPE), which offer enhanced insights into error scaling and directionality (Ng et al., 2023). Model selection criteria, including the Akaike Information Criterion (AIC) and Schwarz Bayesian Criterion (SBC), are also vital, particularly for identifying suitable ARIMA model orders (Huruta, 2024). Moreover, rigorous diagnostic checks such as the Augmented Dickey-Fuller (ADF) test for stationarity and the Ljung-Box test for white noise residuals are utilised for validating model assumptions (Mahmudah, 2017; Istiyani et al., 2023). This integrated approach, encompassing a variety of metrics and diagnostic tools, supports solid predictions for Indonesia's unemployment rate.

While these studies provide the foundation into unemployment forecasting within the context of Indonesia, youth unemployment in particular remains overlooked, despite being a separate and significant policy challenge with distinctive underlying factors. This gap emphasises the significance of not only using classical time series methods, but also exploring machine learning-based forecasting models to assess their comparative efficacy for capturing the dynamic patterns that characterise youth labour market dynamics.

# 3 Methodology

## 3.1 Data

The data used in this research are secondary data from BPS. BPS retrieved the data through the National Labour Force Survey (SAKERNAS), specifically designed to collect employment data on an ongoing basis. The unemployment rates are reported biannually in February and August. The dataset used was data on the labour force by age group for the period of February 2008 to

August 2024. This was retrieved through the Web API of BPS. BPS defined youth as those between the ages of 16 and 30 (Yanindah, 2022). Data access was limited to information on the total labor force, the proportions of employed and unemployed persons, and the employment rate, which was computed as the ratio of employed individuals to the total labor force. Age segmentation within this data was structured in regular five-year bands, beginning at 15–19 and progressing through 60+, with a final category representing the total across all age groups.

$$Youth\ Unemployment\ Rate\ =\ \left(\frac{Unemployed\ Youth}{Total\ Youth\ Labour\ Force}\right)\ \times\ 100\% \qquad \ldots (1)$$

To collect the youth unemployment rate data, the share of the unemployed and the total labour force for the age groups 15-29 were aggregated. Then, it is calculated by dividing the number of unemployed individuals by the total number of people in the labour force and multiplying it by 100 to express it as a percentage, as seen in equation (1).

## 3.2 Research Method

### 3.2.1 Data Collection and Preprocessing

The data collection and analysis for this study were primarily conducted using Python. A python script was created to collect the youth unemployment rate through the Web API from BPS. The collected raw data was preprocessed first by aggregating relevant age groups and calculating the youth unemployment rate, according to the equation (1). The resulting time series data was then formatted into a dataframe consisting of the time periods and the unemployment rate values. For consistency in time series modeling, the biannual periods of February and August were designated as Q1 and Q3, respectively.

### 3.2.2 Exploratory Data Analysis

The analysis started with retrieving descriptive statistics and performing exploratory data analysis (EDA). The EDA includes analysing data distribution, trends and volatility, and inherent seasonality. Stationarity tests, such as the ADF test and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test were performed to ensure the time series met the assumptions for several forecasting models. Additionally, white noise tests and normality tests were conducted to validate model assumptions. Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) analyses were utilised to measure the correlation between the time series and its lagged values, supporting identification of trends, seasonality, and potential autoregressive (AR) and moving average (MA) components (Box et al., 2015; Hyndman & Athanasopoulos, 2021).

### 3.2.3 Forecasting Approach and Model Setup

The study covers univariate time series forecasting, which is to forecast Indonesia's youth unemployment rate. The model training and evaluation was facilitated by PyCaret, an open-source and low-code machine learning library in Python that automates machine learning workflows. There were two different forms of setups.

**Baseline Setup**

The first setup, known as the 'baseline setup', only included the original time series dataset with the unemployment rate values. The forecast horizon was set to six periods ahead, which is approximately 20% of the length of the full data period or equivalent to three years given the biannual frequency.

**Feature-Engineered Setup**

The second setup included explicit feature engineering. This setup involved manually including columns that indicate *Lag 1* (unemployment rate of one period ago), *Lag 2* (unemployment rate of two periods ago), *rolling mean 3* (average unemployment rate of the previous three periods), and *quarter* (an indicator for the February or August period). The inclusion of such autoregressive features were based on the performed EDA. The rows with missing values caused by the feature engineering were subsequently removed. An additional difference was the explicit specification of the seasonal period in the PyCaret configuration.

### 3.2.4 Model Evaluation and Selection

The forecasting models were compared using well-established evaluation metrics, quantifying the accuracy and suitability. These included Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), Symmetric Mean Absolute Percentage Error (SMAPE), Mean Absolute Scaled Error (MASE), Root Mean Squared Scaled Error (RMSSE), and the Coefficient of Determination ($R^2$). The equations for these metrics are provided in the following equations.

$$MAE = \frac{1}{n}\sum_{t=1}^{n}\left|y_t - \hat{y}_t\right| \qquad \ldots(2)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(y_t - \hat{y}_t\right)^2} \qquad \ldots(3)$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_t - \widehat{y}_t}{y_t}\right| \times 100\% \qquad \ldots(4)$$

$$SMAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|y_i - \widehat{y}_i\right|}{\frac{\left|y_i\right| + \left|\widehat{y}_i\right|}{2}} \times 100\% \qquad \ldots(5)$$

$$MASE = \frac{\frac{1}{n}\sum_{t=1}^{n}\left|y_t - \widehat{y}_t\right|}{\frac{1}{n-1}\sum_{t=2}^{n}\left|y_t - y_{t-1}\right|} \qquad \ldots(6)$$

$$RMSSE = \sqrt{\frac{\frac{1}{n}\sum_{t=1}^{n}\left(y_t - \widehat{y}_t\right)^2}{\frac{1}{n-1}\sum_{t=2}^{n}\left(y_t - y_{t-1}\right)^2}} \qquad \ldots(7)$$

$$R^2 = 1 - \frac{\sum_{t=1}^{n}\left(y_t - \widehat{y}_t\right)^2}{\sum_{t=1}^{n}\left(y_t - \widehat{y}\right)^2} \qquad \ldots(8)$$

with $y_t$ = the actual value at time *t*

$\widehat{y}_t$ = the forecasted value at time *t*

$n$ = the number of data points (observations)

MAE provides the average absolute difference between actual values and predictions. RMSE is similar to MAE but the errors are squared, therefore it penalises larger errors more. MAPE shows the average absolute percentage error. SMAPE is a scale-independent version of MAPE, preventing issues when the actual value at time *t* is near zero. MASE scales the MAE, comparing the model to a naive baseline. RMSE is similar to MASE but uses squared errors. Finally, the $R^2$ quantifies the proportion of variance that is explained by the model (Quispe et al., 2024).

### 3.2.5 Finalisation and Future Forecasting

Following the initial comparison, the models exhibiting the higher performance across the metrics were selected for fine-tuning. Based on the outcomes of the fine-tuning, the best-performing model was finalised and consequently employed to forecast future youth unemployment rates with the same forecast horizon that is six periods ahead.

## 3.3 Limitations and Ethical Considerations

The study acknowledges several limitations that prevail. First, the available data was relatively limited, which may affect the generalisability and robustness of the forecasting result. As one of the first attempts to forecast youth unemployment rate for Indonesia, there is limited prior research for direct benchmarking. For the context of Indonesia, there is more research available for forecasting the general unemployment rate or at the provincial level. The models were also only built using the time series itself, or features derived from the time series. It does not incorporate exogenous variables such as Gross Domestic Product (GDP) growth, inflation rate, or education level which are known to be correlated with unemployment trends (Phillips, 1958; Gomez-Salvador & Leiner-Killinger, 2008; Kang, 2021; Sitompul & Athoillah, 2023). Finally, the study mostly utilises the PyCaret library. This streamlined the modeling process, especially the comparisons between different forecasting models. It could be beneficial to see if the depth of custom model tuning or feature experimentation can be better achieved through manual approaches.

The obtained data is publicly available and does not include any sensitive or personally identifiable information. The study complies to ethical research procedures, notably in terms of data management and transparency, by complying with the FAIR principles (Wilkinson et al., 2016). A public GitHub repository hosts all materials. Data and code are stored with metadata in a structured repository with a persistent URL and clear documentation. All resources are made available via open and standardised protocols. Standard file formats and open source Python libraries ensure compatibility across platforms. The repository includes detailed provenance, clear usage terms (MIT license), and follows community standards to support future research and adaptation.

# 4 Results and Discussion

## 4.1 Descriptive Statistics

| | **Youth Unemployment Rate** |
|---|---:|
| count | 34 |
| mean | 14.615368 |
| min | 11.81219 |
| 25% | 13.048628 |
| 50% | 14.249481 |
| 75% | 15.957687 |
| max | 18.374399 |
| std | 1.826307 |

**Table 1.** Descriptive statistics of Indonesia’s youth unemployment rate (2008 - 2024)

Based on the data obtained from BPS on Indonesia's youth unemployment rate, the average unemployment rate across 34 data points is 14.6. The lowest and highest unemployment rate within the timeframe is 11.8 in February 2017 and 18.4 in August 2008 respectively. Table 1 presents the descriptive statistics of this dataset.

