QUANTITATIVE TRADING PORTFOLIO OPTIMIZATION-BASED STOCK PREDICTION USING LONG-SHORT TERM MEMORY NETWORK

by

Ruizhi Hao

A THESIS

Submitted to the Faculty of the Stevens Institute of Technology
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN BUSINESS INTELLIGENCE AND ANALYTICS

____

Ruizhi Hao, Candidate

ADVISORY COMMITTEE

____

Dragos Bozdog, Chairman Date

____

Ionut Florescu, Reader Date

STEVENS INSTITUTE OF TECHNOLOGY
Castle Point on Hudson
Hoboken, NJ 07030
2021

QUANTITATIVE TRADING PORTFOLIO OPTIMIZATION-BASED STOCK PREDICTION USING LONG-SHORT TERM MEMORY NETWORK

ABSTRACT

Stock prediction is critical in quantitative trading for creating an efficient trading strategy that yields a high return. The ability to predict outcomes is also needed for successful portfolio construction and optimization. Stock prediction, on the other hand, is a difficult task due to the numerous factors involved, such as uncertainty and instability. Deep learning techniques, especially the recurrent neural network (RNN), have recently been developed for sequence prediction. A long short-term memory (LSTM) network is proposed in this paper to predict market movement using historical data. Multiple portfolio optimization techniques, such as equal-weighted modeling (EQ) and optimization modeling maximizing Sharpe ratio, are used to optimize portfolio efficiency in order to build an effective portfolio. The results showed that our proposed LSTM prediction model is effective in predicting stock prices with high accuracy. In addition, using maximizing Sharpe ratio method to rebalance the allocation strategy every three month showed a significant improvement in the cumulative return of the constructed portfolios. Furthermore, our constructed portfolios beat the benchmark Sector ETF index in both XLU and XLB.

Author: Ruizhi Hao

Advisor: Dragos Bozdog

Date: May 10, 2021

Degree: Master of Science in Business Intelligence and Analytics

Acknowledgments

First, I would like to express my great gratitude to Professor Bozdog and Professor Florescu. Throughout the study process, you are always trying to assist. Your enthusiasm for science inspires me and makes me more accountable for my own thesis. I would not be able to complete the thesis properly without your encouragement and guidance. Second, I'd like to express my gratitude to the Hanlon Financial Systems Laboratories' faculty. It would be much more difficult to perform the experiment if you do not have the database source for the news papers. Last but not least, I'd like to express my gratitude to all who assisted me with the thesis. It is a good time to have the first experience of research at Stevens Institute of Technology.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

A portfolio is defined as a collection of investment assets. Portfolio management is the method of making investing decisions based on customized tactical investment strategies in order to increase returns over a fixed period of time. The two most popular approaches are traditional and quantitative fund management(1).

Both approaches have several similarities, such as focusing on a small number of key equity-value-driving variables, analyzing historical data to estimate these key drivers, defining stock selection eligibility criteria, and monitoring output over time. Traditional portfolio management, relies heavily on judgment depth analysis, regime transformations, core characteristics, and qualitative considerations, while quantitative portfolio management prioritizes universe exploration, discipline, verification, risk management, and lower fees. It is capable of not only detecting mode openings, but also of assisting in the management of unforeseen risks(3).

Quantitative trading refers to trading strategies that are focused on quantitative investment analysis and depend on mathematical models to construct an automated trading system. Portfolio building, also known as diversification in quantitative investing, is the method of choosing and allocating assets across various stocks to minimize trading risk. The market trend, entry and exit transactions, price background, and volume are all important factors in quantitative trading strategies. The creation of an effective prediction model is the most important step in the quantitative method for building an efficient portfolio. Stock prediction is important for predicting the overall movement of the market or the movement of a single stock. Stock price prediction has been recognized as one of the most challenging problems in the financial industry due to the complexity of multivariate time series properties as well as the amount of financial data involved. In a variety of experiments, statistical and machine learning methods were used to increase prediction accuracy(55). In terms of accuracy and decision support, artificial intelligence (AI) and deep learning algorithms have recently been shown to have a number of advantages over traditional prediction models. Deep learning algorithms allow the development of a variety of trading strategies that can be implemented and adapted to a real-time market on a continuous basis(56).Although the potential of deep learning to predict stocks has been extensively investigated, little attention has been paid to using the stock prediction phase to build successful quantitative portfolios. In this paper, a prediction model for market price prediction is proposed using long short-term memory (LSTM), a special variant of recurrent neural network (RNN), followed by portfolio optimization strategies to leverage the prediction performance. Many quantitative portfolios are built using a strategic asset allocation trading strategy. In contrast to real trading, the prediction model achieves high accuracy in prediction for and trial, and our developed portfolios have a substantial return over many estimated time periods. The designed portfolios outperform the benchmark Level in terms of active return and risk management. The following are the key contributions of this paper:

- The Random Forest, Linear regression XGBoost and LSTM prediction model were proposed to predict stock price in order to choose the best method to construct and optimize portfolios in quantitative trading.

- Presenting a comparison between each method

• In quantitative trading, optimization modeling methods were used to optimize portfolios.

• Finally, portfolio output for the built portfolios was analyzed, and our portfolios outperformed the benchmark in terms of cumulative return.

The remaining part of the paper is structured as follows. Chapter 2 introduces the fundamental principles of quantitative trading and related work. Chapter 3 addresses the proposed Linear regression, XGBoost, Random Forest and LSTM prediction models for stock prediction and portfolio construction and optimization techniques. Chapter 4 outlines the experiment and its consequences. Finally, Chapter 5 summarizes the conclusions and discussions.

Chapter 2

Literature Review

2.1 Fundamentals of Quantitative Trading

The data collection method is the first fundamental component of the system, and data can be obtained from external sources, a data provider, or proprietary analysis. Time-series data and cross-sectional data are the two types of financial data structures(26). The key tasks in getting secure data sources stored in the data warehouse are data cleaning and preprocessing. The modeling method is primarily responsible for designing precise prediction, statistical analysis, and optimization models. Finally, the study' findings are visualized and used as guidelines for making investment decisions. Modeling and analytics are commonly used in an iterative process of analyzing patterns, assessing strategies, backtesting, and evaluating portfolio output.

Quantitative trading is a form of automated trading system in which a series of mathematical models manage the trading strategies and decisions. Quantitative trading is a concept that aims to make high-frequency trading processes more efficient by combining mathematical mathematics, computer algorithms, and computing resources, which aims to minimize risk and maximize return based on the historical performance of the encode strategies tested against historical financial data. Quantitative trading is the latest wave of trading in quantitative portfolio management, offering investors a range of advantages such as efficient execution and lower transaction costs, as well as the opportunity to use technical strategies to boost portfolio efficiency(2). With the advancement of computing capital, trading systems must be able to digest vast amounts of financial data in a variety of formats while reacting rapidly to evolving market conditions. A high-frequency trading system is well suited for quantitative trading. In the early 2000s, it became popular. It accounted for around a quarter of the overall volume by 2005. Quantitative trading volumes rose threefold to 75% in 2009, putting the sector under pressure. Lower commissions, confidentiality, power, discipline, openness, access, competitiveness, and reduced transaction costs are just a few of the advantages of quantitative trading(27). The alpha model, risk model, transaction model, portfolio creation model, and execution model are the five modules of a traditional quantitative trading system. Data collection, data preprocessing, trade analysis, portfolio creation, back-testing, and execution are the six stages of a quantitative trading strategy workflow(28).

2.2 Stock Selection Model

In the process of portfolio construction, it is generally used to predict the stocks within a period, and then select certain stocks according to the predicted results to form a new portfolio according to certain weights.

2.2.1 Statistical Methods

Alizadeh et al. (2) used an adaptive neuro-fuzzy inference system (ANFIS) for stock portfolio prediction. They demonstrated that by using ANFIS and various input features comprising technical factors and fundamental factors, the efficiency of portfolio return prediction could be improved. Deng and Min (3) used a linear regression model comprising ten variables for stock collection in US and global equity. Algorithms shown by Xingyu Fu et al.(20) ranging from traditional statistical learning methods to Logistic Regression, Random Forest, Deep Neural Network and the Stacking, are trained to solve the classification task. Satit Yodmun et al.(21) presents a stock selection approach assisted by fuzzy procedures. Stocks are categorized into categories according to market styles in their approach. The stocks are screened within each category and then rated according to their investment weight gained from blurry quantitative analysis. Groups were also classified by the weight of their group obtained from the fuzzy analytical hierarchy system.

2.2.2 Artificial Intelligence Methods

With the enormous growth of financial data in volume and complexity, machine-learning algorithms provide powerful tools to extract patterns from data processed all across the global. For many years, stock prediction always has drawn attention to the development of intelligent trading systems. There are substantial benefits to be gained from stock prediction for security selection and quantitative investment analysis.

In practice, stock prediction can be conducted by fundamental analysis, technical analysis, and sentiment analysis. Fundamental analysis is the most conventional use, which tries to determine a stock's value or price based on financial statements such as income statement, balance sheet, and cash-flow statement. In other words, the main objective of fundamental analysis is to estimate a company' 's intrinsic value. Fundamental signals have a positive and significant correlation with future earnings performance (41). Fundamental analysis is the prerequisite investigation for value investing as known as long-term investing. In contrast, technical analysis typically begins with charts and technical indicators based on historical data. Technical analysis is usually used to predict short-to medium-term time horizons. An artificial neural network-based stock trading system using technical analysis and big data framework has been proposed in the work of (42). The results have shown that, by choosing the most appropriate technical indicators, the neural network model can obtain comparable results against the buy and hold strategy in most of the cases. Furthermore, fine-tuning the technical indicators and/or optimization strategy can enhance the overall trading performance. In the short-term, the stock market is irrational movement by the effect of emotion trading. Sentiment analysis is the new trend for stock prediction based on finding the correlation between public sentiment and market sentiment. The results show that social media content can give an impact on stock price via sentiment analysis (43; 44). On the effort of improving the prediction accuracy, many studies have been conducted by combining multiple analysis approaches (45; 46).

