CHAPTER I
INTRODUCTION

1.1 Background

Investing in the stock market has become a popular method for individuals and institutions to grow wealth and achieve long-term financial stability. Stocks offer the potential for significant returns through capital appreciation and dividends. However, the stock market is notoriously volatile and unpredictable, making investment decisions challenging. Stock forecasting plays a crucial role in financial decision-making, as it provides investors, traders, and other market participants with valuable insights into the likely direction and magnitude of future price changes [1]. By using stock forecasting techniques, market participants can make more informed decisions about when to buy or sell stocks, which can help to maximize returns and minimize risks.

The need for accurate stock forecasting is particularly important in today's fast-moving and highly competitive financial markets [2]. However, stock forecasting is a challenging task that requires sophisticated statistical and machine learning techniques. Financial markets are complex and highly unpredictable, and many different factors can influence stock prices, including economic indicators, geopolitical events [3], [4] and changes in investor sentiment [5]–[8]. As a result, stock forecasting requires the use of advanced mathematical models and algorithms that can capture the complex relationships between these various factors and predict future price movements with a high degree of accuracy.

Machine learning methods, especially deep learning, have been developed to predict stock prices. The application of the Recurrent Neural Network (RNN) model to process time-series data such as the Long-Short Term Memory (LSTM) is used to predict the close price index of the S&P 500 stock the next day. It was found that single layer LSTM has RMSE value of 42.7093 in 100 epochs while Multi-Layer LSTM has RMSE value of 53.9076 in 100 epochs [9]. Other studies have also compared the LSTM model and Bidirectional Long-Short Term Memory (BiLSTM) to predict stock prices, where with the same parameters, BiLSTM produces a lower RMSE value compared to LSTM [10]. In other research BiLSTM achieve MAE value of 21.952 and RMSE value of 31.694 in predicting Shanghai Composite Index Stock [11]. GRU is also used to predict the Shanghai Composite Index and get the MSE value of 20.3225 and RMSE value of 28.839 [12]. There is also research that implements a combination of Gated Recurrent Unit (GRU) models with four blocks of neural network architecture (NN) to identify patterns of joint movement in the stock market and it produce MAPE Value of 5.61% [13].

Recent advancements in stock forecasting have seen the exploration of various sophisticated models. Transformers, originally designed for natural language processing, have shown promising results in time series forecasting due to their attention mechanism, which allows the model to focus on relevant parts of the input sequence, thus capturing long-range dependencies more effectively than traditional RNNs. Nadeem Malibari et al. utilized transformer to predict Saudi Stock Exchange and get RMSE result of MSE 0.0013, RMSE value of 0.763 and MAPE value of 1.681% [14]. On the other hand, Bidirectional Gated Recurrent Unit (BiGRU) models have been employed to capture information from both past and future data points, enhancing the model's understanding of temporal dependencies also offers a more comprehensive view of the sequential data [15].

In recent years, combining the strengths of various neural network architectures aim to harness the complementary features of different deep learning paradigms. The surge in popularity of these hybrid models can be attributed to their ability to mitigate the limitations of individual architectures, offering a more robust and adaptable framework for forecasting stock prices [16]. Md. Ebtidaul Karim et al. utilized hybrid model combining BiLSTM and GRU for predicting NIFTY-50 stock and get the MAPE value of 117.079% [17]. This thesis aims to explore and compare existing stock prediction models and develop a hybrid model that combines BiGRU and BiLSTM which will represent an approach to time-series prediction, allowing the model to capture both forward and backward dependencies in sequential data.

This hybrid model has the potential to produce lower error of stock price forecasts by effectively learning intricate patterns and dependencies in historical stock data. The historical stock data that will be used in this thesis are Gold, Apple, Oil and Silver. Gold, widely regarded as a "safe-haven" asset, becomes particularly relevant in times of economic uncertainty, with its price often reflective of broader market sentiment and tendencies towards seeking security. Apple, as a leading technology company, serves as an indicator of trends and innovations in the tech sector, carrying significant implications for the overall stock market. Oil prices, functioning as a global economic indicator, are crucial in reflecting the dynamics of global demand and supply, offering insights into broader economic growth prospects. Meanwhile, silver, not only a precious metal but also widely used in various industries, provides valuable insights into industrial demand and investor sentiment towards precious metals. This diversified selection across sectors aims to create a comprehensive portfolio that captures a broad spectrum of market dynamics, contributing to more effective risk management.