A series of time series assumption checks were also performed on the dataset, including testing for white noise, stationarity, and normality. The p-value of the Ljung-Box test shows that it strongly rejects the null hypothesis of white noise. This suggests that the series contains autocorrelation and is not purely random. For stationarity, both the ADF and the KPSS test show that the data is non-stationary. The ADF test fails to reject the null hypothesis of a unit root and the KPSS test rejects the null of stationarity. The Shapiro-Wilk test for normality indicates that the data does have a normal distribution. Table 2 shows the detailed results of these tests.

| **Test** | **Test Name** | **Property** | **Value** |
|---|---|---|---|
| White Noise | Ljung-Box | Test statistic | 68.190447 |
| | | p-value | 0.000004 |
| | | White noise | 0 |
| Stationarity | ADF | Test statistic | -1.951027 |
| | | p-value | 0.30845 |
| | | Stationarity | FALSE |
| | KPSS | Test statistic | 0.163636 |
| | | p-value | 0.035303 |
| | | Stationarity | FALSE |
| Normality | Shapiro-Wilk | p-value | 0.2377 |
| | | Normality | TRUE |

**Table 2.** White noise, stationarity, and normality test results

## 4.2 Trend and Seasonality

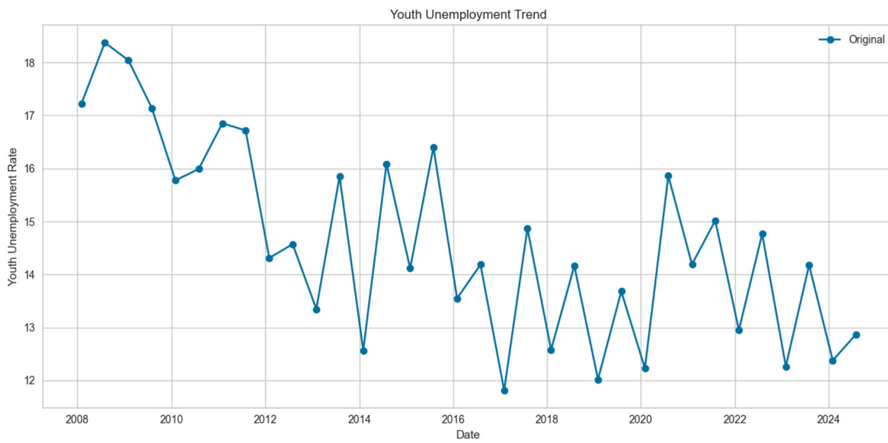


**Figure 1.** Indonesia's youth unemployment trend from 2008 - 2024

Figure 1 shows the trend of the data during the period ranging from February 2008 to August 2024 and Figure 2 shows the results of the Seasonal-trend decomposition using Loess (STL) on the time series. The actual time series shows seasonal fluctuations. The 2008 period was the highest during the global financial crisis. There is also a spike during the COVID-19 period, around 2020. Overall, there is a gradual declining trend which suggests long-term improvement in youth unemployment. The seasonal panel shows strong and consistent seasonal patterns with oscillations between +1 and -1. The residuals appear relatively random and scattered around zero, suggesting that most of the systematic patterns are captured through the decomposition. Figure 3 shows a boxplot and a line plot separating the time series based on its reported time period: February and August. The boxplot shows higher youth unemployment in August compared to February. Both the median and the interquartile range for August is higher. This could be due to new graduates entering the labour market in the middle of the year. The line plot also confirms that August unemployment is consistently higher than February, except for two instances in 2009 and 2011.

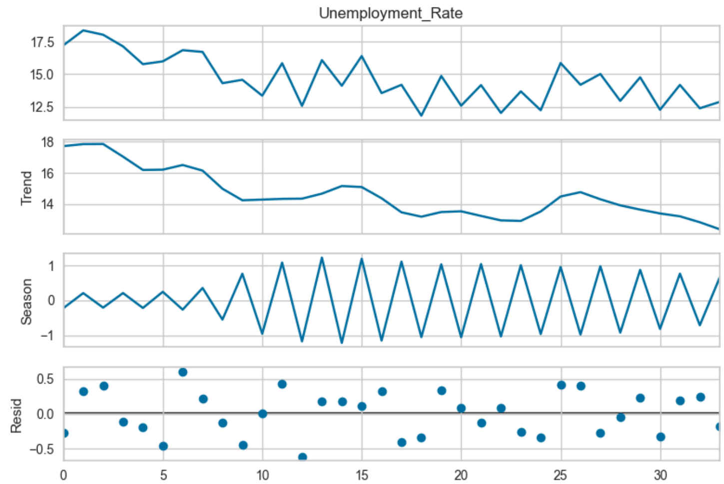


**Figure 2.** STL decomposition of the youth unemployment rate time series

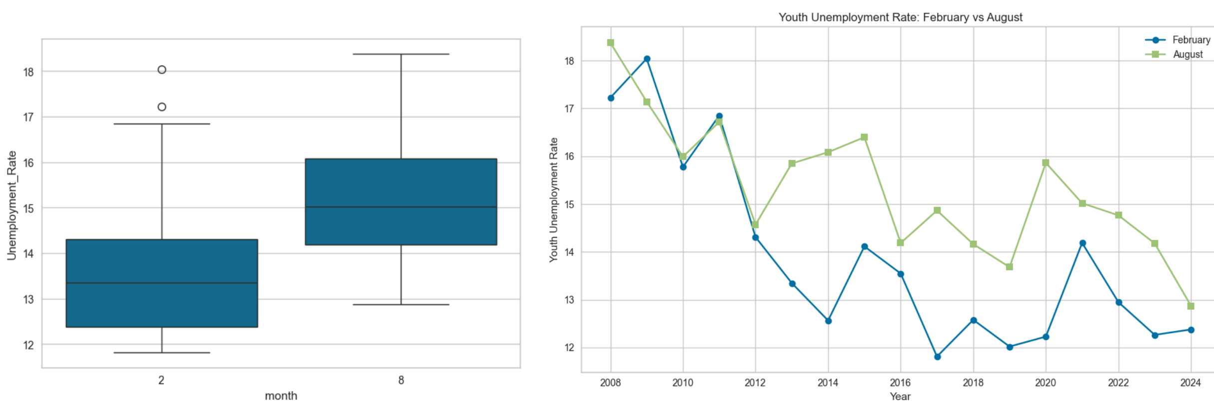


**Figure 3.** Box-plot of time series per month (left). Left box-plot shows February and right box-plot shows August. Youth unemployment trend by month (right). Blue line shows February and green line shows August.

## 4.3 Autocorrelations

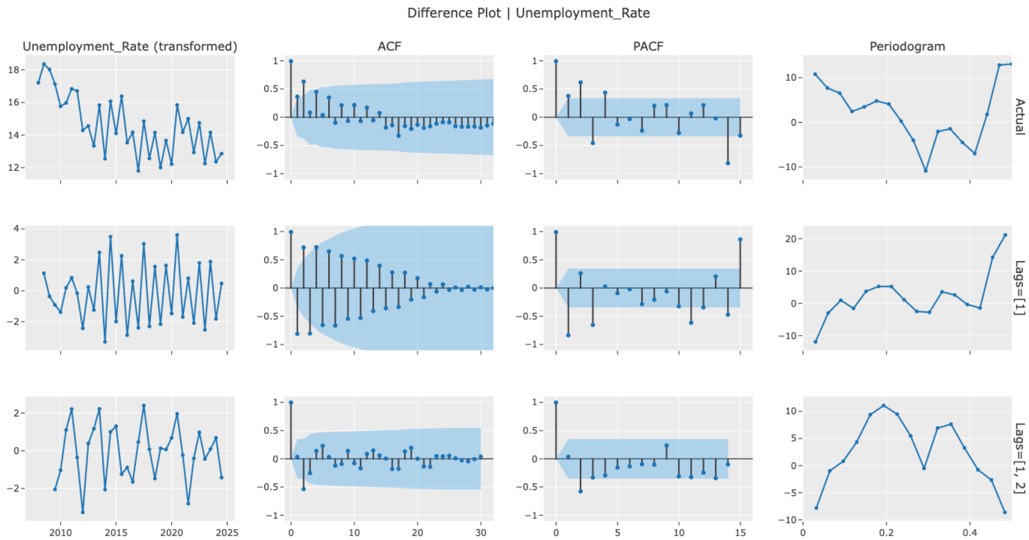


**Figure 4.** Time series, ACF, PACF and periodogram plot for the original data, first order differencing, and second order differencing

ACF measures the correlation between a time series and its lagged values, whereas PACF evaluates the correlation after accounting for intermediate lags. ACF serves in identifying trends and seasonality, as well as MA components. PACF helps detect the AR order (Hyndman & Athanasopoulos, 2021). Figure 4 shows the time series, ACF, PACF, and periodogram of three stages: the original series (first row), the first-order differenced series (second row), and the second-order differenced series (third row). A differenced series is the change between consecutive observations and is commonly used to achieve stationarity (Box et al., 2015; Hyndman & Athanasopoulos, 2021). The ACF plot of the original data shows a gradual decline, indicating non-stationarity and the PACF plot shows significance at Lag 1 and Lag 2. This motivated the application of first-order differencing to transform it into a stationary time series. After first differencing, the values are centered around zero and the variability becomes more visible. There is a significant negative spike at Lag 1 in the ACF plot. Additionally, there are significant spikes at seasonal lags, especially at Lag 2 and 4 in the ACF plot. The second differencing further stabilises the fluctuations, appearing to be largely white noise with most spikes in both ACF and PACF falling within the confidence intervals meaning that it no longer shows strong autocorrelation and is likely stationary.

While PyCaret handles the order selection and differencing internally when using ARIMA models, the results of the initial analysis are what motivated the inclusion of several features for the feature-engineered setup. The significance of Lag 1 and Lag 2 lead to the inclusion of Lag 1 and Lag 2 as lag features to capture short-term dependencies, especially for regression models. The addition of the rolling average was to also capture smoothed trends. Finally, time-based features such as the month were considered to distinguish between seasonal reporting periods.

## 4.4 Model Comparisons

PyCaret streamlines the comparison of the available time series forecasting models. Table 3 shows the detailed results of the model evaluations for the baseline setup and Table 4 shows the detailed results of the model evaluations for the feature engineered setup. The differences in which models are included in the comparison of each setup are automated through PyCaret itself. Some model names are ended with the suffix 'cds_dt'. This stands for conditional deseasonalise and detrending, which refers to the internal preprocessing pipeline that prepares the data before training. PyCaret automatically tests for seasonality, only applying deseasonalisation if needed as well as removing trends. The complete list of model abbreviations names can be found in Appendix A.