Recently, there is considerable interest in stock prediction using deep learning methods. Deep learning techniques have been receiving a lot of attention lately, with breakthroughs in image processing and natural language processing. Lin et al. (4) apply the Elman neural network to learn the dynamic behavior of the stock market and predict future return, and where the cross-covariance matric was used to calculate the stock covariance matrix. Paiva et al. (5) used SVM to choose better properties. However, the deep learning application to finance does not yet seem to be commonplace. It has been used for limit order book modeling, financial sentiment analysis, volatility prediction, and portfolio optimization (27; 30). With the effort to decompose and eliminate the noise of the stock price time -series data, the wavelet transform was used. Features are extracted from the decomposed data using stacked autoencoders, and then the high-level denoising features are fed into long short-term memory (LSTM) to build the model and forecast the next day's closing price (31). Stock price exchange rates are forecasted by improving the deep belief network (DBN). The structure of the DBN is optimally determined through experiments and, to accelerate the speed of learning rate, conjugate gradient methods are applied. The model shows more efficiency at foreign exchange rate prediction compared with the feedforward Network (FFNN)(32) . In the work of (33), the recurrent neural network was introduced and used, however, it suffers from the vanishing gradient problem. The vanishing gradient problem was improved in the LSTM and gated recurrent units(GRU) model. The LSTM model has update, input, forget, and output gates, and maintains the internal memory state and applies a non-linearity(sigmoid) before the output gate, whereas GRU has only update and reset gates. Wang et al. (16) utilize LSTM neural network to forecast and stock's potential moving path. Krauss et al. (7) presented and compared the output for statistical arbitrage of multilayer perceptron (MLP), gradient-boosted tree, random forest, and some ensembles of these models. Jiang et al. (8) realized their model in three instants in this work with a convolutional neural network (CNN), a basic RNN, and a LSTM and examined in three back-test experiments with a trading period of 30 minutes in a cryptocurrency market. Malandri et al. (9) use public financial sentiment and historical prices collected from the New York Stock Exchange (NYSE) to train multiple machines learning models for automatic wealth allocation across a set of assets. They show that long short-term memory networks are superior to multilayer perceptron and random forests producing. Freitas et al. (10) used the autoregressive neural network to forecast expected returns. For future stock return estimation, Fischer and Krauss (18) first used LSTM neural network, random forest, MLP and logistic regression. The LSTM neural network worked better than the other memory-free versions, they observed. To restructure the LSTM and render RFG-LSTM, Han et al. (19) use a rectified forgetting gate (RFG) and demonstrate that RFG-LSTM also has the capacity to process sequenced results.

2.3 Quantitative Portfolio Management

2.3.1 Portfolio Construction

Portfolio construction aims to create a portfolio that maximizes expected return for a given level of risk, or, alternatively, minimizes risk for a given expected return over a fixed investment time period. Portfolio construction, in general, is the process of deciding on asset allocation and security selection. In an investment portfolio, asset allocation is a money management technique that determines how capital should be allocated across different asset classes, or broad types of assets, such as stocks, bonds, commodities, and cash. Strategic asset allocation, tactical asset allocation, dynamic asset allocation, constant-weight asset allocation, insured asset allocation, and integrated asset allocation are the six different methods that most asset allocation approaches fall under (29). The portfolio development strategy may be categorized as active or passive depending on the investment strategy, risk tolerance, and liability utilization. The process of selecting individual securities within a specific asset class that will make up the portfolio is known as security selection. After the asset allocation has been determined, the security selection process begins. Following the development of the asset allocation strategy, securities must be chosen to create the portfolio and populate the allocation targets according to the strategy. Though asset allocation is focused on investment strategies, security selection is heavily reliant on forecasting. As a result, determining the expected portfolio return requires a specific investing plan that ensures a portfolio has the right combination of assets to fit individual circumstances, investment goals, and risk tolerance, as well as a highly accurate prediction model. Expected return, asset return variance (volatility), and asset return correlation (or covariance) are the three main inputs for portfolio construction. The estimated return of a portfolio is an estimation of how much profit a portfolio will produce. The variance is a measure of how much risk an investor is willing to take when keeping a portfolio. The portfolio's returns and risk are determined by the returns and risks of the individual stocks and their corresponding portfolio shares.

Statistical measurements of the spread, tails, or distribution of portfolio returns are often used in quantitative portfolio risk management. Variance and standard deviation (spread), coefficient of variance (risk relative to mean), and percentiles of the distribution are examples of such measures (tails). The definition of risk in finance and investment can be expressed in a variety of ways. The most basic and commonly used one, on the other hand, is concerned with risk as an unknown variable that can deviate from expectations. As a result, the average spread or dispersion of a distribution is a natural way to describe a measure of uncertainty. The distances between possible values and the expectation, as well as the probabilities of achieving the various possible values, are two aspects of risk. The variance and standard deviation are two measures that characterize the spread of a distribution, with the standard deviation being the square root of the variance. The higher the spread or dispersion, the higher the variance/standard deviation, and therefore the higher the risk.

The concept behind covariance is to compare the deviations from the mean of two random variables at the same time. Covariance has the problem that its units are products of the original units or the two random variables, making the value difficult to interpret. The correlation coefficient is determined by dividing the covariance by the product of the two random variables' standard deviations.

2.3.2 Portfolio Selection and Optimization

Modern portfolio theory is a theory that explains how risk-averse investors can build portfolios to optimize or maximize expected return based on a given level of risk, emphasizing that higher reward often comes with higher risk. Sharpe also introduced the capital asset pricing model (CAPM) to the industry, which was, in its most basic form, a technique for combining a market portfolio with a risk-free asset to increase the collection of risk-returns above the effective frontier(30). Modern portfolio theory and capital market theory offer a basis for assessing and evaluating investment risk, as well as establishing risk-return relationships. Asset pricing models are the names for these types of partnerships. The arbitrage pricing theory (APT) was established as an alternative to the CAPM by(31). APT was a multi-factor asset pricing model built on the assumption that an asset's returns can be estimated using a linear relationship between the asset's expected return and a variety of macroeconomic variables that capture systemic risk, unlike the CAPM. The Fama French three-factor model was an asset pricing model that built on the CAPM by using size and value risk factors in addition to market risk. This formula takes into account the fact that small-cap and value stocks consistently outperform the index. The model adjusts for this outperforming propensity by adding these two additional variables, which is thought to make it a stronger method for assessing manager success (32). The Black-Litterman model was basically a hybrid of the CAPM and modern portfolio theory(33). The Black-Litterman model's key advantage is that it can be used by a portfolio manager to generate a collection of estimated returns within the mean-variance optimization system. Multiple optimization techniques have been proposed to expand the influence of modern portfolio theory, in addition to developing portfolio theories as the principle of portfolio management. A 60-year review of different approaches developed to address the challenges encountered when using portfolio optimization in practice, such as the transaction costs, portfolio constraints, and estimates errors was provided in (34).

Multi-objective optimization has sparked broad interest in mathematical optimization. Convex programming, integer programming, linear programming, and stochastic programming are only a few of the optimization algorithms that have been developed to solve optimization problems with both linear and random constraints (35)(36). Metaheuristics is a subfield of computational intelligence that represents an effective way to deal with complex optimization problems. It can be used to solve both continuous and combinatorial optimization problems. On complex goals and constraint optimization tasks, evolutionary algorithms such as genetic algorithms have shown to be efficient (37).

In recent years, a lot of research has been performed on financial investment volatility. In order to facilitate portfolio selection, probabilistic programming approaches have been used to deal with the volatility of financial markets. Many practical problems, such as financial risk management, have been solved using fuzzy set theory. Quantitative and qualitative research, expert experience, and investors' subjective strategies can all be better incorporated into a portfolio selection model using fuzzy approaches (38). The input values for optimization models, such as expected returns and risk, which are either determined by a mathematical or statistical model, are based on historical data, which is a major difference between the discussed approaches and this work. In quantitative trading, the alpha and risk models measure the expected return and risk, respectively. In other words, the prediction model is used to determine the input values for the optimization model. In quantitative trading, where complex and large-scale portfolio optimization is the top priority, optimization is performed on expected data, which is an essential prerequisite for successful portfolio management.

Many portfolio selections models based on machine learning method have been proposed during the past few decades. In the following, some recent research related to this study are presented in two perspectives.

2.3.3 Portfolios Based on Machine Learning Models and Mean Variance

The mean–variance approach was proposed by Markowitz to deal with the portfolio selection problem (1). A decision-maker can determine the optimal investing ratio to each security based on the sequent return rate. It can be seen that this portfolio selection method is considered as a static situation. However, this assumption is truly against the real situation. People always vary their optimal portfolio selection with time. Many methods have been proposed to deal with the dynamic portfolio selection problem, several restricted assumptions.

