CHAPTER I
INTRODUCTION

1.1.Background

Unemployment is one of the crucial indicators that reflects the quality of regional development. The high unemployment rate not only shows an imbalance between the number of job seekers and the availability of jobs, but also has a direct impact on social stability, a decline in people's purchasing power, and a slowdown in economic growth[1]. Sidoarjo Regency as one of the largest industrial and trade centers in East Java faces serious challenges in maintaining this balance[2]. Sidoarjo's strategic position as an industrial urban area makes it an area that continues to receive inflows of labor from various regions, while employment growth is not always able to keep pace with the increase in the number of labor force[3].

The development of the unemployment rate in Sidoarjo Regency during 2005 to 2024 shows significant dynamics, with a downward trend interrupted by external shocks such as the COVID-19 pandemic in 2020. Based on BPS data in 2024, the Open Unemployment Rate (TPT) of Sidoarjo Regency was recorded at 6.49%. However, these TPT fluctuations do not occur in a vacuum; It correlates closely with other macroeconomic indicators. For example, Sidoarjo's economic growth (GDP) shows a pattern that is generally inversely proportional to TPT, where a sharp decline in GDP to -3.69% in 2020 followed by a surge in TPT to 7.99%. Similarly, the realization of investment, which tends to increase in the last decade, also supports labor absorption. This condition confirms that the unemployment problem in Sidoarjo is a multidimensional phenomenon influenced by the complex interaction between labor supply (indicated by the number of working-age population), labor demand (driven by economic growth and investment), and the quality of human resources (as reflected in the Human Development Index)[4].

In the context of data-driven planning, the ability to predict TPT should not rely only on its own historical patterns (univariate), but must utilize information from supporting variables that have a strong causal or correlational relationship. Multivariate approach *time series forecasting* become a must. However, regional economic data such as Sidoarjo is often available on an annual frequency with a limited number of observations (in this study 20 data points). These limitations demand the

selection of modeling methods that are not only sophisticated but also efficient and robust against risk *overfitting*. This is where the combination of models comes in. *deep learning* which is capable of capturing complex non-linear relationships and intelligent hyperparameter optimization techniques is crucial [5].

This study specifically uses the Open Unemployment Rate (TPT) in percentage (%) as the main variable. The selection of TPT was carried out because this ratio is more representative in describing employment conditions, more stable to demographic changes, and is most suitable for time series modeling. The use of TPT allows for more precise trend analysis so that the predictions can be scientifically accounted for and relevant to be used as a basis for policy.

Along with the development of artificial intelligence technology, modern forecasting methods are based on *deep learning* is starting to be widely used due to its ability to capture non-linear relationships and complex time series patterns[6]. One of the widely used models is *Gated Recurrent Unit (GRU)*, a derivative architecture of *Recurrent Neural Network (RNN)* designed to study long-term dependency on time series data. GRU has a simpler and more efficient structure than LSTM, but is still able to produce competitive prediction performance, especially on small to medium-scale datasets such as annual TPT data[7].

In order to optimize the forecasting results, the GRU model in this study is combined with Bayesian Hyperparameter Optimization, a modern optimization approach that can find the best combination of parameters systematically, efficiently, and probabilically-based[8]. Bayesian optimization helps improve model accuracy by minimizing risk *overfitting*, speed up the parameter search process, and produce a more stable model than manual parameter selection or conventional approaches such as *grid search*[9].

By utilizing datasets enriched by key macroeconomic variables (GDP Growth, Investment, Working-Age Population, and HDI), this study aims to build a Bayesian-optimized Multivariate GRU forecasting model. The resulting predictions are expected to be not only more accurate but also more interpretive, as they can provide an idea of how changes in these determinants may affect future TPT projections. Thus, this research contributes to providing a more comprehensive and evidence-based decision support tool for the Sidoarjo Regency government in formulating effective and anticipatory employment policies.

1.2.Problem Formulation

Based on this background, several research problems that will be answered in the thesis can be formulated, including:

1. How to design and implement a Multivariate *Gated Recurrent Unit (GRU)* model that integrates TPT historical data with supporting variables (GDP Growth, Investment, Working Age Population, and HDI) for TPT forecasting in Sidoarjo Regency?
2. How can the Bayesian Hyperparameter optimization process improve the performance of Multivariate GRU models?

1.3.Research Objectives

The purpose of the thesis with the title "Forecasting the Unemployment Rate in Sidoarjo Regency Using the Model *Gated Recurrent Unit (GRU)* with Bayesian Hyperparameter Optimization" are as follows:

1. Designing and implementing a Multivariate GRU model that integrates TPT data and related macroeconomic variables for the period 2005–2024.
2. Evaluate and compare the performance of the optimized Multivariate GRU model using RMSE metrics.

1.4.Research Benefits

Some of the benefits that can be taken from the thesis entitled "Forecasting the Unemployment Rate in Sidoarjo Regency Using the Model *Gated Recurrent Unit (GRU)* with Bayesian Hyperparameter Optimization" including:

1. Provide an understanding of the application of *the Gated Recurrent Unit (GRU)* method in forecasting the unemployment rate. This research can provide a clear picture of the technical process of developing the GRU model and how the *deep learning model* is able to study the trend patterns and dynamics of TPT data in Sidoarjo Regency in a non-linear manner.
2. It is a reference in the selection of effective and efficient forecasting methods. The results of this study can be a reference for researchers and practitioners in comparing the advantages of GRU models optimized with Bayesian Hyperparameters against other forecasting methods, especially in the context of employment time sequence data.
3. Encourage the development of deep learning-based predictive analytics systems at the regional level. This research has the potential to be the first step

in the development of an automated forecasting system based on the GRU model to support the government's decision-making process more quickly, accurately, and sustainably.

By understanding the advantages and limitations of the GRU method, policymakers or system developers can design prediction solutions that are more adaptive, accurate, and relevant to changes in employment conditions in Sidoarjo Regency.

1.5.Problem Limitations

Limitations of the problem in the thesis research with the title "Forecasting the Unemployment Rate in Sidoarjo Regency Using the Model *Gated Recurrent Unit (GRU)* with Bayesian Hyperparameter Optimization" are as follows:

1. This research focuses specifically on Sidoarjo Regency and analyzes data in the period of 2005 to 2024. The resulting findings and models cannot necessarily be generalized to other regions or periods outside of that range without further study.
2. This study limited the predictor variables on the Open Unemployment Rate (TPT) as the target variable, and the four main explanatory variables, namely GDP growth, investment realization, the number of working-age population, and the Human Development Index (HDI). Other macroeconomic or social variables that may have an influence (such as inflation, interest rates, or specific fiscal policies) were not included in the modeling.
3. All data used are secondary data sourced from the official publication of the Central Statistics Agency (BPS) of Sidoarjo Regency, so that it follows the definition, methodology, and data quality applicable in the institution. The study did not use primary survey data or estimation data from non-official sources.
4. This study focuses on aspects of multivariate predictive modeling and hyperparameter optimization for short-term forecasting. Although using economic variables, the main goal is to produce accurate predictions, not to conduct causal analysis or test economic theories in depth regarding the factors that cause unemployment.

This page was deliberately blanked