Baseline Model Comparisons

| | MASE | RMSSE | MAE | RMSE | MAPE | SMAPE | R2 |
|---|---|---|---|---|---|---|---|
| gbr_cds_dt | 0.833 | 0.9551 | 1.092 | 1.4067 | 0.0743 | 0.0779 | -0.1589 |
| ada_cds_dt | 0.8409 | 0.9835 | 1.1013 | 1.4448 | 0.0746 | 0.0789 | -0.2393 |
| rf_cds_dt | 0.9436 | 1.0214 | 1.2417 | 1.5066 | 0.0847 | 0.089 | -0.3301 |
| knn_cds_dt | 0.9632 | 1.038 | 1.2737 | 1.53 | 0.0869 | 0.0916 | -0.3866 |
| theta | 0.9769 | 1.0526 | 1.282 | 1.5464 | 0.0888 | 0.0913 | -0.4254 |
| dt_cds_dt | 1.0302 | 1.1336 | 1.3261 | 1.6492 | 0.091 | 0.0962 | -0.6565 |
| exp_smooth | 1.0487 | 1.1314 | 1.3759 | 1.6604 | 0.0932 | 0.0997 | -0.6452 |
| ets | 1.0487 | 1.1314 | 1.3759 | 1.6605 | 0.0932 | 0.0997 | -0.6452 |
| lr_cds_dt | 1.0521 | 1.1315 | 1.3771 | 1.6599 | 0.0938 | 0.0996 | -0.6527 |
| omp_cds_dt | 1.0598 | 1.1344 | 1.3861 | 1.6645 | 0.0944 | 0.1003 | -0.6612 |
| ridge_cds_dt | 1.0668 | 1.136 | 1.3988 | 1.6673 | 0.0953 | 0.1013 | -0.6668 |
| huber_cds_dt | 1.0696 | 1.1431 | 1.3987 | 1.6764 | 0.0953 | 0.1014 | -0.6867 |
| br_cds_dt | 1.0785 | 1.1427 | 1.4128 | 1.6765 | 0.0964 | 0.1025 | -0.6814 |
| en_cds_dt | 1.0906 | 1.1479 | 1.4309 | 1.6851 | 0.0975 | 0.104 | -0.6983 |
| lasso_cds_dt | 1.0906 | 1.1479 | 1.4309 | 1.6851 | 0.0975 | 0.104 | -0.6983 |
| llar_cds_dt | 1.0906 | 1.1479 | 1.4309 | 1.6851 | 0.0975 | 0.104 | -0.6983 |
| lightgbm_cds_dt | 1.0906 | 1.1479 | 1.4309 | 1.6851 | 0.0975 | 0.104 | -0.6983 |
| et_cds_dt | 1.1619 | 1.2504 | 1.4907 | 1.8135 | 0.1022 | 0.1092 | -0.9731 |
| arima | 1.1761 | 1.2914 | 1.5496 | 1.911 | 0.1053 | 0.1116 | -1.1563 |
| polytrend | 1.2239 | 1.271 | 1.593 | 1.8563 | 0.1092 | 0.1158 | -1.0236 |
| croston | 1.3969 | 1.467 | 1.803 | 2.1441 | 0.1374 | 0.1247 | -2.1796 |
| naive | 1.4274 | 1.381 | 1.8151 | 1.9833 | 0.1348 | 0.1252 | -2.5713 |
| grand_means | 1.473 | 1.5536 | 1.8909 | 2.2602 | 0.1447 | 0.1306 | -2.4637 |

**Table 3.** PyCaret time series forecasting model comparisons for the baseline setup

Feature Engineered Model Comparisons

| | MASE | RMSSE | MAE | RMSE | MAPE | SMAPE | R2 |
|---|---|---|---|---|---|---|---|
| et_cds_dt | 0.8726 | 0.9339 | 1.1416 | 1.3724 | 0.078 | 0.0817 | -0.2018 |
| omp_cds_dt | 0.8839 | 0.9887 | 1.1541 | 1.4532 | 0.0791 | 0.0822 | -0.2562 |
| ridge_cds_dt | 0.8945 | 0.9805 | 1.1792 | 1.4487 | 0.0824 | 0.0835 | -0.257 |
| en_cds_dt | 0.9207 | 1.0103 | 1.2113 | 1.4836 | 0.0827 | 0.0869 | -0.3235 |
| lasso_cds_dt | 0.9207 | 1.0103 | 1.2113 | 1.4836 | 0.0827 | 0.0869 | -0.3235 |
| llar_cds_dt | 0.9207 | 1.0103 | 1.2113 | 1.4836 | 0.0827 | 0.0869 | -0.3235 |
| lightgbm_cds_dt | 0.9207 | 1.0103 | 1.2113 | 1.4836 | 0.0827 | 0.0869 | -0.3235 |
| dt_cds_dt | 0.9411 | 1.0412 | 1.2028 | 1.51 | 0.0822 | 0.0848 | -0.3459 |
| rf_cds_dt | 0.9533 | 1.0171 | 1.244 | 1.4882 | 0.0852 | 0.0896 | -0.3517 |
| knn_cds_dt | 0.981 | 1.0219 | 1.2801 | 1.4987 | 0.088 | 0.0925 | -0.3618 |
| br_cds_dt | 0.9907 | 1.0623 | 1.3184 | 1.571 | 0.0923 | 0.0927 | -0.5031 |
| gbr_cds_dt | 1.0411 | 1.1095 | 1.3442 | 1.612 | 0.0926 | 0.0963 | -0.5256 |
| ada_cds_dt | 1.0415 | 1.0984 | 1.3299 | 1.5862 | 0.0933 | 0.094 | -0.4947 |
| arima | 1.1356 | 1.3067 | 1.4905 | 1.962 | 0.1029 | 0.1133 | -1.4251 |
| croston | 1.383 | 1.4493 | 1.7338 | 2.0847 | 0.1324 | 0.1205 | -2.0071 |
| huber_cds_dt | 1.7964 | 2.3043 | 2.2243 | 3.2259 | 0.1553 | 0.1402 | -12.4106 |

**Table 4.** PyCaret time series forecasting model comparisons for the feature engineered setup

For the baseline setup, the best performing models were Gradient Boosting and AdaBoost, with MASE values of 0.833 and 0.841 respectively. Both models also achieve a MAPE of under 8%. The best performing models for the feature engineered setup were Extra Trees and Orthogonal Matching Pursuit, with MASE values of 0.8726 and 0.8839 respectively. The MAPE for these two models are also just above 8%. $R^2$ should be interpreted with caution in this context due to the relatively small size of the dataset. With fewer data points, $R^2$ becomes more sensitive to slight variations and can produce misleading results. Therefore, greater emphasis will be placed on the other metrics, mainly on MASE and MAPE for its scale-independence and comparability. While the top models in the baseline setup perform better than those in the feature engineered setup, the MASE of a larger group of models are consistently below 1.0 in the feature engineered setup. Boosting models that performed best in the baseline, became less competitive after feature engineering, which may be overfitting due to the engineered features being redundant information or that the models had already captured interactions and patterns yet this conflicted with the manual features. Meanwhile, models such as 'et_cds_dt', 'omp_cds_dt', and 'ridge_cds_dt' improved as they may have benefited from richer features. However, the difference in performance between the two setups are not that significant despite the additional features. Furthermore, based on the results of these comparisons, several top performing models from both setups are fine-tuned.