Alizadeh et al. (2) used adaptive network-based fuzzy inference system(ANFIS) for stock portfolio prediction. The suggested method outperformed the classical mean variance model, neural networks and the Sugeno-Yasukawa method, experiments showed. Deng and Min (3) developed a portfolio using the MV model. Experimental findings found that the suggested model outperformed that of the US equity universe in the global equity universe. Lin et al. (4) show a dynamic portfolio selection problem. They obtained optimal dynamic portfolio selection models. Experimental findings found that this model worked better than the model of vector autoregression and provided better results for the problem of dynamic portfolio optimization. Paiva et al. (5) suggested a decision-making process called SVM+MV for stock market financial trading by using the stock price prediction support vector machine (SVM) and the portfolio optimization MV model. This model first used SVM to choose better properties, then used the portfolio optimization MV model. Experimental results showed that SVM was able to boost the portfolio's overall performance and its decision-making model had adequate performance in the Brazilian market. In the portfolios based on machine learning models and mean variable part, Wang et al. (16) merged the neural network of LSTM with the MV paradigm for portfolio construction. They picked the top k stock using the MV model to construct the portfolio. To illustrate its supremacy, they compared their proposed model with four MV models based on three ML models and the autoregressive integrated moving average model.

2.3.4 Portfolios Based on the Predictive Results of Machine Learning Models

In order to forecast potential market return, Yang et al. (6) implemented extreme learning machines and used predictive return as a predictor to create a portfolio optimization model in conjunction with other technical indices. To solve the portfolio optimization problem, differential evolution algorithms were used. The findings revealed that the proposed model outperformed conventional models by using China's A-share market as experimental evidence, which indicated the promising impact of stock estimation for stock selection. Krauss et al. (7) develops portfolios on the basis of the predictive effects of various models by long the top k inventories and short the bottom k inventories. Experiments revealed that the portfolio produced returns of more than 0.45 percent per day prior to the transaction fee, based on an equal weighted ensemble model including MLP, gradient-boosted tree and random forest. Experimental findings found that from 1992 to 2009, this portfolio outperformed the general market but worsened in 2010. Jiang et al. (8) presents a financial-model-free Reinforcement Learning framework consisted of the Ensemble of Identical Independent Evaluators (EIIE) topology, a Portfolio-Vector Memory (PVM), an Online Stochastic Batch Learning (OSBL) scheme, and a fully exploiting and explicit reward function. All three instances of the framework monopolize the top three positions in all experiments, outdistancing other compared trading algorithms and the framework is able to achieve at least 4-fold returns in 50 days. Malandri et al. (9) makes experiments performed on five portfolios. An average increase in the revenue across the portfolios ranging between 5% (without financial mood) and 19% (with financial mood) compared to the equal-weighted portfolio. Moreover, they find that among the employed machine learning algorithms, long short-term memory networks are better suited for learning the impact of public mood on financial time series.

Some studies not only apply predictive return as expected return, but also use the predictive errors to build portfolio optimization model. A novel portfolio optimization model was proposed by Freitas et al. (10), which used predictive errors for portfolio optimization. Experimental findings found that the proposed model outperformed the MV model and yielded a better return for the same risk. A prediction-based portfolio optimization model was proposed by Freitas et al. (11) by using the autoregressive moving reference neural network AR-MRNN model as a predictor. This paper first used the AR-MRNN model to predict future stock returns and then used the variance of predictive error as a risk to set up a model for portfolio optimization. Experimental results showed that the proposed model exceeded the classical mean variance model based on an effective border and real stock market performance analysis. By using

SVR for future stock return prediction and the variance of predictive errors as risk for portfolio optimization, Hao et al. (12) presented a predictive-based portfolio selection model, comparing their proposed model with the model in (11). Experimental findings demonstrated that their model performed better. They also mentioned that the better prediction of future stock return gave their model better performance.

While in many domains, LSTM networks have been seen to be more precise, they still lack implementations in the financial industry. Heaton et al. (13; 14; 15) started research on network construction for financial markets. They also identified LSTM templates for all capital exchanges, such as NYSE, AMEX, and Nasdaq, through their research. This capital exchanges have a valuation of more than 5 billion US dollars (the total monthly data is for 848,000 stocks).

In the Portfolios based on the predictive results of machine learning models' part, the efficiency of the recurrent neural network gated recurrent unit and LSTM neural network for stock return prediction was first compared by Lee and Yoo (17). Experimental findings found that the other models outperformed the LSTM neural network. Predictive threshold-based portfolios with the predictive results of the LSTM neural network were also proposed and satisfactory output was produced. Fischer and Krauss (18) developed a portfolio focused on the LSTM neural network's predictive results using the same approach in (7). Han et al. (19) are building a multi-factor alpha portfolio with RFG-LSTM based on the current trading scenario of China's A financial market. Experimental findings show that the RFG-LSTM model can learn the dynamics and laws of the A Financial Market critically, and this can lead to the investing policy of the portfolio.

2.3.5 Portfolio Performance Evaluation

The concept of market efficiency is put to the test by evaluating portfolio results. The evaluation is carried out for three key reasons: to enhance performance, to track risk, and to evaluate returns. The efficiency of a portfolio can be measured using a number of different metrics. Two desirable characteristics for an efficient portfolio are the ability to generate above-average returns for a given risk class and the ability to fully diversify the portfolio to minimize all unsystematic risk relative to the portfolio's benchmark. Performance assessment approaches are classified into two categories: traditional and risk-adjusted methods (39). Benchmark comparison and style comparison are two of the most commonly used traditional approaches. Returns are adjusted using risk-adjusted approaches to account for variations in risk levels between the managed and benchmark portfolios. Traditional approaches are favored over risk-adjusted methods. The work of (40) lists some of the most important portfolio efficiency metrics.

Chapter 3

Methodology

Our proposed methodology architecture is developed based on the typical quantitative investment management system. Historical data were collected from multiple resources. Multiple prediction models were conducted to predict stock prices such as Linear Regression, XGBoost, and LSTM. On the basis of the predicted results for each period, the cumulative return and Sharpe ratio were calculated. The portfolio was constructed by selecting the outperform stocks from the predicted result in terms of the highest cumulative return and lowest risk. Optimal stock allocation for the constructed portfolio was evaluated by optimization modeling. Maximizing the Sharpe ratio were used to evaluate the optimal stock allocation weights.

3.1 Prediction Model

In this section, we proposed LSTM network to predict the stock price. Random Forest, XGBoost, Linear Regression model are the benchmarks.

3.1.1 LSTM Neural Network

The LSTM network is an RNN variant with recurrently linked memory blocks in the hidden layer. The cell state and the hidden state are the two states that are passed from one cell to the next. Memory blocks are in charge of remembering information, and they are manipulated through three main mechanisms known as gates. The addition of information to the cell state is handled by a forget gate. The output gate determines which hidden state will be chosen next. Operations performed on LSTM network units are explained in the following formulas, where x_t is the input at time t and f_t is the forget gate at time t, which clears information from the memory cell when needed and keeps a record of the previous frame whose information from the memory. The output gate o_t keeps the information about the upcoming step, where g is the recurrent unit with the activation function "tanh," and is computed from the current frame's input and the previous frame's state h_{t-1}. In all input (I_t), forget (f_t), and output (O_t) gates, as well as the recurrent unit (g_t), we use (W_i, W_f, W_o, W_g) and (b_i, b_f, b_o, b_g) as weights and bias, respectively. The input gate decides the components of the transformed input g_t should be applied to the long-term state c_t. This procedure changes the long-term state c_t, which is then passed on to the next cell. Finally, the output gate filters the modified long-term state c_t, transforms it into $tanh(.)$, and generates the output y_t, which is sent to the next cell as the short-term state h_t.

$$i_t = \sigma(W_{x_i}^T x_t + W_{h_i}^T h_{t-1} + b_i), \tag{2}$$

$$f_t = \sigma(W_{x_f}^T x_t + W_{h_f}^T h_{t-1} + b_f), \tag{3}$$

$$o_t = \sigma(W_{x_o}^T x_t + W_{h_o}^T h_{t-1} + b_o), \tag{4}$$

$$g_t = tanh(W_{x_g}^T x_t + W_{h_g}^T h_{t-1} + b_g), \tag{5}$$

$$c_t = f_t c_{t-1} + i_t g_t, \tag{5}$$

$$y_t = h_t = o_t tanh c_t, \tag{6}$$

Where $\sigma(.)$ is the logistic function, and $tanh(.)$ Is the hyperbolic tangent function. The gate controllers f_t,i_t, and o_t, which are all completely connected layers of neurons, monitor how the three gates open and close. Since they use the logistic function for activation, the range of their outputs is $[0, 1]$. The outputs of each gate are fed into element-wise multiplication operations in each gate, so if the output is close to 0, the gate is narrowed and less memory is stored in c_t, while if the output is close to 1, the gate is more widely accessible, allowing more memory to flow through the gate. When using LSTM cells, it's popular to stack multiple layers of cells to make the model deeper and capture the data's nonlinearity. The computation in an LSTM cell is shown in Figure 3.1. To maintain the wealth of a stock market, we must have an effective prediction model that can forecast based on previous stock market results. In this paper, we used LSTM networks to build a model that can predict the stock price(52). On the basis of the output of the prediction price, a portfolio is constructed.

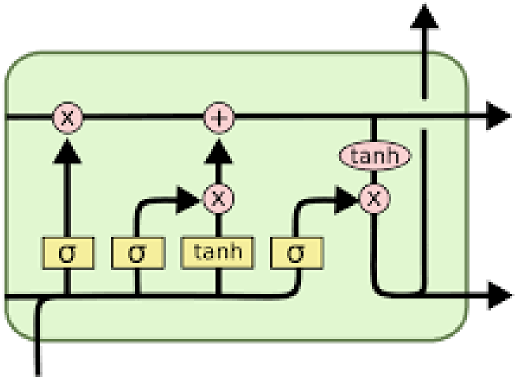

Figure 3.1: LSTM network architecture

3.1.2 Linear Regression

In statistics, linear regression(LR) is a linear approach to modelling the relationship between a scalar response and one or more explanatory variables (also known as dependent and independent variables). The case of one explanatory variable is called simple linear regression; for more than one, the process is called multiple linear regression.

