CHAPTER I
INTRODUCTION

1.1 Background

In the contemporary era, investment activities have gained significant interest among various segments of society, including both the general public and students. This development is bolstered by rising financial literacy and simplified access to various investment instruments such as stocks, bonds, real estate, and precious metals. A comprehensive understanding of investment is a fundamental requirement for individuals before entering the capital market, as it encompasses the ability to evaluate investment feasibility, analyze risks, and estimate expected rates of return. Sound investment knowledge, acquired through formal education or outreach by capital market institutions, has been shown to positively influence public interest and participation in investing [1]. The rise in investment interest is further driven by technological advancements, specifically the emergence of various digital investment platforms that facilitate transactions within the capital market. According to data from the Kustodian Sentral Efek Indonesia (KSEI), the number of stock investors in 2022 reached 10,311,152, representing a 37.68% increase from the previous year's figure of 7,489,337 [2]. One of the most prominent investment instruments in the Indonesian capital market is the LQ45 Index, which consists of 45 stocks characterized by high liquidity and large market capitalization. This index is regarded as a primary indicator reflecting national stock market performance [3]. Among the various sectors within this index, the banking sector holds a particular appeal due to its critical role in mobilizing and allocating public funds, while simultaneously serving as a barometer for national economic stability [4]. Nevertheless, stock prices in the banking sector remain highly volatile, influenced by diverse economic factors, political policies, interest rate fluctuations, and investor sentiment.

In recent years, advancements in artificial intelligence technology, particularly machine learning and deep learning, have significantly contributed to the field of financial data analysis. One of its primary applications is the prediction of stock prices, which are characterized as time-series data information that evolves over time. The Recurrent Neural Network (RNN) model is an artificial neural network architecture specifically designed to handle sequential data due to its ability to retain information from previous inputs. However, conventional RNNs are limited by the vanishing gradient problem, which makes it difficult for the model to learn long-term dependencies within historical data [5].

To address these issues, two primary variants of RNN were developed, namely Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU). The LSTM model possesses the capability to retain long-term information through a complex memory management mechanism, making it superior in capturing long-term fluctuating patterns. Conversely, GRU offers a simpler architecture while remaining reliable in capturing short-term fluctuations. With the growing demand for higher predictive accuracy, approaches integrating both architectures into a hybrid LSTM-GRU model have emerged. This integrated architecture is designed to optimize the respective strengths of LSTM and GRU in analyzing sequential financial data. The resulting hybrid model demonstrates more robust predictive performance by integrating LSTM's capacity for understanding long-term dependencies with the training efficiency inherent in GRU [6]. This combination produces a more resilient model for identifying dynamic and recurring stock price patterns. Furthermore, several previous studies indicate that hybrid LSTM-GRU models can provide more accurate prediction results than standalone models by balancing long-term memory capacity with training efficiency. Nevertheless, the optimal performance of this model is highly dependent on the hyperparameter configurations employed. Inappropriate hyperparameter selection can lead to model instability or issues such as overfitting and underfitting [7].

In the context of deep learning, hyperparameters play a vital role in determining how effectively a model can learn patterns from the provided data. Unlike parameters acquired during the training process, hyperparameters must be defined before training commences, and each combination of values can significantly yield different model performances [9]. The process of determining appropriate hyperparameters is often conducted using conventional methods such as Grid Search or Random Search. However, Grid Search possesses a computational complexity that makes it impractical for moderate values and remains highly sensitive to definition. Meanwhile, Random Search has a limitation where entire regions within the search space may be under-explored due to its randomized nature [10]. To overcome these challenges, a more adaptive and intelligent optimization method is needed to explore the search space and find the optimal hyperparameter combination. Metaheuristic-based optimization methods have become widely adopted alternatives because they can balance the exploration and exploitation of the solution space without necessitating an exhaustive search [11]. One metaheuristic method that has proven effective is the Genetic Algorithm (GA). The strength of GA lies in its ability to explore the solution space extensively and identify optimal values without depending on gradients or mathematical derivatives [12]. Previous research has successfully utilized GA to optimize hyperparameters in RNN models. Consequently, the application of GA to a hybrid LSTM-GRU model is expected to improve the accuracy and stability of prediction results by identifying the optimal parameter configurations.

Based on the aforementioned description, this research focuses on the application of a hybrid LSTM-GRU model to predict the stock prices of the banking sector listed in the LQ45 Index. This model was selected due to its capability to combine the strengths of two RNN architectures that are effective in capturing patterns within time-series data. To achieve optimal predictive performance, this study integrates the Genetic Algorithm (GA) as a hyperparameter optimization method, which is expected to produce a model that is more accurate and efficient than those without optimization. The model's performance will be evaluated using three primary metrics, namely Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE), which are utilized to measure prediction error rates from various perspectives.

1.2 Problem Formulation

Based on the background previously described, the problems addressed in this research are as follows:

1. How to develop a hybris LSTM-GRU model for predicting the stock prices of banking sector companies listed in the LQ45 Index?
2. How is the Genetic Algorithm applied to optimize the hyperparameters of the hybrid LSTM-GRU model?

3. How does the performance of the hybrid LSTM-GRU model optimized using the Genetic Algorithm compare to the non-optimized hybrid model based on evaluation metrics?

1.3 Research Objectives

The objectives of this research are as follows:

1. To develop a hybrid LSTM-GRU model for predicting the stock prices of banking sector companies listed in the LQ45 Index.
2. To implement the Genetic Algorithm for the hyperparameter optimization of the hybrid LSTM-GRU model.
3. To compare the performance of the hybrid model optimized using the Genetic Algorithm with the non-optimized hybrid model.

1.4 Research Benefits

The expected benefits of this research are as follows:

1. To contribute to the development of stock price prediction methods for banking sector companies listed in the LQ45 Index, optimized through an optimization approach.
2. To demonstrate the application of combining deep learning methods with evolution-based optimization within the field of investment.
3. To provide insights for investors, academics, and financial practitioners regarding the effectiveness of utilizing the hybrid LSTM-GRU model.

1.5 Scope and Limitations

This research includes several limitations as follows:

1. The data used is sourced from Yahoo Finance for historical daily stock price data of banking companies listed in the LQ45 Index and the Otoritas Jasa Keuangan (OJK) for corporate internal factor data. The timeframe of the data spans from November 2020 to June 2025.
2. The analyzed stocks are limited to the banking sector companies listed in the LQ45 Index, specifically BBCA, BBNI, BBRI, BBTN, BMRI, and BRIS.
3. The analyzed internal corporate factors are limited to two financial ratios, namely Return on Assets (ROA) and Return on Equity (ROE), which represent the companies performance.