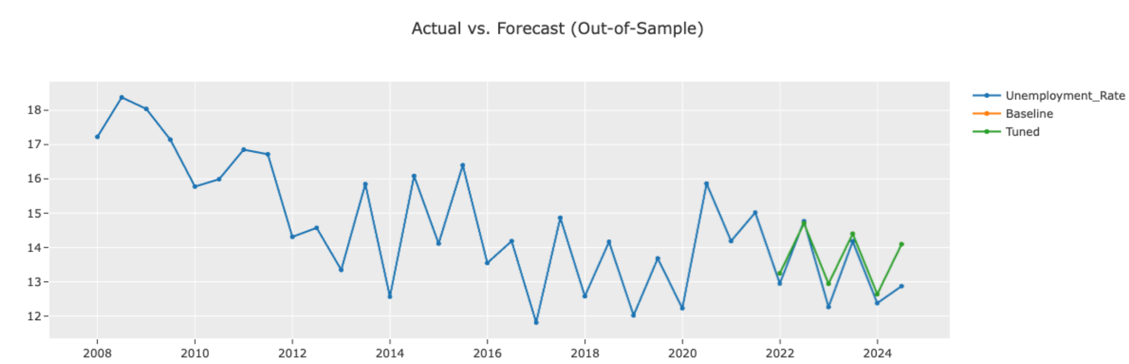


**Figure 5.** Hold-out set forecast for Gradient Boosting with Cond. Deseasonalise & Detrending (Baseline setup)

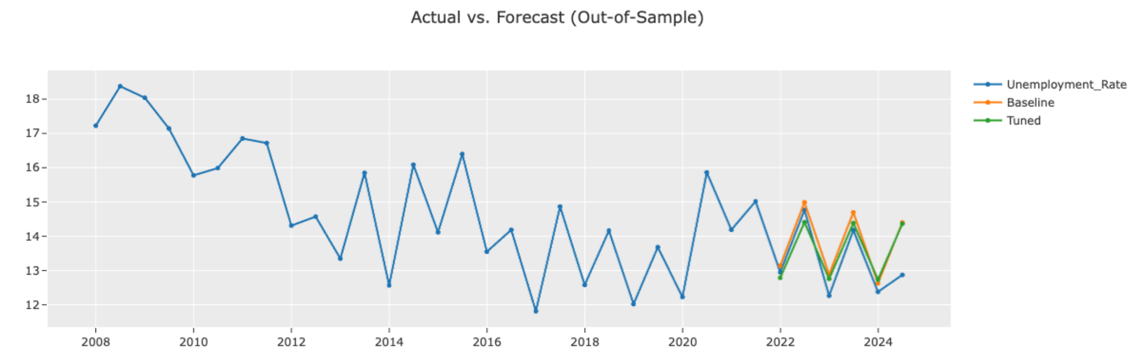


**Figure 6.** Hold-out set forecast for Exponential Smoothing (Baseline setup)

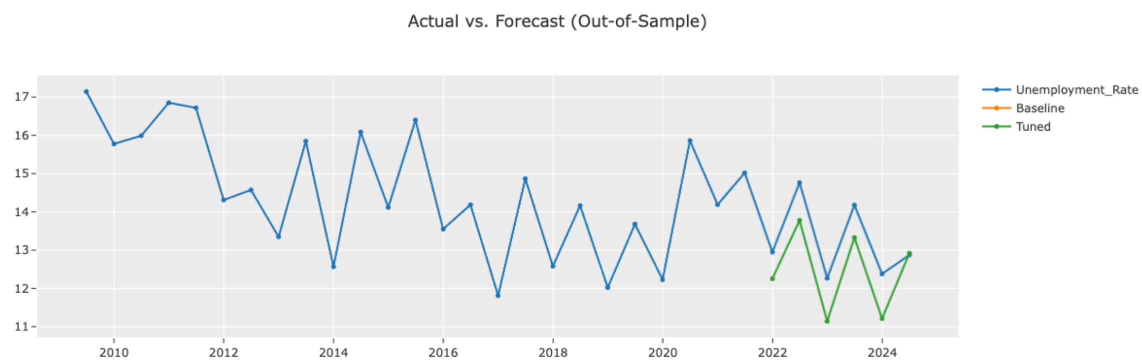


**Figure 7.** Hold-out set forecast for Extra Trees with Cond. Deseasonalise & Detrending (Feature engineered setup)

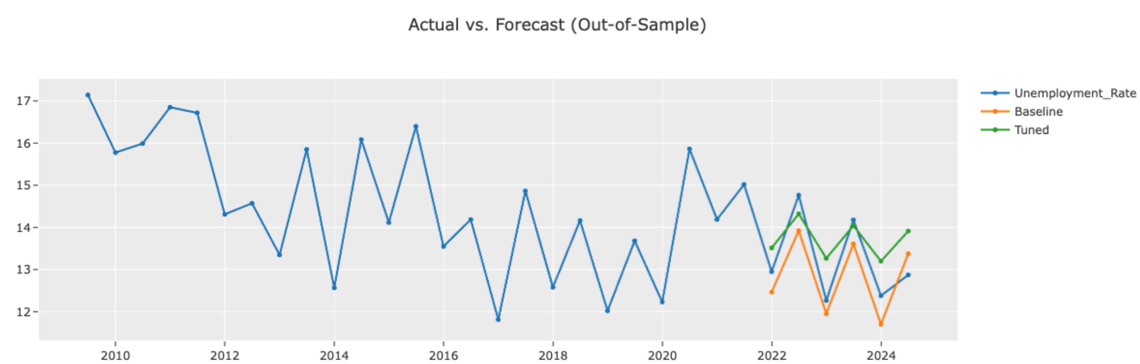


**Figure 8.** Hold-out set forecast for Orthogonal Matching Pursuit with Cond. Deseasonalise & Detrending (Feature engineered setup)

The evaluations of the top performing models and the fine-tuned versions reveal the best performing models by average MASE. The models included for fine tuning from the baseline setup are Gradient Boosting, AdaBoost, Random Forest, Exponential Smoothing, and ETS. This selection was a mix of including the top performing models as well as more traditional forecasting models to see how they perform after fine-tuning. The models included for fine tuning from the feature engineered setup were the top three models: Extra Trees, Orthogonal Matching Pursuit, and Ridge.

Tuned Exponential Smoothing achieved an average MASE of 0.7678. The baseline Gradient Boosting followed with 0.833. Figures 5 and 6 show the hold-out set forecast for the baseline and tuned model for Gradient Boosting and Exponential Smoothing respectively. Figures 7 and 8 show the hold-out set forecast using the feature engineered setup for Extra Trees and Orthogonal Matching Pursuit. The detailed cross-validation and hold-out set evaluation for the selected models can be seen in Appendix B. The visualisation for the hold-out set forecasts of the remaining fine-tuned models can be seen in Appendix C. The overall results also show a paradox in hyperparameter tuning. This can either hurt or improve the model. It was shown that simple models benefit more from tuning, while complex models may be overfitting considering the small data size. There were also insights from incorporating cross validation in the model evaluations. The first set period (2012Q3 cutoff) was the post-financial crisis period. This period showed relatively higher errors where most models seem to struggle more. Majority of the models perform well in the second set period (2015Q3 cutoff), this period is economically stable favoring all approaches. The final set period (2018Q3 cutoff) had mixed results, it can either be higher or lower than the first period. It was moderately challenging across different models. All models still performed best in the second period. For the hold-out set, the best performing models were the baseline Gradient Boosting (0.3967 MASE) and Random Forest (0.3862 MASE). Exponential smoothing maintains competitive performance (0.4429 MASE).

Exponential smoothing benefited significantly from fine-tuning. In exponential smoothing techniques, future values are estimated through a weighted average of historical data, with the weighting scheme characterized by an exponential decay, ensuring that more recent data exerts a stronger influence on the forecast (Hyndman & Athanasopoulos, 2021). The performance of the fine-tuned Exponential Smoothing model is very consistent across time periods. The mean and standard deviation metrics also show that this model had the lowest variability. Conforming with the analysis from the decomposition, the error, trend, and seasonal components can be captured well by the model. Tree-based methods such as Gradient Boosting, AdaBoost, Random Forest perform quite well but are more inconsistent. Despite the time series diagnostics suggesting suitability for ARIMA, it performed worse compared to the other models. The possible issues might lie in the implementation, inadequate feature engineering, or improper handling of structural breaks. The findings align with the parsimony principle in time series analysis, where simpler, well-calibrated models often outperform complex alternatives (Box et al., 2015). Based on the results of the fine-tuning, the tuned Exponential Smoothing model will be finalised and used for forecasting future periods. This model is configured with a multiplicative trend and an additive seasonal component with a seasonal period of two. Consequently, this implies that the seasonal fluctuations of the time series are assumed to be relatively constant over time. The trend's effect on the youth unemployment rate, however, is proportional to the current level of the series.

## 4.5 Future Predictions

| | y_pred |
|---|---|
| 2025Q1 | 12.5365 |
| 2025Q3 | 14.1263 |
| 2026Q1 | 12.5151 |
| 2026Q3 | 14.1073 |
| 2027Q1 | 12.4981 |
| 2027Q3 | 14.0921 |

**Table 5.** Future predictions with final model (tuned Exponential Smoothing) for six-periods ahea

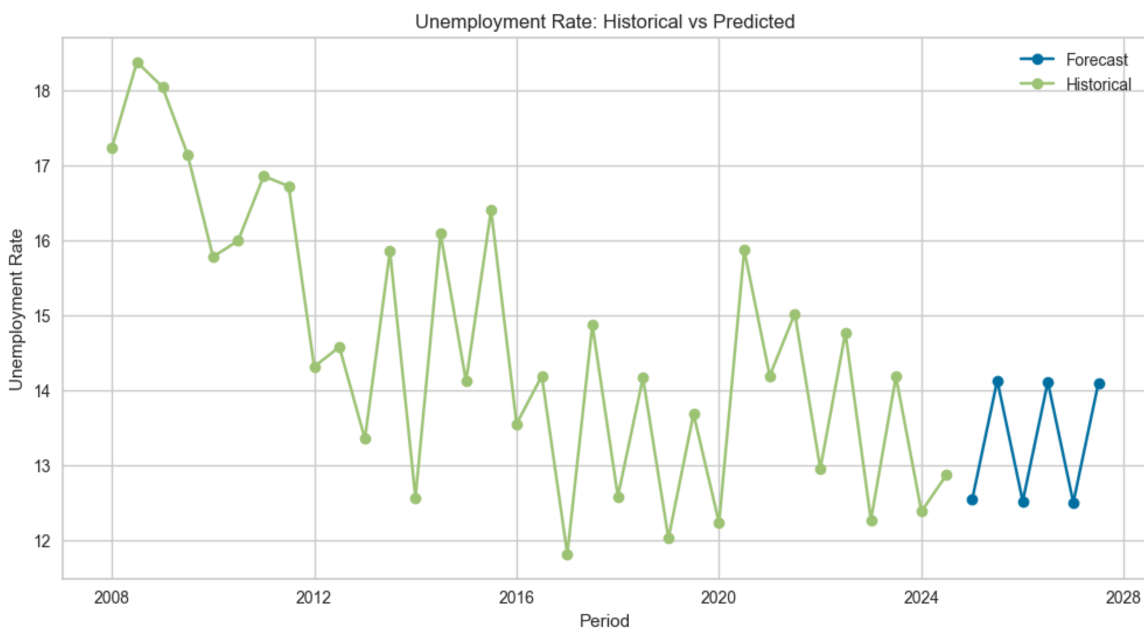


**Figure 9.** Plot of historical data and future predictions with final model (tuned Exponential Smoothing) for six-periods ahead

The tuned Exponential Smoothing model generates future youth unemployment rate forecasts that exhibit a distinct seasonal pattern with relatively stable long-term dynamics. The model predicts unemployment rates between approximately 12.5% in Q1 periods (February) and 14.1% in Q3 periods (August) across 2025 - 2027. This reflects the seasonal pattern of the historical data showing higher rates during the third quarter. There is also a slight downward trend in both seasonal components indicating gradual improvement.

This study contributes to the field by providing a novel benchmark for forecasting youth unemployment in Indonesia, a domain previously unexplored. A critical aspect is the limited data size of approximately 34 data points from 2008 to 2024. This aligns with several comparable studies in the literature ranging approximately from 20 to 50 data points (Mahmudah, 2017; Didiharyono & Syukri, 2020; Gustriansyah et al., 2023; Istiyani et al., 2023; Syafwan et al., 2023; Huruta, 2024). The consistent challenge of achieving positive $R^2$ highlights the inherent difficulty of capturing significant variance with sparse data. However, the fine-tuned Exponential Smoothing model consistently shows that it is possible to achieve positive predictive power even with a limited dataset. The top performing models in this study reach a MAPE of 7-9%, which according to literature, a MAPE value under 10% shows high performance or high accuracy (Fajar et al., 2020; Gustriansyah et al., 2023; Syafwan et al., 2023; Huruta, 2024). The tuned Exponential Smoothing reached a MAPE of 7.1% which is comparable to or better than other studies' results in forecasting Indonesia's (general) unemployment rate. This research offers a valuable tool for informing targeted policy interventions aimed at mitigating the significant economic and social risks associated with high youth unemployment, especially within

Indonesia’s unique emerging market context characterised by skill mismatches, regional, and demographic disparities.