Given a data set $\{y_i, x_{i1}, \ldots x_{ip}\}_{i=1}^{n}$ of n statistical units, a linear regression model assumes that the relationship between the dependent variable y and the p-vector of regressors x is linear. This relationship is modeled through a disturbance term or error variable ε — an unobserved random variable that adds "noise" to the linear relationship between the dependent variable and regressors. Thus, the model takes the form

$$y_i = \beta_0 + \beta_1 x_{i1} + +_p x_{ip} \varepsilon_i = x_i^T + \varepsilon_i \qquad i = 1, \ldots n \tag{1}$$

3.1.3 Random Forest

Random forest(RF) is an ensemble machine learning algorithm extended from the classification and regression trees. To solve a classification problem, a classification tree is built from an independently sampled subset from the original dataset. Instead of growing only one classification tree by classification and regression trees(CART), RF constructs a hundred of classification trees using randomly selected subsets of training samples and predictor variables for binary splits. The modes of all the predicted classes from all trees are determined as the model output. The principle of RF is to construct “a forest of classification/regression trees” through combining a set of “weak learners” to form a “strong learner” for improving predictive performance. RF models are capable of handling missing data, correlated predictor variables and non-linearity, insensitive to noise, and can handle very large numbers of input variables.

3.1.4 XGBoost

XGBoost algorithm was proposed by Chen and Guestrin (53) based on the Gradient Boosting Decision Trees(GBDT) structure. Different from the GBDT, XGBoost introduces the regularization term in the objective function to prevent overfitting. The objective function is defined as

$$O = \sum_{i=1}^{n} L(y_i, F_{n-1}(x)) + \sum_{k=1}^{t} R(f_k) + C \tag{7}$$

Where $R(f_k)$ denotes the regularization term at the k time iteration, and C is a constant term, which can be selectively omitted. The regularization term $R(f_k)$ is expressed as

$$R(f_k) = \alpha H + \frac{1}{2}\eta \sum_{j=1}^{H} \omega_j^2 \tag{8}$$

Where α represents the complexity of leaves, H indicates the number of leaves, η denotes the penalty parameter, and ω_j is the output result of each leaf node. In particular, the leaves indicate the predicted categories following the classification rules, and the leaf node indicates the node of the tree that cannot be split.

Moreover, as opposed to use the first-order derivative in GBDT, a second-order Taylor series of objective function is adopted in XGBoost. Suppose the mean square error is used as the loss function, then the objective function can be derived as

$$O = \sum_{i=1}^{n} [p_i \omega_{q(x_i)} + \frac{1}{2}(q_i \omega_{q(x_i)}^2)] + \alpha H + \frac{1}{2}\eta \sum_{j=1}^{H} \omega_j^2 \tag{9}$$

where $q(x_i)$ indicates a function that assigns data points to the corresponding leaves, and g_i and h_i denote the first and second derivative of loss function, respectively.

The final loss value is calculated based on the sum of all loss values. Because samples correspond to leaf nodes in the decision tree, the ultimate loss value can be determined by summing the loss values of leaf nodes. Therefore, the objective function is also expressed as

$$O = \sum_{i=1}^{n} [P_j \omega_j + \frac{1}{2}(Q_j + \eta)\omega_(j)^2] + \alpha H \tag{10}$$

Where $P_j = \sum_(i \in I_j) p_j, Q_j = \sum_{i \in I_j} q_j$, and I_j indicates all samples in leaf node j.

To conclude, the optimization of objective function is converted to a problem of determining the minimum of a quadratic function. In addition, because of the introduction of regularization term, XGBoost has a better ability to against overfitting.

3.2 Quantitative Models

3.2.1 Multiple Assets Portfolio Construction

Assume a portfolio of N stocks, and S_0 is the set of initial values for each stock in the portfolio, denoted by $S_0 = (s_1^0, \ldots s_N^0).X = (x_1, \ldots, x_N)$ represents the number of stocks in the portfolio. The following formula is used to determine the portfolio's initial value,

$$V_0 = x_1 s_1^0 + ... + x_N s_N^0 = \sum_{i=1}^{N} x_i s_i^0 \tag{11}$$

The number of shares of each commodity will be determined after the division of our money, which is our primary concern, and will be expressed as the weights $W = (\omega_1, \dots \omega_N)$ with the constraint $\sum_{i=1}^{N} \omega_i = 1$, defined by $\omega_i = (x_i s_i^0)/V_0$ with $i = 1, \dots, N$ At the end of the period t, the values of the stocks change $S_t = (s_1^t, \dots, s_N^t)$, which gives the final value of the portfolio V_t as a random variable,

$$V_t = x_1 s_1^t + ... + x_N s_N^t = \sum_{i=1}^{N} x_i s_i^t \tag{12}$$

The set of random returns on each stock in the portfolio is $R_P = (r_1, \dots r_N)$ and the vector of expected return is $\mu = (\mu_1, \dots \mu_N)$ with $\mu_i = E(r_i)$ for $i = 1, 2, \dots N$. The real return on a portfolio of multiple assets over a given time span can be estimated simply as follows:

$$R_P = \omega_1 r_1 + \omega_2 r_2 + ... + \omega_N r_N \tag{13}$$

The cumulative return on a portfolio can be estimated as follows:

$$R_C = \omega_1(1 + r_1)x_1 s_1^0 + \omega_2(1 + r_2)x_2 s_2^0 + ...\omega_N(1 + r_N)x_N s_N^0 \tag{14}$$

The weighted average of the expected returns of each asset in the portfolio is the expected portfolio return. The weight assigned to each asset's projected return is the ratio of the asset's market value to the portfolio's overall market value. As a result, the portfolio's estimated return $E(R_P) = \mu_P$ at the end of time t is determined as follows:

$$E(R_P) = \omega_1 E(r_1) + \omega_2 E(r_2) + ... + \omega_N E(r_N) = \sum_{i=1}^{N} {}_i\mu_i \tag{15}$$

Variance of return for the portfolio used above part as follows:

$$Var(R_P) = E(R_P - \mu_P)^2 = E(R_P^2) - \mu_P^2 \tag{16}$$

The variance of the return can be computed from the variance of S_t,

$$Var(R_P) = Var(\frac{(S_t - S_0)}{S_0}) = \frac{1}{S_0^2} Var(S_t - S_0) = \frac{1}{S_0^2} Var(S_t) \tag{17}$$

$\sqrt{(Var(R_P))}$ are the standard deviations of various random returns. The covariance between asset returns will be denoted by $Cov(r_i, r_j)$, in particular $\sigma_{ii} = \sigma_i^2 = Var(r_i)$. These are the entries of the NN covariance matrix Cov,

$$Cov(r_i, r_j) = E[(r_i - \mu_i)(r_j - \mu_j)] \tag{18}$$

$$Cov = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1N} \\ \vdots & \ddots & \vdots \\ \sigma_{N1} & \cdots & \sigma_{NN} \end{bmatrix} \tag{19}$$

3.3 Stock Prediction Evaluation

The mean absolute error (MAE) and root mean squared error (RMSE) were used to calculate the difference between the expected and actual data in order to assess predicted error rates and model efficiency. The following formulas were used to measure MAE, RMSE:

$$MAE = \frac{\sum_{i=1}^{T} |y_i - y_i^{'}|}{T} \tag{20}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{T} (y_i - y_i^{'})^2}{T}} \tag{21}$$

3.4 Portfolio Optimization

The aim of portfolio optimization is to determine the best asset allocation based on stock price prediction phrases. Top-down investing was modified to select the best-performing stock based on the prediction model in order to construct a multi-asset portfolio. The estimated return and standard deviation for each stock was determined using the stock prediction results. The top predicted output stocks with the highest predicted expected returns will be chosen for each time span to construct a portfolio with initial weights, with EQ being the most frequently allocated. The portfolio total return is measured by multiplying the number of stocks by the correlated weights. The optimal set is a collection of current allocation weights for the built portfolio's selected stocks. We can determine the optimal weights for the selected stocks in the built portfolio by changing the model parameters of the portfolio optimizers. Instead of using the traditional EQ approach and optimization techniques were used to find the best weights for the built portfolio.

3.4.1 Maximizing the Sharpe ratio

A portfolio constructed from N different assets can be described by means of the vector of weights $W = (W_1, W_2, ...W_N)$, with the constraint given $\sum_{i=1}^{N} W_i = 1$. The N-dimensional vector $I = (1, 1, ., 1)$ is denoted by I. Therefore, the constraint can conveniently be written as $W^T I = 1$. Denote the random returns on the stocks by $r_1, ..., r_N$, and the vector of expected return by $mu = (\mu_1, \mu_2, ..., \mu_N)$ with $\mu_i = E(r_i)$ for$i = 1, 2, , N$. The covariances between returns will be denoted by $\sigma_{ij} = Cov(r_i, r_j)$,in particular $\sigma_{ii} = \sigma_i^2 = Var(r_i)$ these are the entries of the covariance matrix Φ.

$$\Phi = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1N} \\ \vdots & \ddots & \vdots \\ \sigma_{N1} & \cdots & \sigma_{NN} \end{bmatrix} \tag{22}$$

The expected return $\mu_P = E(R_P)$ and covariance $\sigma_P^2 = Var(R_P)$ of a portfolio with weights W are given by

$$\mu_P = \sum_{i=1}^{N} W_i \mu_i = W_T \mu \tag{23}$$

$$\sigma_P^2 = Var(R_P) = W_T \Phi W \tag{24}$$

To get weight of the maximum Sharpe Ratio, let r_f be the risk-free interest rate. Consider:

$$\max_{W} \quad \frac{\mu^T W - r_f}{\sqrt{W^T \Phi W}} \tag{25}$$

$$s.t. \quad W_i > 0 \quad W_T I = 1$$

After solve formula 25, we can get the optimal asset allocation weight.