1.2 Problem Identification

The stock market is characterized by its inherent volatility and complexity, posing significant challenges for accurate price prediction. Traditional linear models often fall short in capturing the nonlinear and intricate patterns of stock price movements. Table 1.1 shows the performance of several previous work.

Table 1.1. Previous Work Summary

Model	Dataset	Performance
LSTM [9]	S&P 500	RMSE : 42.7093
BiLSTM [18]	Shanghai Composite Index	MAE: 21.952 RMSE:31.694
GRU [12]	Shanghai Composite Index	RMSE: 28.839 MSE: 20.3225
Transformer [14]	Tadawul (Saudi Stock Exchange)	RMSE: 0.763 MSE: 0.0013 MAPE: 1.681
BiGRU [19]	Britannia Stock	RMSE: 0.016 MSE: 0.00027
BiLSTM-GRU [17]	NIFTY-50	MAPE: 117.079

Despite the development of various predictive models, many fail to adequately address these complexities, leading to suboptimal predictions that can adversely impact investment decisions. Single deep learning models, including RNNs, LSTMs, GRUs, and Transformers, exhibit their own set of limitations. RNNs, for instance, are susceptible to vanishing gradient problems, hindering their ability to capture long-term dependencies effectively. While LSTMs address this issue to some extent, they may still struggle with certain types of temporal dependencies. Similarly, GRUs designed to overcome specific challenges, may not fully capture the complexity of non-linear relationships in stock price data. The Transformer architecture, effective for sequence-to-sequence tasks, demands extensive data and computational resources. These limitations highlight the need for a more advanced and robust forecasting approach to navigate the intricacies of financial markets.

To address these challenges, an ensemble learning approach is proposed for stock price forecasting. The ensemble model under consideration combines the Bidirectional Gated Recurrent Unit (BiGRU) and Bidirectional Long Short-Term Memory (BiLSTM) architectures. The empirical evidence from previous studies, which indicates the effectiveness of combining BiLSTM and GRU in ensemble models, serves as a foundation for this proposed approach [17]. By strategically combining these bidirectional architectures, the thesis aims to evaluate the effectiveness of BiGRU-BiLSTM models in stock price forecasting and enhance predictive capabilities through empirical assessment and comparison with other benchmark model.

1.3 Objectives

This thesis considers following assumptions:

1. Design a hybrid deep learning model of BiGRU-BiLSTM for stock forecasting.
2. Test the hybrid model on different types of stocks, including gold, oil, silver, and Apple, to assess its performance across various market conditions.
3. Evaluate the performance of the BIGRU-BiLSTM model. The evaluation based on various metrics such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE).

4. Conduct empirical comparison between BiGRU-BiLSTM models and other deep learning models: BILSTM, GRU, Transformer, BiGRU and BiLSTM-GRU for stock forecasting to determine the relative performance and advantages of the proposed architecture.

1.4 Scope of Work

1. Implementing BiGRU-BiLSTM model for stock closing price forecasting.
2. The stock used in this thesis are Gold, Silver, Oil and Apple. The feature used from historical stock prices are Open, High, Low, Adj Close and Volume, acquired from yahoo finance.
3. Evaluating the performance of the BiGRU-BiLSTM model on the testing data using various evaluation metrics such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) using Python programming.

1.5 Research Methodology

In this thesis, we use fundamental study and experiment based on work-packages (WP). These are the following WP for this thesis:

WP1: Collect historical stock price data for a selected set of gold, silver, oil and apple stocks from a reliable financial data source, such as Yahoo Finance.

WP2: Clean and pre-process the collected data to remove any missing values, outliers, or other irregularities that could affect the performancy of the model.

WP3: Develop a hybrid BiGRU-BiLSTM model.

WP4: Train the BiGRU-BiLSTM model using the Pre-processed data, tuning the model hyperparameters to optimize its performance.

WP5: Evaluate the performance of the BiGRU-BiLSTM model using various metrics.

WP6: Interpret the results of the experiments and analyse the model performance.