## 4.6 Recommendations

To address the limitations, further research should prioritise expanding the dataset to include a longer time series if data becomes available, which would enhance model robustness and generalisability beyond the current available data points. Another aspect is to incorporate relevant exogenous variables, such as GDP growth, inflation rate, education levels, share of the population per age group, or even the share of the service sector in the economy (Gomez-Salvador & Leiner-Killinger, 2008) that correlate with unemployment trends could improve predictive accuracy. Furthermore, exploring manual, in-depth custom model tuning beyond PyCaret's automated processes, particularly for models like ARIMA, could result in additional performance advancements, especially given that this model and its variations performed well in previous research forecasting Indonesia's general unemployment rate. This includes looking further into the suitability of different seasonal periods for the ARIMA models, which was a point of interest in the exploratory analysis. It would be valuable to manually search for the optimal order for the ARIMA model or to select the most appropriate variation for this time series.

# 5 Conclusion

The study aimed to address the research question, “How can the youth unemployment rate in Indonesia be effectively forecasted?” This was achieved through a structured framework like PyCaret, streamlining the forecasting process, making use of both traditional forecasting models and machine learning models. The research contributed a novel benchmark for forecasting youth unemployment in Indonesia, a domain previously unexplored, providing critical insights for a demographic group disproportionately affected by labour market shifts. The forecasting framework could also be applied to predict the general unemployment rate or those characterised by other demographics. The results suggest that the tuned Exponential Smoothing model is the best method for forecasting youth unemployment rate in Indonesia, maintaining low error rates and stable performance across different time periods. The MAPE for this model is 7.11%.

By providing actionable insights into future unemployment trends among Indonesian youth, policymakers can proactively address critical issues. Effective forecasts can guide the development of programs aimed at skill development, vocational training, and active labour market policies, mitigating the consequences of prolonged youth unemployment. For future research, it is recommended to expand the dataset with a longer time series, incorporate exogenous variables correlated with youth unemployment trends, further exploration into manual and in-depth custom model tuning especially for traditional time series models.

## Code and Data Availability

The data collection and cleaning script, exploratory data analysis, model training, feature engineering, model evaluation, and forecasting notebooks, along with the processed dataset used in this study, are openly available in a GitHub repository at https://github.com/imajirzl/thesis-youth-unemployment-id.

# References

Alfonsi, L., Bandiera, O., Bassi, V., Burgess, R., Rasul, I., Sulaiman, M., & Vitali, A. (2020). Tackling youth unemployment: evidence from a labor market experiment in Uganda. *Econometrica*, *88*(6), 2369–2414. https://doi.org/10.3982/ecta15959

Allen, E. R. (2016). *Analysis of trends and challenges in the Indonesian labor market* (ADB Papers on Indonesia No. 16). Asian Development Bank. https://www.adb.org/publications/analysis-trends-and-challenges-indonesian-labor-market

Atanasova-Pacemska, T., Mitreva, E., & Lapevski, M. (2015). Application of economic-mathematical models for assessment of unemployment of young people. In: SYM OP IS 2015, *Proceedings of the XLII International Symposium on Operational Research*. Matematički Institut, SANU, Beograd, pp. 610-614. ISBN 978-86-80593-55-5

Başol, O., Sevgi, H., & Yalçın, E. C. (2023). The effect of digitalization on youth unemployment for EU countries: treat or threat? *Sustainability*, *15*(14), 11080. https://doi.org/10.3390/su151411080

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time Series Analysis: Forecasting and Control* (5th ed.). John Wiley & Sons.

BPS. (n.d.-a). *Labor Force (LF) by age group - statistical data*. BPS-Statistics Indonesia. https://www.bps.go.id/en/statistics-table/2/Njk4IzI=/labor-force--lf--by-age-group.html

BPS. (n.d.). *Percentage of work force on working age population by age group - statistical data*. BPS-Statistics Indonesia. https://www.bps.go.id/en/statistics-table/2/NzE0IzI=/percentage-of-work-force-on-working-age-population-by-age-group.html

Davidescu, A. A., Apostu, S., & Paul, A. (2021). Comparative Analysis of Different Univariate Forecasting Methods in Modelling and Predicting the Romanian Unemployment Rate for the Period 2021–2022. *Entropy*, *23*(3), 325. https://doi.org/10.3390/e23030325

Destefanis, S., & Mastromatteo, G. (2010). Labour‐market performance in the OECD: some recent cross‐country evidence. *International Journal of Manpower*, 31(7), 713–731. https://doi.org/10.1108/01437721011081563

Didiharyono, D., & Syukri, M. (2020). Forecasting with ARIMA Model In Anticipating Open Unemployment Rates in South Sulawesi. In Central Statistics Agency (BPS) South Sulawesi, *International Journal of Scientific & Technology Research* (Vol. 9, Issue 03, pp. 3838–3839). https://www.ijstr.org/final-print/mar2020/Forecasting-With-Arima-Model-In-Anticipating-Open-Unemployment-Rates-In-South-Sulawesi.pdf

European Commission: Directorate-General for Education, Youth, Sport and Culture & Petkovic, S. (2021). *Proposal for an updated dashboard of EU youth indicators : support services to Expert Groups and mutual learning activities in the field of youth policy (June 2020-December 2021) : final report*, Publications Office of the European Union. https://data.europa.eu/doi/10.2766/357817

Fajar, M., Prasetyo, O.R., Nonalisa, S., & Wahyudi, (2020). Forecasting Unemployment Rate in the Time of COVID-19 Pandemic Using Google Trends Data (Case of Indonesia). *International Journal of Scientific Research in Multidisciplinary Studies* , 6(11), 29-33.

Fenga, L., & Son-Turan, S. (2022). Forecasting youth unemployment in the aftermath of the COVID-19 pandemic: the Italian case. *International Journal of Scientific and Management Research*, *5*(1), 75–91. https://doi.org/10.37502/ijsmr.2022.5105

Glen, J., & Singh, A. (2004). Comparing capital structures and rates of return in developed and emerging markets. *Emerging Markets Review*, *5*(2), 161–192. https://doi.org/10.1016/j.ememar.2004.01.001

Gomez-Salvador, R., & Leiner-Killinger, N. (2008). An analysis of youth unemployment in the euro area. *ECB Occasional Paper*, No. 89, European Central Bank (ECB), Frankfurt a. M.

Görlich, D., Stepanok, I., & Al-Hussami, F. (2013). Youth unemployment in Europe and the world: Causes, consequences and solutions. *Kiel Policy Brief, No. 59.* Kiel Institute for the World Economy (IfW), Kiel.

Gustriansyah, R., Alie, J., & Suhandi, N. (2023). Modeling the number of unemployed in South Sumatra Province using the exponential smoothing methods. *Quality & Quantity*, *57*(2), 1725–1737. https://doi.org/10.1007/s11135-022-01445-2

Hasan, Zainul & Sasana, Hadi. (2020). Determinants Of Youth Unemployment Rate In Asean. *International Journal of Scientific & Technology Research*, 9(3), 6687-6691.

Huruta, A. D. (2024). Predicting the unemployment rate using autoregressive integrated moving average. *Cogent Business & Management*, 11(1). https://doi.org/10.1080/23311975.2023.2293305

Hyndman, R. J. (2006). Another Look at Forecast Accuracy Metrics for Intermittent Demand. *Foresight: The International Journal of Applied Forecasting*, 4, 43-46.

Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice* (3rd ed.). OTexts. OTexts.com/fpp3.

International Labour Organization (ILO). (2022). *Global employment trends for youth 2022: Investing in transforming futures for young people*. International Labour Office. Geneva. https://doi.org/10.54394/QSMU1809

Istiyani, N., Hanim, A., Dewi, I. R., & Najib, M. K. (2023). Accuracy of predicting the unemployment rate in Indonesia. *Journal of Economics Management and Trade*, *29*(6), 18–26. https://doi.org/10.9734/jemt/2023/v29i61096

Jung, J. U. (2018). Comparative usefulness of Naver and Google Search information in predictive models for youth unemployment rate in Korea. *Journal of Digital Convergence*, *16*(8), 169–179. https://doi.org/10.14400/jdc.2018.16.8.169

Kang, Y. (2021). Determinants of youth unemployment: Empirical analysis of OECD and EU member countries. *Journal of Economic Development,* The Economic Research Institute, Chung-Ang University, 46(3), 111-133. https://ideas.repec.org/a/ris/jecdev/0032.html

Kokotović, F. (2016). An empirical study of factors influencing total unemployment rate in comparison to youth unemployment rate in selected EU member-states. *Theoretical and Applied Economics,* Asociatia Generala a Economistilor din Romania / Editura Economica, vol. 0(3(608), A), 79-92. https://ideas.repec.org/a/agr/journl/vxxiiiy2016i3(608)p79-92.html

*Labour Force Statistics (LFS, STLFS, RURBAN databases) - ILOSTAT*. (2024, January 22). ILOSTAT. https://ilostat.ilo.org/methods/concepts-and-definitions/description-labour-force-statistics/

Mahmudah, U. (2017). Predicting unemployment rates in Indonesia. *Economic Journal of Emerging Markets*, *9*(1), 20–28. https://doi.org/10.20885/ejem.vol9.iss1.art3

Marelli, E., Choudhry, M. T., & Signorelli, M. (2013). YOUTH AND TOTAL UNEMPLOYMENT RATE: THE IMPACT OF POLICIES AND INSTITUTIONS. *Rivista Internazionale Di Scienze Sociali*, *121*(1), 63–86. http://www.jstor.org/stable/26477700