Chapter 4

Experiment

4.1 Data Collection and Experiment Design

4.1.1 Data Collection

In this section, Yahoo Finance has collected the daily earnings price of the shares of the top 10 weight utilities companies on Sector SPDR ETF XLU as of October 18, 2008 at solstice as of September 29, 2020, for a total of 3009 trading days. The Symbol and Company Name are shown in the Table below. The key input values to the data set are the regular daily close price and the last twenty days close price of each stock and last five days' INX and DJIA close price.The model employs an 80:20 train-to-test split ratio.

Table 4.1: Index and Company Name

Index	Company Name
NEE	NextEra Energy Inc
DUK	Duke Energy Corp
SO	Southern Co
D	Dominion Energy Inc
EXC	Exelon Corp
AEP	American Electric Power
SRE	Sempra Energy
XEL	Xcel Energy Inc
PEG	Public Service Enterprise Grp
WEC	WEC Energy Group Inc

4.1.2 Optimized Hyperparameter

To find the best model parameters, we iterated the number of neurons from 50 to 300 and iterated the number of epochs from 5 to 30. Rooted mean square error (RMSE) and mean absolute error(MAE) were used to measure prediction errors; In figure 4.1, the least loss error was obtained at 25 epochs. In figure 4.2, The minimum prediction failure error was observed at 250 neurons. We use a layered LSTM architecture with two LSTM hidden layers (57). Adam optimization, as reported in (51), is better suited to deep learning problems with larger datasets. As a result, in our tests, we used the Adam optimizer with default Keras parameters. Table 4.2 summarizes the details of the selected hyperparameters. The scikit-learn library was used to train prediction models using machine learning algorithms.

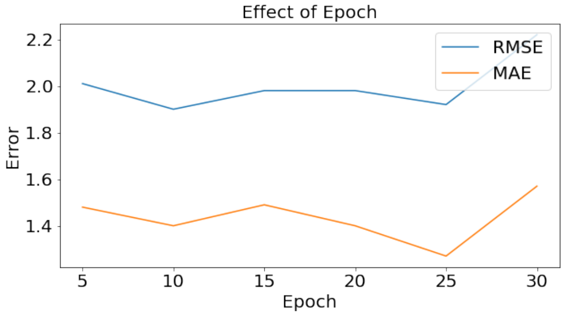

Figure 4.1: Epoch

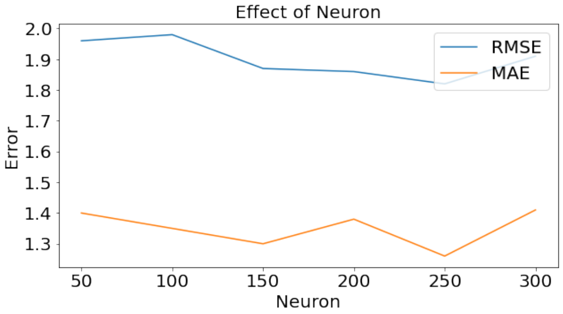

Figure 4.2: Neuron

Table 4.2: LSTM Hyperparameters setup

Categories	Hyperparameter
Optimizer	Adam
The number of Neuron	250
the number of Hidden Layer	2
Number of Epochs	25

We also find the optimal hyperparameters for the benchmarks. According to (57), there is no need to find the hyperparameters for Linear Regression model. So, we just find the hyperparameters for random forest and XGBoost. Table 4.3 and Table 4.4 summarizes the details of the selected hyperparameters of the two methods.

Table 4.3: XGBoost Hyperparameters setup

Categories	Hyperparameter
The number of estimators	20
max depth	10
Learning rate	0.2

Table 4.4: Random Forest Hyperparameters setup

Categories	Hyperparameter
The number of estimator	30
Max depth	20

4.1.3 Stock Prediction Results

The prediction procedure follows the described in Chan et al.(54) called rolling window method. This method is used in the prediction process. We have made a minor adjustment to the protocol. This technique, in particular, is divided into two sections. The first section is the training section, which is where the model is trained, and model parameters are updated. The second part is the test, in which we use the best model to predict results. In particular, our time frame for each portion differs from that of Chan et al.(54). In the training phase, we use data from the previous 10 years to train the models. The test part is put to the test for the next three months (a calendar quarter). In the test part, in line with popular portfolio management practice, we predict the quarterly performance of each model. This process continues for three years on each quarter. The prediction procedure is illustrated in figure 4.3.

Table 4.5: Prediction Period

Period	Start	End
1	2018-05-16	2018-10-05
2	2018-10-06	2019-03-01
3	2019-03-02	2019-07-23
4	2019-07-24	2019-12-11
5	2019-12-12	2020-05-05
6	2020-05-06	2020-09-29

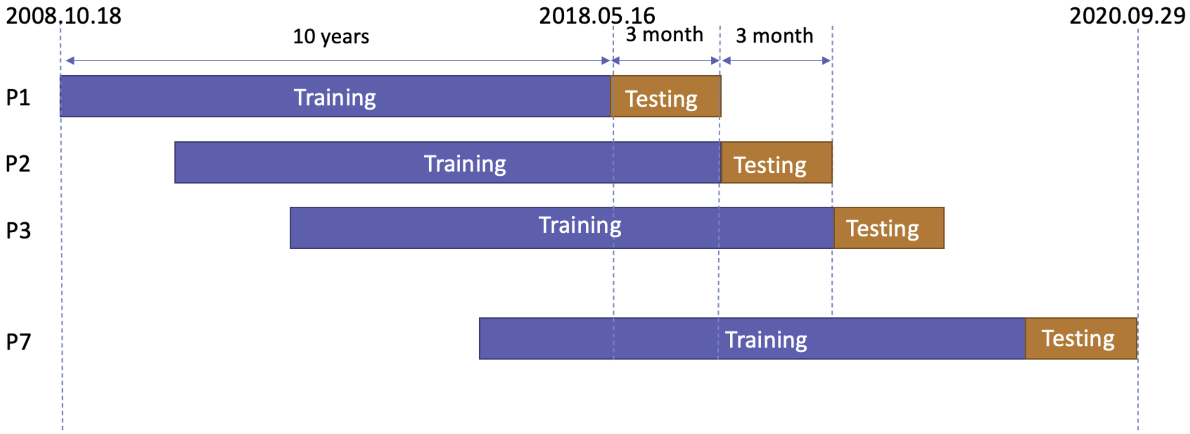

Figure 4.3: Rolling window for prediction

We random choose a stock index and show the yearly predicted data from the four models and the corresponding actual data in the graph. Figure 4.4 is the result. According to figure 4.4, we can find that XGBoost and Random Forest have larger variations and distances to the actual data than LSTM and Linear Regression. Furthermore, comparing LSTM with Linear Regression, the former out- performs the latter: LSTM has less volatility and is closer to the actual trading data than Linear Regression.

Table 4.5 records the model performance in prediction the stock price. It can be

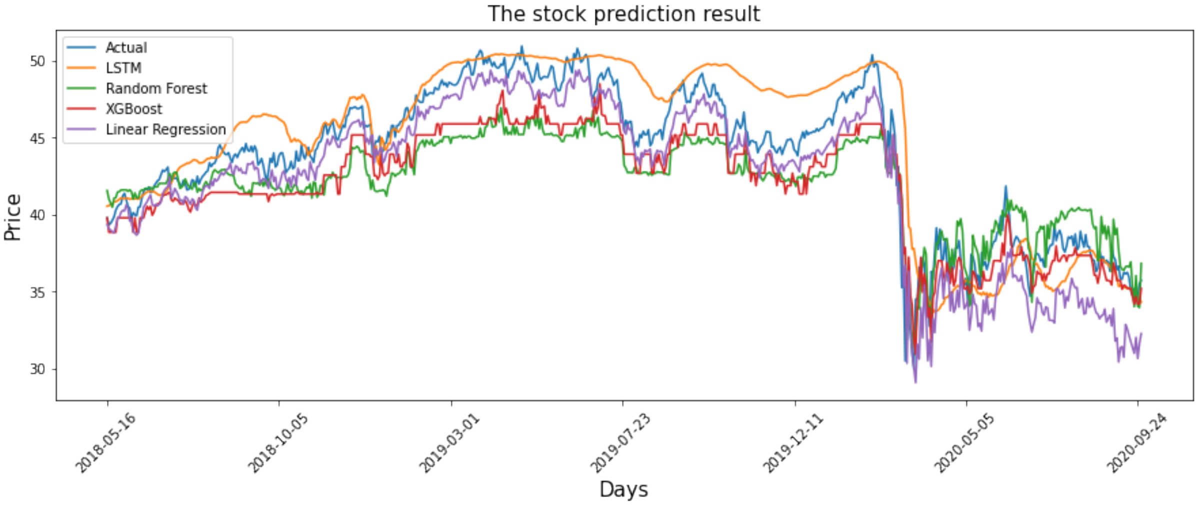

Figure 4.4: Prediction Result

Table 4.6: The error of each method

Method	LR	RF	XGBoost	LSTM
RMSE	2.88	2.88	2.64	2.6
MAE	2.19	2.35	2.15	1.68

seen from the table that LSTM shows much better performance than the other three models in predicting both stock indices. For example, the value of RMSE and MAE of LSTM reach 2.6 and 1.68, respectively, which is much less than those of the other three models.

4.1.4 Portfolio Construction

On the basis of predicting stock prices by using LSTM, we use different investment strategies to construct investment portfolios. First, we will use the LSTM model to predict the price changes of ten stocks in the next three years. After that, the maximizing Sharpe ratio optimization method was used to optimize the weight of the investment portfolio and finally buy stocks with the Buy and Hold strategy. The weight will not change in the entire test period. The optimal weight is $W_{Buyandhold} = [0.47567, 0, 0, 0, 0.13284, 0.32276, 0.06873, 0, 0, 0]$

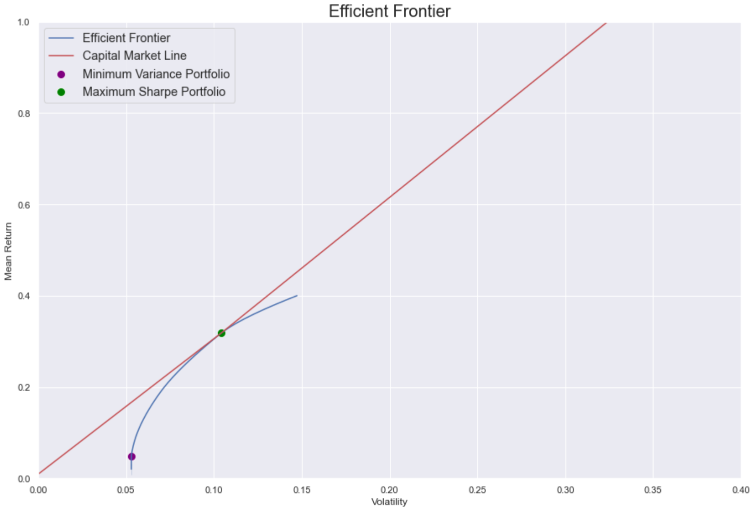

Figure 4.5: Efficient frontier Line and the maximum sharpe ratio portfolio

In contrast, the second investment strategy does not use any portfolio weight optimization method, namely the equal-weight portfolio method(EQ). The weights chosen here are $W_{EQ} = [0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1]$

We also use the top 10 Sector ETF XLU weights to make the third benchmark at the actual result. We normalize the weight and get the

$$w_{XLU} = [0.258, 0.129, 0.116, 0.107, 0.074, 0.074, 0.069, 0.064, 0.053, 0.051]$$

The last strategy is to use the maximizing Sharpe ratio method to build the portfolio at the beginning point and rebalance the allocation strategy after the rolling window moves to another period.

Through the experiment, the weight of the investment portfolio in each period is shown in Table 4.7. During Period 4 because the cumulative return of all stocks at the end of Period 4 is lower than at the beginning of Period 4. So, we don't invest anything in Period 4.

Table 4.7: Weights in different period

Period	P1	P2	P3	P4	P5	P6
NEE	0.0	0.0	0.41191	0.78366	0.0	0.47285
DUK	0.0	0.0	0.13424	0.0	0.0	0.02622
SO	0.0	0.0	0.0	0.13529	0.0	0.09605
D	0.0	0.0	0.0	0.0	0.0	0.05134
EXC	0.28924	0.93942	0.31443	0.04093	0.0	0.04563
AEP	0.0	0.0	0.0	0.04010	0.0	0.0
SRE	0.58223	0.06057	0.08265	0.0	0.0	0.21848
XEL	0.0	0.0	0.0	0.0	0.51769	0.01641
PEG	0.12851	0.0	0.05673	0.0	0.0	0.07299
WEC	0.0	0.0	0.0	0.0	0.48230	0.0

4.2 Experiment Results

4.2.1 Portfolio Performance Evaluation

We evaluated the constructed portfolios based on prediction models. In order to optimize the performance of the constructed portfolio, maximizing Sharpe ratio optimization method were employed to evaluate the impact of optimization on portfolio performances. In the majority of cases, expected returns showed a tendency to increase.

According to figure 4.6 and table 4.8, the maximizing Sharpe ratio optimization method is obviously better than the equal weight method. The EQ method has the lowest expected cumulative return. This indicates that the optimization method can improve the performance of the constructed portfolio by increasing the SR value.

In other words, optimization techniques are not only guaranteed to increase returns, but also to reduce trading risk. The buy and hold approach have more volatility. On the premise that transaction costs are not taken into account, the use of maximizing Sharpe ratio method can significantly improve the expected cumulative return of trading and can greatly reduce the investment risk.

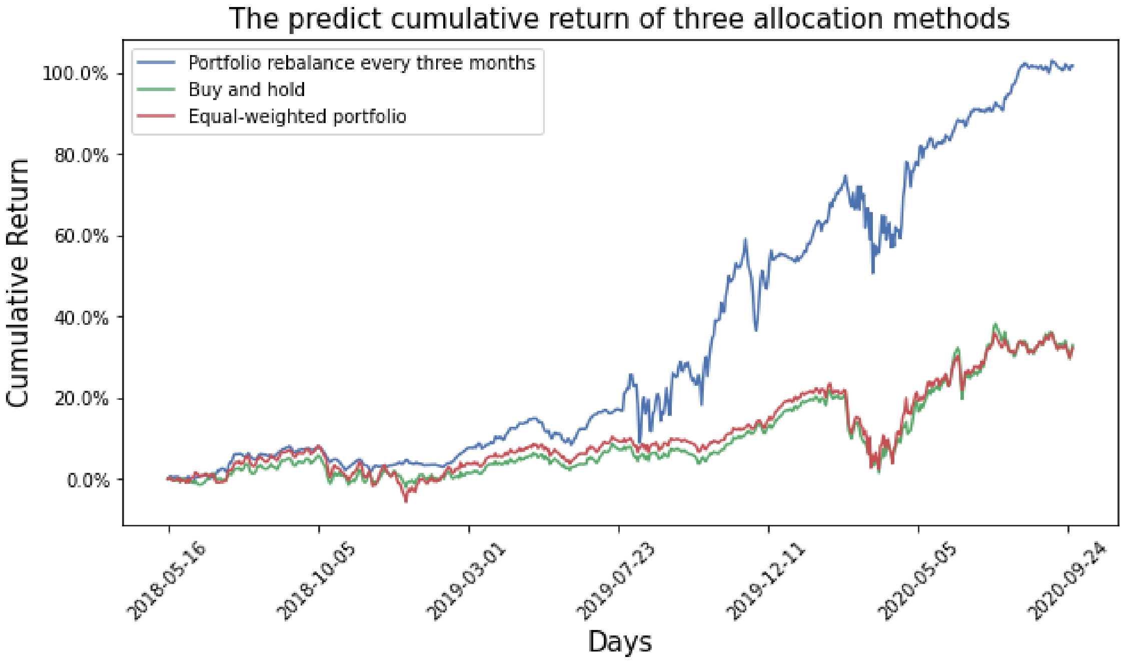

Figure 4.6: The predicted result of portfolio performance

Table 4.8: The predicted cumulative return of each allocation method

Allocation methods	Cumulative return
Portfolio rebalance every three month	101.7926%
Buy and hold	32.9575%
Equal-weigthed portfolio	32.209%

Secondly, the portfolio constructed in the prediction section is tested in the actual transaction. According to figure 4.7 and table 4.9, Due to the impact of the epidemic in 2020, the stock market has been seriously affected. But resetting the allocation of different assets in your portfolio every three months can still get higher returns than the other method. As shown in figure 4.7, it has the highest cumulative returns. Therefore, the portfolio constructed based on LSTM prediction model and maximizing Sharpe ratio optimization method is better than other prediction models and allocation strategies.

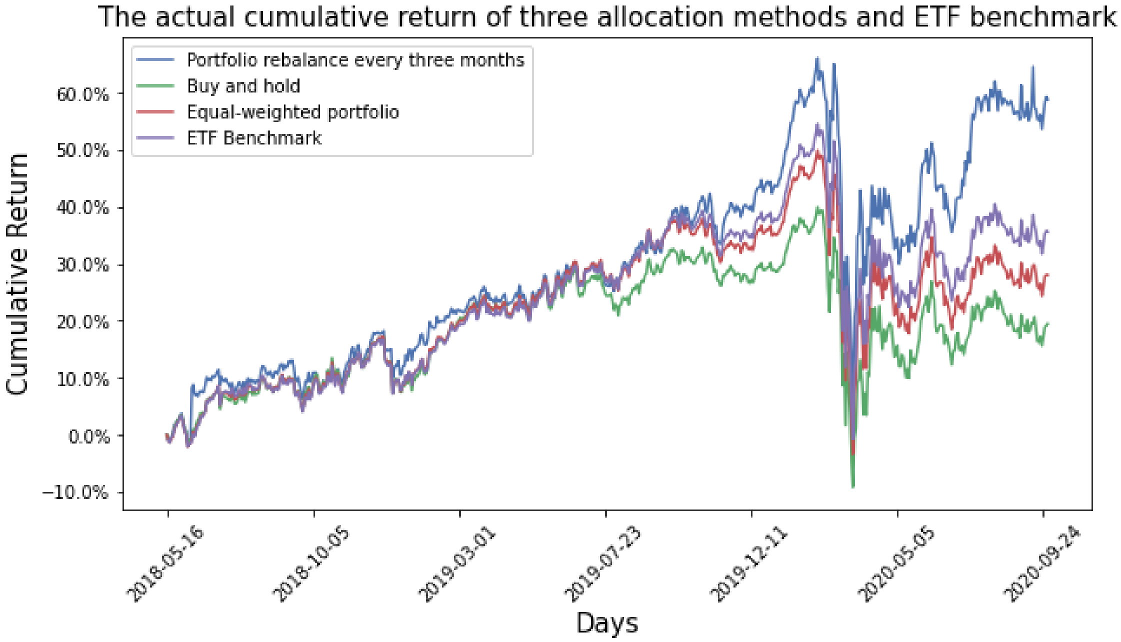

Figure 4.7: The actual result of portfolio performance

Table 4.9: The actual cumulative return of each allocation method

Allocation methods	Cumulative return
Portfolio rebalance every three month	58.7866%
Buy and hold	19.4529%
Equal-weigthed portfolio	28.0164%
ETF Benchmark	35.588%

The validity of the proposed prediction and optimal allocation model in prediction and actual transaction is evaluated. The results show that the maximizing Sharpe ratio method is superior to the proposed allocation strategies. Effective prediction is not only prediction, but also the support of the optimization allocation. These constructed portfolios were selected as efficient portfolios for quantitative trading. And the portfolio we built that also outperformed the benchmark Sector ETF XLU.