Muladi, N., Wibawa, A. P., & Moses, K. M. (2018). A study of the impact of cooperation between vocational high school and industries in Malang City. *Journal of Physics Conference Series*, *1028*, 012077. https://doi.org/10.1088/1742-6596/1028/1/012077

Ng, K. Y., Zainal, Z., & Samsudin, S. (2023). COMPARATIVE PERFORMANCE OF ARIMA, SARIMA AND GARCH MODELS IN MODELLING AND FORECASTING UNEMPLOYMENT AMONG ASEAN-5 COUNTRIES. *International Journal of Business and Society*, *24*(3), 967–994. https://doi.org/10.33736/ijbs.6393.2023

Organisation for Economic Co-operation and Development (OECD). (n.d.). *Youth unemployment rate*. https://doi.org/10.1787/c3634df7-en

Orisa, M., & Faisol, A. (2024). Forecasting Indonesia's unemployment rates using moving average methods. *Journal of Computer Science and Engineering (JCSE)*, *5*(1), 39–45. https://doi.org/10.36596/jcse.v5i1.793

O'Higgins, N. (2007). The challenge of youth unemployment. *International Social Security Review*, *50*(4), 63–93. https://doi.org/10.1111/j.1468-246x.1997.tb01084.x

Phillips, A. W. (1958). The Relation between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861-1957. *Economica*, *25*(100), 283-299. https://doi.org/10.2307/2550759

Quispe, J. O. Q., Quispe, A. C. F., Calvo, N. C. L., & Toledo, O. C. (2024). Analysis and Selection of Multiple Machine Learning Methodologies in PyCaret for Monthly Electricity Consumption Demand Forecasting. *Materials Proceedings,* 18(1), 5. https://doi.org/10.3390/materproc2024018005

Sachs, A., & Smolny, W. (2015). Youth unemployment in the OECD: the role of institutions. *Jahrbücher Für Nationalökonomie Und Statistik*, *235*(4–5), 403–417. https://doi.org/10.1515/jbnst-2015-4-505

Shankar, V., & Narang, U. (2020). Emerging market innovations: unique and differential drivers, practitioner implications, and research agenda. *Journal of the Academy of Marketing Science*, *48*(5), 1030–1052. https://doi.org/10.1007/s11747-019-00685-3

Sharma, A. N. (2022). Youth employment and unemployment in India: issues and challenges. *Indian Journal of Labour Economics*, *65*(2), 237–267. https://doi.org/10.1007/s41027-022-00387-5

Sharma, S., & Soni, H. K. (2021). An unemployment prediction rate for Indian Youth through time series Forecasting. In *Algorithms for intelligent systems* (pp. 315–335). https://doi.org/10.1007/978-981-33-4893-6_29

Sheth, J. N. (2011). Impact of emerging markets on marketing: Rethinking existing perspectives and practices. *Journal of marketing*, 75(4), 166-182. https://doi.org/10.1509/jmkg.75.4.166

Simionescu, M., & Cifuentes-Faura, J. (2022). Forecasting national and regional youth unemployment in Spain using Google Trends. *Social Indicators Research*, *164*(3), 1187–1216. https://doi.org/10.1007/s11205-022-02984-9

Sitompul, N. K., & Athoillah, M. (2023). Indonesia's Youth Unemployment Rate: Evidence from Sakernas Data. *Journal of International Conference Proceedings*, *6*(1), 347–357. https://doi.org/10.32535/jicp.v6i1.2346

*Statistics on youth - ILOSTAT*. (2025, April 29). ILOSTAT. https://ilostat.ilo.org/topics/youth/

Syafwan, H., Putri, P., & Syafwan, M. (2023). FORECASTING UNEMPLOYMENT IN INDONESIA USING WEIGHTED MOVING AVERAGE METHOD. *JURTEKSI (Jurnal Teknologi Dan Sistem Informasi)*, 9(4), 699–706. https://doi.org/10.33330/jurteksi.v9i4.2624

Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al*. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*(1). https://doi.org/10.1038/sdata.2016.18

World Bank. (2023). *The World Bank in Indonesia: Overview*. https://www.worldbank.org/en/country/indonesia/overview

Yanindah, A. (2022). An insight into Youth Unemployment in Indonesia. *Proceedings of the International Conference on Data Science and Official Statistics*, *2021*(1), 666–682. https://doi.org/10.34123/icdsos.v2021i1.229

# Appendix

## Appendix A. Model Abbreviations

| | Model |
|---|---|
| ada_cds_dt | AdaBoost w/ Cond. Deseasonalise & Detrending |
| arima | ARIMA |
| br_cds_dt | Bayesian Ridge w/ Cond. Deseasonalise & Detrending |
| croston | Croston |
| dt_cds_dt | Decision Tree w/ Cond. Deseasonalise & Detrending |
| en_cds_dt | Elastic Net w/ Cond. Deseasonalise & Detrending |
| et_cds_dt | Extra Trees w/ Cond. Deseasonalise & Detrending |
| et_cds_dt | Extra Trees w/ Cond. Deseasonalise & Detrending |
| ets | ETS |
| exp_smooth | Exponential Smoothing |
| gbr_cds_dt | Gradient Boosting w/ Cond. Deseasonalise & Detrending |
| grand_means | Grand Means Forecaster |
| huber_cds_dt | Huber w/ Cond. Deseasonalise & Detrending |
| knn_cds_dt | K Neighbors w/ Cond. Deseasonalise & Detrending |
| lasso_cds_dt | Lasso w/ Cond. Deseasonalise & Detrending |
| lightgbm_cds_dt | Light Gradient Boosting w/ Cond. Deseasonalise & Detrending |
| llar_cds_dt | Lasso Least Angular Regressor w/Cond. Deseasonalise & Detrending |
| lr_cds_dt | Linear w/ Cond. Deseasonalise & Detrending |
| naive | Naive Forecaster |
| omp_cds_dt | Orthogonal Matching Pursuit w/ Cond. Deseasonalise & Detrending |
| polytrend | Polynomial Trend Forecaster |
| rf_cds_dt | Random Forest w/ Cond. Deseasonalise & Detrending |
| ridge_cds_dt | Ridge w/ Cond. Deseasonalise & Detrending |
| theta | Theta Forecaster |

**Table A1** PyCaret forecasting model abbreviations

# Appendix B. Fine-Tuned Model Performances

***Baseline setup***

gbr_cds_dt

| | cutoff | MASE | RMSSE | MAE | RMSE | MAPE | SMAPE | R2 |
|---|---|---|---|---|---|---|---|---|
| 0 | 2012Q3 | 0.9606 | 1.1173 | 1.4358 | 1.8271 | 0.0943 | 0.0979 | -0.5596 |
| 1 | 2015Q3 | 0.4294 | 0.5628 | 0.5238 | 0.7833 | 0.0379 | 0.0395 | 0.431 |
| 2 | 2018Q3 | 1.1089 | 1.1851 | 1.3164 | 1.6096 | 0.0907 | 0.0962 | -0.3481 |
| Mean | NaT | 0.833 | 0.9551 | 1.092 | 1.4067 | 0.0743 | 0.0779 | -0.1589 |
| SD | NaT | 0.2917 | 0.2787 | 0.4047 | 0.4496 | 0.0258 | 0.0271 | 0.426 |
| holdout set | | 0.3967 | 0.448 | 0.4574 | 0.6011 | 0.0357 | 0.0346 | 0.5764 |

tuned gbr_cds_dt

| | cutoff | MASE | RMSSE | MAE | RMSE | MAPE | SMAPE | R2 |
|---|---|---|---|---|---|---|---|---|
| 0 | 2012Q3 | 1.2821 | 1.3323 | 1.9164 | 2.1787 | 0.1261 | 0.1335 | -1.2175 |
| 1 | 2015Q3 | 0.7662 | 0.7894 | 0.9347 | 1.0987 | 0.0698 | 0.07 | -0.1193 |
| 2 | 2018Q3 | 1.4889 | 1.653 | 1.7675 | 2.2449 | 0.1197 | 0.1315 | -1.6226 |
| Mean | NaT | 1.1791 | 1.2582 | 1.5395 | 1.8408 | 0.1052 | 0.1116 | -0.9865 |
| SD | NaT | 0.3039 | 0.3564 | 0.432 | 0.5254 | 0.0252 | 0.0295 | 0.6351 |
| holdout set | | 0.3967 | 0.448 | 0.4574 | 0.6011 | 0.0357 | 0.0346 | 0.5764 |

ada_cds_dt

| | cutoff | MASE | RMSSE | MAE | RMSE | MAPE | SMAPE | R2 |
|---|---|---|---|---|---|---|---|---|
| 0 | 2012Q3 | 0.9629 | 1.1207 | 1.4392 | 1.8327 | 0.0945 | 0.0981 | -0.5691 |
| 1 | 2015Q3 | 0.3973 | 0.4967 | 0.4847 | 0.6913 | 0.0352 | 0.0365 | 0.5569 |
| 2 | 2018Q3 | 1.1625 | 1.333 | 1.3801 | 1.8104 | 0.0939 | 0.102 | -0.7055 |
| Mean | NaT | 0.8409 | 0.9835 | 1.1013 | 1.4448 | 0.0746 | 0.0789 | -0.2393 |
| SD | NaT | 0.3241 | 0.3549 | 0.4367 | 0.5329 | 0.0278 | 0.03 | 0.5657 |
| holdout set | | 0.5393 | 0.5628 | 0.6218 | 0.7551 | 0.0478 | 0.0462 | 0.3314 |