4.3 Test in Sector ETF XLB

In order to make our proposed method more general. We tested our approach on sector ETFs XLB. We use Yahoo Finance has collected the daily earnings price of the shares of the top 10 U.S. utilities companies on Sector SPDR ETF XLB as of October 18, 2008 at solstice as of February 23, 2021. The Symbol and Company Name are shown in the table 4.10.

Table 4.10: Index and Company Name

Index	Company Name
LIN	Linde plc
SHW	The Sherwin-Williams Company
APD	Air Products and Chemicals, Inc
ECL	Ecolab Inc
FCX	Freeport-McMoRan Inc
NEM	Newmont Corporation
DD	DuPont de Nemours, Inc
PPG	PPG Industries, Inc
IFF	International Flavors Fragrances Inc.
BLL	Ball Corporation

We use the same process to analysis Sector ETF XLB dataset. Firstly, according to figure 4.8 and table 4.11, we can see from the predicted results that the use of maximizing Sharpe ratio method can significantly improve the cumulative return of the transaction without considering the transaction cost. And volatility is much lower than other methods, which suggests that reweighting a portfolio every three months can significantly reduce risk.

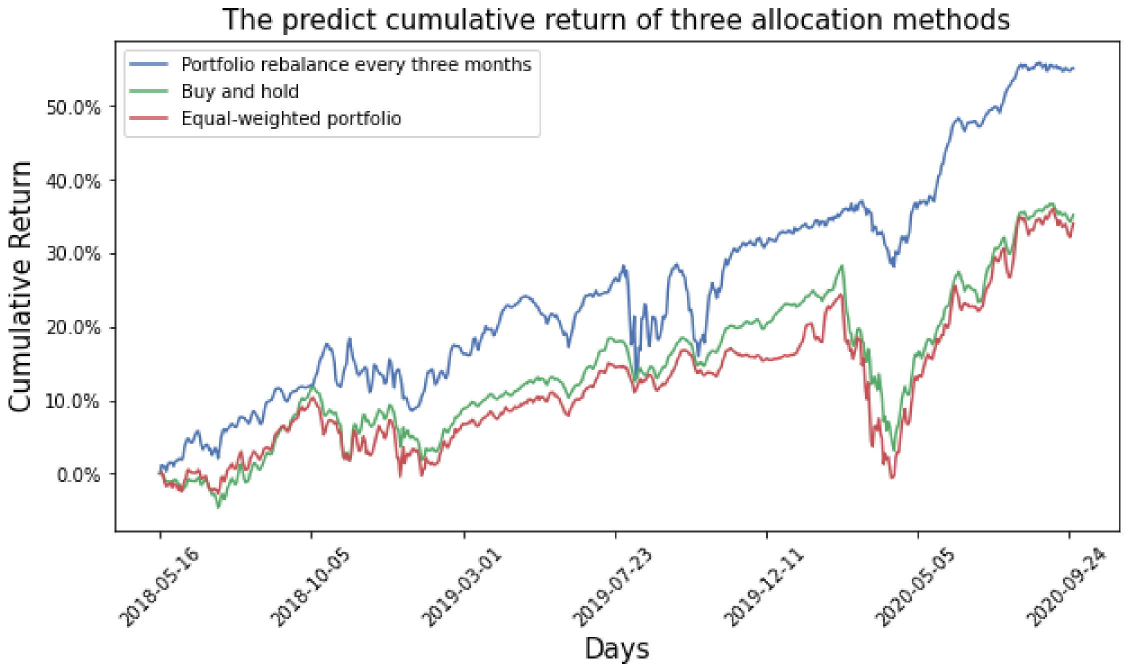

Figure 4.8: The actual result of XLB

Table 4.11: The predicted cumulative return of each allocation method

Allocation methods	Cumulative return
Portfolio rebalance every three month	55.1672%
Buy and hold	35.1743%
Equal-weigthed portfolio	34.0457%

Secondly, accoring to figure 4.9 and table 4.12, we can see from the actual dataset that, due to the impact of the epidemic, although the stock market has been impacted, the method of adjusting the weight of the investment portfolio every three months is not only still significantly better than other methods, including the Sector ETF XLB, but also performs better than that on the sector ETF XLU dataset. So our method can be generalized to other datasets.

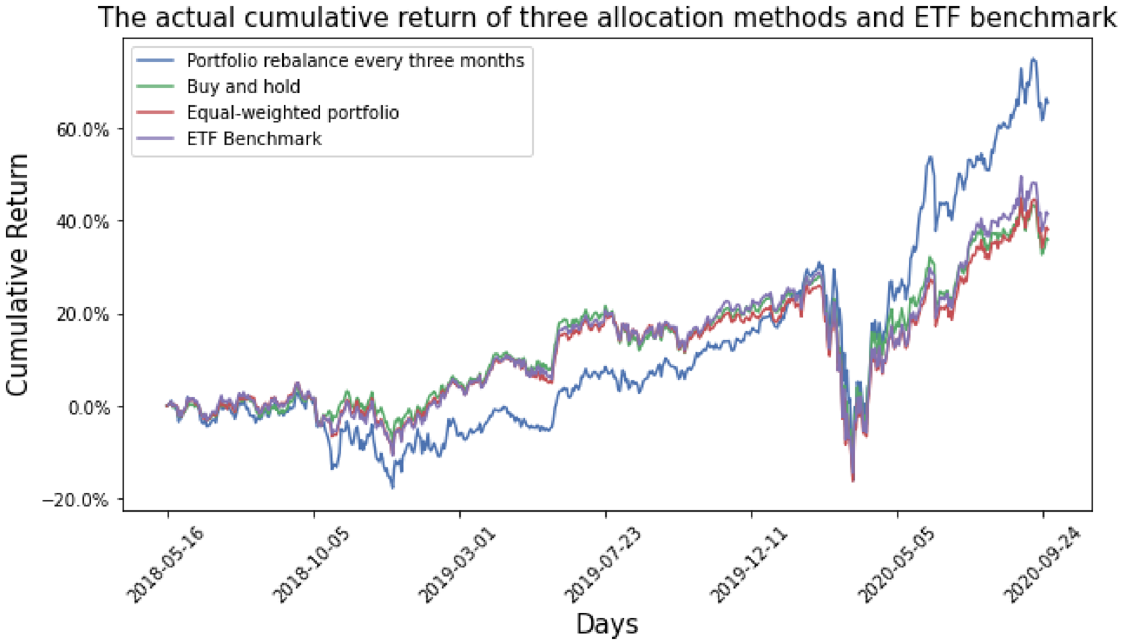

Figure 4.9: The actual result of XLB

Table 4.12: The actual cumulative return of each allocation method

Allocation methods	Cumulative return
Portfolio rebalance every three month	65.4738%
Buy and hold	35.9791%
Equal-weigthed portfolio	38.1279%
ETF Benchmark	41.4411%

Chapter 5

Conclusions and Discussions

Stock prediction plays a major role in building an investment portfolio. In order to illustrate a conventional quantitative trading technique, this paper used the LSTM network, a form of recurrent neural network, to forecast the stock price. As opposed to other machine learning techniques such as Linear Regression and XGBoost, the proposed model performs well. As a consequence, we will use the forecast results to construct a predictive portfolio for each time period projected. In comparison to the Sector ETF XLU, our built portfolios performed well using optimization methods, achieving larger returns in both predicted and actual result.

Deep learning presents a number of problems when it comes to developing successful quantitative trading strategies. To begin with, market data has a high noise-to-signal ratio. On the historical data collection, the prediction models perform well. The stock market, on the other hand, is constantly fluctuating due to factors such as market psychology, macroeconomics, and even political issues. As a result, strong output on a historical dataset does not guarantee a profitable outcome in real-world trading. Second, backtesting is useful not only for evaluating the newly discovered technique, but also for avoiding false positives. In conclusion, the deep learning approach has a significant impact on stock prediction efficiency, which may be a prerequisite for portfolio creation and optimization in quantitative trading.

Bibliography

[1] Markowitz H. (1952) Portfolio selection. J Finance 7:77–91.

[2] Alizadeh, M., Rada, R., Jolai, F., & Fotoohi, E. (2011). An adaptive neuro-fuzzy system for stock portfolio analysis. International Journal of Intelligent Systems, 26(2), 99-114.

[3] Deng, S., & Min, X. (2013). Applied optimization in global efficient portfolio construction using earning forecasts. The Journal of Investing, 22(4), 104-114.

[4] Lin, C. M., Huang, J. J., Gen, M., & Tzeng, G. H. (2006). Recurrent neural network for dynamic portfolio selection. Applied Mathematics and Computation, 175(2), 1139-1146.