tuned ada_cds_dt

| | cutoff | MASE | RMSSE | MAE | RMSE | MAPE | SMAPE | R2 |
|---|---|---|---|---|---|---|---|---|
| 0 | 2012Q3 | 0.9606 | 1.1173 | 1.4358 | 1.8271 | 0.0943 | 0.0979 | -0.5596 |
| 1 | 2015Q3 | 0.8596 | 0.9228 | 1.0485 | 1.2843 | 0.0755 | 0.0777 | -0.5293 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 2 | 2018Q3 | 1.1641 | 1.2949 | 1.3819 | 1.7587 | 0.0943 | 0.1017 | -0.6096 |
| Mean | NaT | 0.9947 | 1.1117 | 1.2887 | 1.6234 | 0.088 | 0.0924 | -0.5662 |
| SD | NaT | 0.1266 | 0.152 | 0.1713 | 0.2414 | 0.0089 | 0.0105 | 0.0331 |
| holdout set | | 0.5393 | 0.5628 | 0.6218 | 0.7551 | 0.0478 | 0.0462 | 0.3314 |

rf_cds_dt

| | cutoff | MASE | RMSSE | MAE | RMSE | MAPE | SMAPE | R2 |
|---|---|---|---|---|---|---|---|---|
| 0 | 2012Q3 | 1.117 | 1.2043 | 1.6695 | 1.9694 | 0.1099 | 0.1148 | -0.8119 |
| 1 | 2015Q3 | 0.6423 | 0.7194 | 0.7835 | 1.0013 | 0.0569 | 0.0597 | 0.0704 |
| 2 | 2018Q3 | 1.0716 | 1.1406 | 1.2721 | 1.5492 | 0.0873 | 0.0925 | -0.2489 |
| Mean | NaT | 0.9436 | 1.0214 | 1.2417 | 1.5066 | 0.0847 | 0.089 | -0.3301 |
| SD | NaT | 0.2139 | 0.2151 | 0.3624 | 0.3964 | 0.0217 | 0.0226 | 0.3648 |
| holdout set | | 0.3862 | 0.4621 | 0.4453 | 0.6199 | 0.0348 | 0.0337 | 0.5494 |

tuned rf_cds_dt

| | cutoff | MASE | RMSSE | MAE | RMSE | MAPE | SMAPE | R2 |
|---|---|---|---|---|---|---|---|---|
| 0 | 2012Q3 | 1.0829 | 1.2001 | 1.6186 | 1.9625 | 0.107 | 0.1113 | -0.7992 |
| 1 | 2015Q3 | 0.8367 | 0.8105 | 1.0206 | 1.128 | 0.0737 | 0.0759 | -0.1799 |
| 2 | 2018Q3 | 1.5176 | 1.6512 | 1.8015 | 2.2425 | 0.1228 | 0.1353 | -1.617 |
| Mean | NaT | 1.1457 | 1.2206 | 1.4802 | 1.7777 | 0.1011 | 0.1075 | -0.8654 |
| SD | NaT | 0.2815 | 0.3435 | 0.3335 | 0.4734 | 0.0205 | 0.0244 | 0.5886 |
| holdout set | | 0.3862 | 0.4621 | 0.4453 | 0.6199 | 0.0348 | 0.0337 | 0.5494 |

exp_smooth

| | cutoff | MASE | RMSSE | MAE | RMSE | MAPE | SMAPE | R2 |
|---|---|---|---|---|---|---|---|---|
| 0 | 2012Q3 | 1.2239 | 1.2739 | 1.8294 | 2.0832 | 0.1197 | 0.1277 | -1.0274 |
| 1 | 2015Q3 | 0.5009 | 0.5518 | 0.611 | 0.768 | 0.0453 | 0.0457 | 0.453 |
| 2 | 2018Q3 | 1.4212 | 1.5684 | 1.6872 | 2.1301 | 0.1145 | 0.1257 | -1.3611 |
| Mean | NaT | 1.0487 | 1.1314 | 1.3759 | 1.6604 | 0.0932 | 0.0997 | -0.6452 |
| SD | NaT | 0.3956 | 0.4271 | 0.5439 | 0.6313 | 0.0339 | 0.0382 | 0.7884 |
| holdout set | | 0.4812 | 0.5389 | 0.5549 | 0.723 | 0.0426 | 0.0411 | 0.387 |

tuned exp_smooth

| | cutoff | MASE | RMSSE | MAE | RMSE | MAPE | SMAPE | R2 |
|---|---|---|---|---|---|---|---|---|
| 0 | 2012Q3 | 0.7614 | 0.7573 | 1.1381 | 1.2383 | 0.0778 | 0.0778 | 0.2836 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 1 | 2015Q3 | 0.6842 | 0.7256 | 0.8346 | 1.0099 | 0.0648 | 0.0617 | 0.0543 |
| 2 | 2018Q3 | 0.8578 | 0.8789 | 1.0183 | 1.1936 | 0.0708 | 0.0735 | 0.2586 |
| Mean | NaT | 0.7678 | 0.7872 | 0.997 | 1.1473 | 0.0711 | 0.071 | 0.1988 |
| SD | NaT | 0.071 | 0.0661 | 0.1248 | 0.0989 | 0.0053 | 0.0068 | 0.1027 |
| holdout set | | 0.4429 | 0.5083 | 0.5107 | 0.6819 | 0.0393 | 0.0381 | 0.4547 |

ets

| | **cutoff** | **MASE** | **RMSSE** | **MAE** | **RMSE** | **MAPE** | **SMAPE** | **R2** |
|---|---|---|---|---|---|---|---|---|
| 0 | 2012Q3 | 1.224 | 1.2739 | 1.8294 | 2.0833 | 0.1197 | 0.1277 | -1.0275 |
| 1 | 2015Q3 | 0.5009 | 0.5518 | 0.611 | 0.768 | 0.0453 | 0.0457 | 0.453 |
| 2 | 2018Q3 | 1.4212 | 1.5684 | 1.6872 | 2.1301 | 0.1145 | 0.1257 | -1.3612 |
| Mean | NaT | 1.0487 | 1.1314 | 1.3759 | 1.6605 | 0.0932 | 0.0997 | -0.6452 |
| SD | NaT | 0.3956 | 0.4271 | 0.5439 | 0.6313 | 0.0339 | 0.0382 | 0.7884 |
| holdout set | | 0.4546 | 0.5151 | 0.5241 | 0.691 | 0.0405 | 0.0391 | 0.4401 |

tuned ets

| | **cutoff** | **MASE** | **RMSSE** | **MAE** | **RMSE** | **MAPE** | **SMAPE** | **R2** |
|---|---|---|---|---|---|---|---|---|
| 0 | 2012Q3 | 1.121 | 1.1409 | 1.6756 | 1.8657 | 0.1105 | 0.1161 | -0.6261 |
| 1 | 2015Q3 | 0.4383 | 0.4307 | 0.5346 | 0.5994 | 0.04 | 0.04 | 0.6669 |
| 2 | 2018Q3 | 1.2671 | 1.4233 | 1.5042 | 1.933 | 0.1024 | 0.1117 | -0.9445 |
| Mean | NaT | 0.9421 | 0.9983 | 1.2381 | 1.466 | 0.0843 | 0.0893 | -0.3012 |
| SD | NaT | 0.3612 | 0.4176 | 0.5024 | 0.6134 | 0.0315 | 0.0349 | 0.6968 |
| holdout set | | 0.5784 | 0.6263 | 0.6669 | 0.8401 | 0.0518 | 0.0497 | 0.1723 |

***Feature engineered setup***

et_cds_dt

| | **cutoff** | **MASE** | **RMSSE** | **MAE** | **RMSE** | **MAPE** | **SMAPE** | **R2** |
|---|---|---|---|---|---|---|---|---|
| 0 | 2012Q3 | 1.1098 | 1.1586 | 1.6967 | 1.9441 | 0.1143 | 0.117 | -0.7658 |
| 1 | 2015Q3 | 0.2785 | 0.263 | 0.3232 | 0.3529 | 0.0234 | 0.0234 | 0.8845 |
| 2 | 2018Q3 | 1.2294 | 1.38 | 1.4051 | 1.8202 | 0.0964 | 0.1048 | -0.7242 |
| Mean | NaT | 0.8726 | 0.9339 | 1.1416 | 1.3724 | 0.078 | 0.0817 | -0.2018 |
| SD | NaT | 0.4229 | 0.4829 | 0.5909 | 0.7227 | 0.0393 | 0.0415 | 0.7683 |
| holdout set | | 0.7282 | 0.6838 | 0.8125 | 0.8958 | 0.0617 | 0.0642 | 0.059 |

tuned et_cds_dt

| | cutoff | MASE | RMSSE | MAE | RMSE | MAPE | SMAPE | R2 |
|---|---|---|---|---|---|---|---|---|
| 0 | 2012Q3 | 1.1912 | 1.2073 | 1.8211 | 2.0259 | 0.1218 | 0.126 | -0.9174 |
| 1 | 2015Q3 | 0.3432 | 0.3808 | 0.3983 | 0.511 | 0.0305 | 0.0298 | 0.7579 |
| 2 | 2018Q3 | 1.1986 | 1.3362 | 1.3698 | 1.7625 | 0.0939 | 0.1017 | -0.6165 |
| Mean | NaT | 0.911 | 0.9748 | 1.1964 | 1.4331 | 0.0821 | 0.0859 | -0.2587 |
| SD | NaT | 0.4015 | 0.4233 | 0.5937 | 0.6608 | 0.0382 | 0.0409 | 0.7292 |
| holdout set | | 0.7282 | 0.6838 | 0.8125 | 0.8958 | 0.0617 | 0.0642 | 0.059 |

omp_cds_dt

| | cutoff | MASE | RMSSE | MAE | RMSE | MAPE | SMAPE | R2 |
|---|---|---|---|---|---|---|---|---|
| 0 | 2012Q3 | 1.0984 | 1.2136 | 1.6792 | 2.0366 | 0.1105 | 0.1149 | -0.9377 |
| 1 | 2015Q3 | 0.4493 | 0.5113 | 0.5214 | 0.6861 | 0.0409 | 0.0394 | 0.5635 |
| 2 | 2018Q3 | 1.1039 | 1.2411 | 1.2616 | 1.6371 | 0.0859 | 0.0924 | -0.3946 |
| Mean | NaT | 0.8839 | 0.9887 | 1.1541 | 1.4532 | 0.0791 | 0.0822 | -0.2562 |
| SD | NaT | 0.3073 | 0.3377 | 0.4788 | 0.5664 | 0.0288 | 0.0316 | 0.6206 |
| holdout set | | 0.5071 | 0.4498 | 0.5658 | 0.5893 | 0.0424 | 0.0431 | 0.5928 |

tuned omp_cds_dt

| | cutoff | MASE | RMSSE | MAE | RMSE | MAPE | SMAPE | R2 |
|---|---|---|---|---|---|---|---|---|
| 0 | 2012Q3 | 1.0249 | 1.1445 | 1.5668 | 1.9205 | 0.1029 | 0.1071 | -0.7231 |
| 1 | 2015Q3 | 0.4415 | 0.4903 | 0.5123 | 0.6579 | 0.0399 | 0.0386 | 0.5987 |
| 2 | 2018Q3 | 1.1038 | 1.2411 | 1.2616 | 1.637 | 0.0859 | 0.0924 | -0.3946 |
| Mean | NaT | 0.8567 | 0.9586 | 1.1135 | 1.4051 | 0.0762 | 0.0794 | -0.173 |
| SD | NaT | 0.2954 | 0.3335 | 0.4431 | 0.5409 | 0.0266 | 0.0294 | 0.5619 |
| holdout set | | 0.5982 | 0.5652 | 0.6675 | 0.7404 | 0.052 | 0.0505 | 0.3571 |

ridge_cds_dt

| | cutoff | MASE | RMSSE | MAE | RMSE | MAPE | SMAPE | R2 |
|---|---|---|---|---|---|---|---|---|
| 0 | 2012Q3 | 1.1984 | 1.266 | 1.8321 | 2.1245 | 0.1266 | 0.1256 | -1.1086 |
| 1 | 2015Q3 | 0.4744 | 0.5133 | 0.5505 | 0.6887 | 0.042 | 0.0407 | 0.5602 |
| 2 | 2018Q3 | 1.0106 | 1.1621 | 1.155 | 1.5328 | 0.0785 | 0.0843 | -0.2227 |
| Mean | NaT | 0.8945 | 0.9805 | 1.1792 | 1.4487 | 0.0824 | 0.0835 | -0.257 |
| SD | NaT | 0.3068 | 0.3331 | 0.5235 | 0.5892 | 0.0346 | 0.0347 | 0.6817 |
| holdout set | | 0.6053 | 0.5348 | 0.6754 | 0.7007 | 0.0507 | 0.0517 | 0.4243 |

tuned ridge_cds_dt

| | cutoff | MASE | RMSSE | MAE | RMSE | MAPE | SMAPE | R2 |
|---|---|---|---|---|---|---|---|---|
| 0 | 2012Q3 | 1.1847 | 1.2264 | 1.8111 | 2.0581 | 0.1217 | 0.1251 | -0.9788 |
| 1 | 2015Q3 | 0.4332 | 0.4939 | 0.5026 | 0.6627 | 0.039 | 0.0377 | 0.5928 |
| 2 | 2018Q3 | 1.1051 | 1.2734 | 1.263 | 1.6797 | 0.0856 | 0.0924 | -0.4682 |
| Mean | NaT | 0.9076 | 0.9979 | 1.1922 | 1.4668 | 0.0821 | 0.0851 | -0.2847 |
| SD | NaT | 0.3371 | 0.3569 | 0.5365 | 0.5892 | 0.0338 | 0.0361 | 0.6546 |
| holdout set | | 0.6053 | 0.5348 | 0.6754 | 0.7007 | 0.0507 | 0.0517 | 0.4243 |

**Table B1.** Full hold-out set model evaluations including fine-tuned model for Baseline setup and Feature-engineered setup

# Appendix C. Hold-Out Set Forecast Visualisations

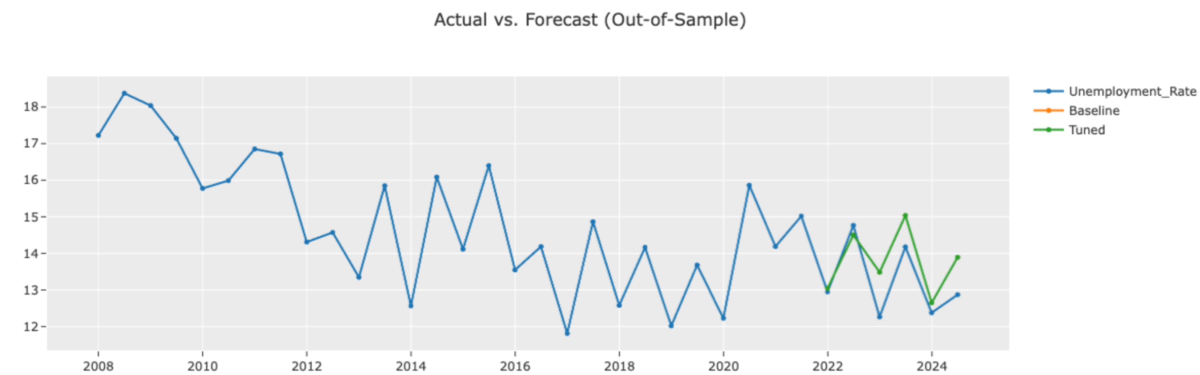


**Figure C1** Hold-out set forecast for AdaBoost with Cond. Deseasonalise & Detrending (Baseline setup)

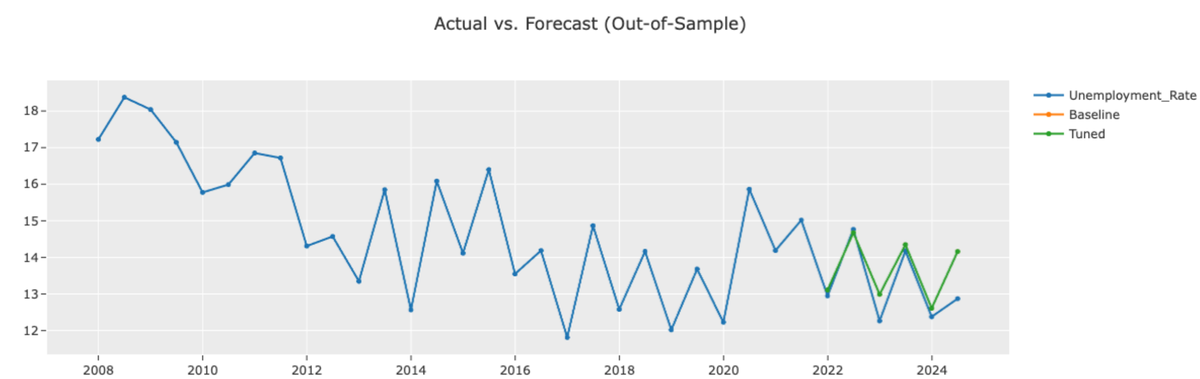


**Figure C2** Hold-out set forecast for Random Forest with Cond. Deseasonalise & Detrending (Baseline setup)

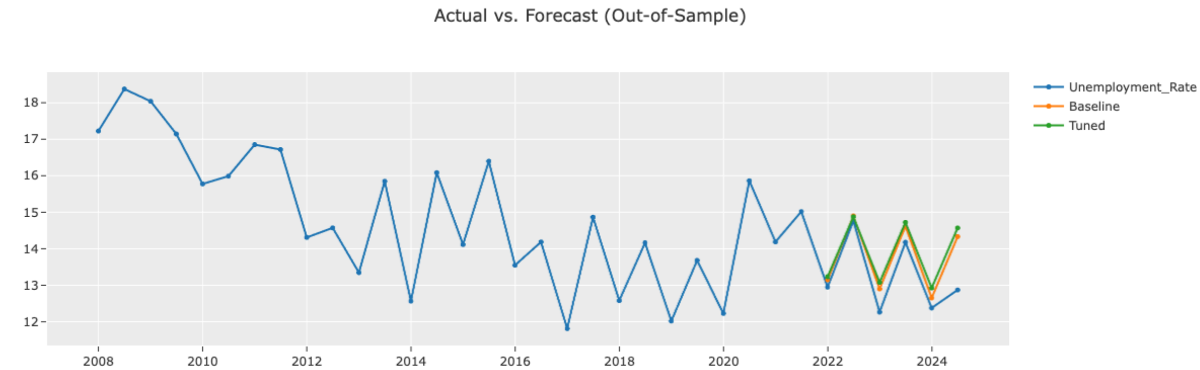


**Figure C3** Hold-out set forecast for ETS (Baseline setup)

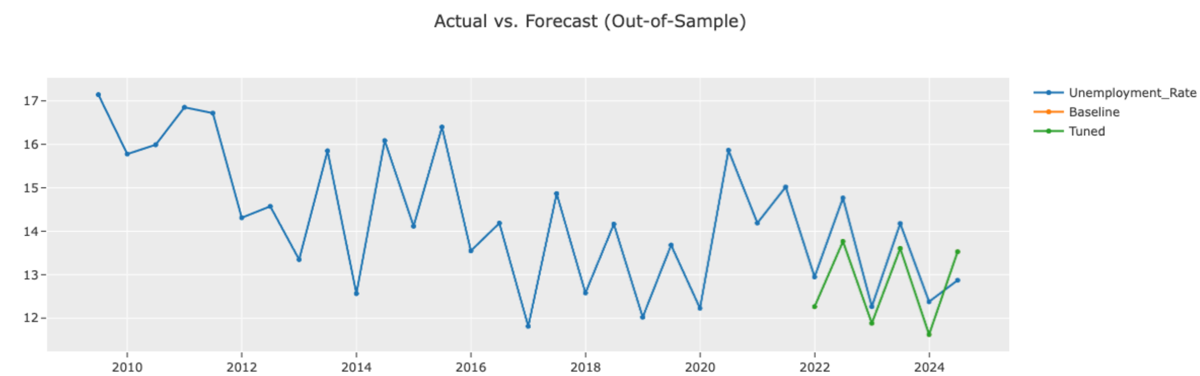


**Figure C4** Hold-out set forecast for Ridge with Cond. Deseasonalise & Detrending (Feature engineered setup)