[5] Paiva, F. D., Cardoso, R. T. N., Hanaoka, G. P., & Duarte, W. M. (2019). Decision-making for financial trading: A fusion approach of machine learning and portfolio selection. Expert Systems with Applications, 115, 635-655.

[6] Yang, F., Chen, Z., Li, J., & Tang, L. (2019). A novel hybrid stock selection method with stock prediction. Applied Soft Computing, 80, 820-831.

[7] Krauss, C., Do, X. A., & Huck, N. (2017). Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. European Journal of Operational Research, 259(2), 689-702.

[8] Jiang, Z., Xu, D., & Liang, J. (2017). A deep reinforcement learning framework for the financial portfolio management problem. arXiv preprint arXiv:1706.10059.

[9] Malandri, L., Xing, F. Z., Orsenigo, C., Vercellis, C., & Cambria, E. (2018). Public mood–driven asset allocation: the importance of financial sentiment in portfolio management. Cognitive Computation, 10(6), 1167-1176.

[10] de Freitas, F. D., De Souza, A. F., & de Almeida, A. R. (2006). A prediction-based portfolio optimization model.

[11] Hao, C., Wang, J., Xu, W., & Xiao, Y. (2013, November). Prediction-based portfolio selection model using support vector machines. In 2013 Sixth International Conference on Business Intelligence and Financial Engineering (pp. 567-571). IEEE.

[12] Freitas, F. D., De Souza, A. F., & de Almeida, A. R. (2009). Prediction-based portfolio optimization model using neural networks. Neurocomputing, 72(10-12), 2155-2170.

[13] Heaton, J. B., Polson, N. G., & Witte, J. H. (2016). Deep learning in finance. arXiv preprint arXiv:1602.06561.

[14] Heaton, J. B., Polson, N. G., & Witte, J. H. (2016). Deep portfolio theory. arXiv preprint arXiv:1605.07230.

[15] Heaton, J. B., Polson, N. G., & Witte, J. H. (2017). Deep learning for finance: deep portfolios. Applied Stochastic Models in Business and Industry, 33(1), 3-12.

[16] Wang, W., Li, W., Zhang, N., & Liu, K. (2020). Portfolio formation with preselection using deep learning from long-term financial data. Expert Systems with Applications, 143, 113042.

[17] Lee, S. I., & Yoo, S. J. (2018). Threshold-based portfolio: the role of the threshold and its applications. The Journal of Supercomputing, 1-18.

[18] Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, 270(2), 654-669.

[19] Su, Z., Xie, H., & Han, L. (2020). Multi-Factor RFG-LSTM Algorithm for Stock Sequence Predicting. Computational Economics, 1-18.

[20] Fu, X., Du, J., Guo, Y., Liu, M., Dong, T., & Duan, X. (2018). A machine learning framework for stock selection. arXiv preprint arXiv:1806.01743.

[21] Yodmun, S., & Witayakiattilerd, W. (2016). Stock selection into portfolio by fuzzy quantitative analysis and fuzzy multicriteria decision making. Advances in Operations Research, 2016.

[22] Fabozzi, F.J.; Markowitz, H.M. The Theory and Practice of Investment Management: Asset Allocation, Valuation, Portfolio Construction, and Strategies, 2nd ed.; John Wiley and Sons: Hoboken, NJ, USA, 2011; Volume 198, pp. 289- -290.

[23] Adebiyi, A.A.; Adewumi, A.O.; Ayo, C.K. Comparison of ARIMA and artificial neural networks models for stock price prediction.]. Appl. Math. 2014.

[24] Cumming, J.; Alrajeh, D.D.; Dickens, L. An Investigation into the Use of Reinforcement Learning Techniques Within the Algorithmic Trading Domain. Master's Thesis, Imperial College London, London, UK, 2015.4.

[25] Chong, E.; Han, C.; Park, F.C. Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. Expert Syst. Appl. 2017, 83, 187-205.

[26] Fabozzi, F.J.; Pachamanova, D.A. Portfolio Construction, and Analytics; John Wiley Sons: Hoboken, NJ, USA, 2016; pp. 111-112.

[27] Kissell, R.L. The Science of Algorithmic Trading and Portfolio Management; Academic Press: Cambridge, MA, USA, 2013; pp.111-112.

[28] Ta, V.D.; Liu, C.M.; Addis, D. Prediction and Portfolio Optimization in Quantitative Trading Using Machine Learning Techniques. In Proceedings of the Ninth International Symposium on Information and Communication Technology, Da Nang, Vietnam, 6 -7 December 2018; pp.98- -105.

[29] Six Asset Allocation Strategies that Work. Available online: https://www.investopedia.com/investing/6- asset- allocation-strategies-work/ (accessed on 4 October 2019).

[30] Sharpe, W.F.; Sharpe, W.F. Portfolio Theory and Capital Markets; McGraw-Hill: New York, NY, USA, 1970; I Volume 217.

[31] Roll, R.; Ross, S.A. An empirical investigation of the arbitrage pricing theory. J. Financ. 1980, 35, 1073- -1103.

[32] Fama, E.F.; French, K.R. Common risk factors in the returns on stocks and bonds.]. Financ. Econ. 1993, 33, 35- -36.

[33] 13. He, G.; Litterman, R. The Intuition Behind Black- Litterman Model Portfolios; Goldman Sachs Investment Management Research: New York, NY, USA, 1999.

[34] Kolm, P.N.; Tutuncu, R.; Fabozzi, F.J. 60 Years of portfolio optimization: Practical challenges and current trends. Eur. J. Oper. Res. 2014, 234, 356- -371.

[35] Ahmadi-Javid, A.; Fallah-Tafti, M. Portfolio optimization with entropic value-at-risk. Eur. J. Oper. Res. 2019, 279, 225- -241.

[36] Lejeune, M.A.; Shen, S. Multi-objective probabilistically constrained programs with variable risk: Models for multi-portfolio financial optimization. Eur. J. Oper. Res. 2016, 252, 522- -539.

[37] Lwin, K.T.; Qu, R.; Mac Carthy, B.L. Mean-VaR portfolio optimization: A nonparametric approach. Eur.]. Oper. Res. 2017, 260, 751- -766.

[38] Qin, Z. Mean-variance model for portfolio optimization problem in the simultaneous presence of random and uncertain returns. Eur. J. Oper. Res. 2015, 245, 480- -488.

[39] Samarakoon, L.P.; Hasan, T. Portfolio performance evaluation. Encyclopedia of Finance, 2nd ed.; Springer: New York, NY, USA, 2006; pp.617- -622.

[40] Aragon, G.O.; Ferson, W.E. Portfolio performance evaluation. Found. Trends Financ. 2007, 2, 831- -890.

[41] Elleuch, J.; Trabelsi, L. Fundamental analysis strategy and the prediction of stock returns. Int. Res. J. Financ. Econ.2009, 30, 95- -107.

[42] Sezer, O.B.; Ozbayoglu, A.M.; Dogdu, E. An artificial neural network-based stock trading system using technical analysis and big data framework. In Proceedings of the South East Conference, Haines, AK, USA, 4- -12 April 2017; pp.223- -226.

[43] Fang, L.; Yu, H.; Huang, Y. The role of investor sentiment in the long term correlation between US stock and bond markets. Int. Rev. Econ. Financ. 2018, 58, 127-139.

[44] Nguyen, T.H.; Shirai, K.; Velcin, J. Sentiment analysis on social media for stock movement prediction. Expert Syst. Appl. 2015, 42, 9603- -9611.

[45] Lam, M. Neural network techniques for financial performance prediction: Integrating fundamental and technical analysis. Decis. Support Syst. 2004, 37, 567- -581.

[46] Deng, S.; Mitsubuchi, T.; Shioda, K.; Shimada, T.; Sakurai, A. Combining technical analysis with sentiment analysis for stock price prediction. In Proceedings of the 2011 IEEE Ninth International Conference on Dependable, Autonomic and Secure Computing, Sydney, Australia, 12- -14 December 2011; pp.800 -807.

[47] Sirignano, J.A. Deep learning for limit order books. Quant. Financ. 2019, 19, 549- -570.

[48] Sohangir, S.; Wang, D.; Pomeranets, A.; Khoshgoftaar, T.M. Big Data: Deep Learning for financial sentiment analysis. J. Big Data 2018, 5, 3.

[49] Xiong, R.; Nichols, E.P.; Shen, Y. Deep Learning Stock Volatility with .Google Domestic Trends. arXiv 2015,arXiv:1512.04916.

[50] Heaton, J.B.; Polson, N.G.; Witte, J.H. Deep learning for finance: Deep portfolios. Appl. Stoch. Models Bus. Ind.2017, 33, 3- -12.

[51] Chung, J., Gulcehre, C., Cho, K., Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv preprint arXiv:1412.3555.

[52] Nguyen, T. T., Yoon, S. (2019). A novel approach to short-term stock price movement prediction using transfer learning. Applied Sciences, 9(22), 4745.

[53] Chen, T., Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining (pp. 785-794).

[54] Chan Phooi M'ng, J., Mehralizadeh, M. (2016). Forecasting East Asian indices futures via a novel hybrid of wavelet-PCA denoising and artificial neural network models. PloS one, 11(6), e0156338.

[55] Adebiyi, A. A., Adewumi, A. O., Ayo, C. K. (2014). Comparison of ARIMA and artificial neural networks models for stock price prediction. Journal of Applied Mathematics, 2014.

[56] Chong, E., Han, C., Park, F. C. (2017). Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. Expert Systems with Applications, 83, 187-205.

[57] Ta, V. D., Liu, C. M., Tadesse, D. A. (2020). Portfolio optimization-based stock prediction using long-short term memory network in quantitative trading. Applied Sciences, 10(2), 437.