Master's Thesis

A Study on the Prediction of USD/KRW Exchange Rate Using MC(Multi-channel & Coupled)-LSTM

Dae Kyem Kim

The Graduate School

Sungkyunkwan University

Department of Applied Data Science

Master's Thesis

A Study on the Prediction of USD/KRW Exchange Rate Using MC(Multi-channel & Coupled)-LSTM

2021

Dae Kyem Kim

Master's Thesis

A Study on the Prediction of USD/KRW Exchange Rate Using MC(Multi-channel & Coupled)-LSTM

Dae Kyem Kim

The Graduate School

Sungkyunkwan University

Department of Applied Data Science

A Study on the Prediction of USD/KRW Exchange Rate Using MC(Multi-channel & Coupled)-LSTM

Dae Kyem Kim

A Master's Thesis Submitted to the Department of Applied Data Science and the Graduate School of Sungkyunkwan University in partial fulfillment of the requirements for the degree of Master of Engineering in Applied Data Science

April 2021

Supervised by

Jang Hyun Kim

Major Advisor

This certifies that the Master's Thesis of

Dae Kyem Kim is approved.

[signature]

Committee Chair :

[signature]

Committee Member :

[signature]

Major Advisor :

The Graduate School

Sungkyunkwan University

June 2021

Table of Contents

List of Tables

List of Figures

Abstract

A Study on the Prediction of USD/KRW Exchange Rate Using MC(Multi-channel & Coupled)-LSTM

The exchange rates are an essential variable that has a profound impact on trade between countries. Exchange risk management is crucial in small open economies, such as Korea, where export dependence is high and foreign exchange risk significantly impacts financial and real asset valuation. However, interest in the exchange rate has decreased a lot as the limitations of the existing financial time series model which was widely used to predict volatility in the financial and economic sectors have been revealed. However, as artificial intelligence technology develops, research using deep learning models that show strength in time series analysis is becoming more active. Nevertheless, research on exchange rates is still lacking. This work conducts exchange rate prediction by proposing XGBoost-LSTM, an ensemble LSTM model rarely utilized in existing studies. Furthermore, The author demonstrates the superiority of the proposed model by comparing the results with traditional machine learning classifiers and deep learning classifiers. The exchange rate forecasting model presented in this study used quoted exchange rates as input variables, and additionally, foreign exchange reserves, market rates (on a government bond basis), and current account balances. First, the exchange rate was predicted with

the regression model, and the trend of the exchange rate was predicted as rising, maintaining, and falling using the classification model. Empirical analysis of regression models and classification models has been performed using data from 30 years since 1990 and sliding window techniques.

Keywords : Exchange Rate Prediction, Machine Learning, Deep Learning, Long Short-Term Memory, XGBoost

Chapter 1. Introduction

One of the most important issues of recent academic circles due to the development of artificial intelligence technology is the study of predictability using various financial variables. Risk management arising from volatility in securities is an essential element of establishing investment strategies, such as stock prices, pricing derivatives, and evaluating the performance of funds. Various financial time series models have been proposed to predict future volatility, especially since financial variables have the characteristics of a time series. Among them are ARCH (Automatic Conditional Heteroscedasticity: Engle, 1982), GARCH (Generalized ARCH: Bolleslev, 1986), and JP MorganRiskMetrics (JP Morgan and Reuters, 1996). Financial time series models using metering techniques have the advantage of being able to provide systematic explanations based on statistical logic. However, the rapid changes in the market environment have resulted in much noise for most market variables, and there is a need for qualitative and quantitative variables to be considered in forecasting variability. Under these circumstances, many studies are being conducted on predicting financial time series using artificial intelligence techniques such as artificial neural networks, machine learning, and deep learning. Deep learning technologies such as Deep Neural Network (DNN) and Long Short-Term Memory (LSTM) show good performance in forecasting financial volatility (Dixon *et al.*, 2017; Jeong & Kim, 2019). However, most domestic financial volatility forecasting studies are centered on stock price forecasting studies (Shin *et al.*, 2017). In the situation where there is a relative lack of research on currency forecasting, this study seeks to conduct a study on predicting exchange rate variability.

Research using DNN and LSTM to predict the financial market has been brisk recently. Dixon et al., (2017) and LSTM Research (Fischer *et al*., 2018) for analyzing characteristics of financial time series in predicting financial markets are typical. Stock price prediction studies using RNN and LSTM were also carried out in Korea (Shin *et al*., 2017). Meanwhile, some studies have been conducted, but extremely rare exchange rate prediction studies using LSTM have been published. Cao (2020) conducted a USD/CNY exchange rate prediction study using deep LSTM, while Cocianu and Avramescu (2020) combined LSTM to the NARX(Nonlinear AutoRegressive with eXternal input) model to carry out a study demonstrating model excellence.

This study aims to build a prediction model for exchange rate(closing price) using DNN, LSTM, and Ensemble Deep Learning models based on LSTM to predict the exchange rate for the following day. The input variables are the market price, high price, low price, and closing price of the exchange rate, and the learning and prediction of the model were carried out using the sliding window technique. The DNN model consists of five nodes in the input layer, ten nodes in the three middle layers respectively, and one node in the output layer. The LSTM model sets the node and sequence length of the five input layers to 7. Each cell has ten hidden nodes, and three LSTM layers. The activation function of the two models was set up ReLU(Rectified Linear Unit) and the study was conducted by applying Adam Optimization Algorithm(AdamOptimizer) as an optimization function.

This study consists of six chapters : Chapter 2 examines the background of the research; Chapter 3 describes the research methods DNN and LSTM models; Chapter 4 conducts empirical analysis based on the proposed model; Chapter 5 reports the results of analysis; The last chapter describes the conclusion of empirical investigation and the utilization, implications, and limitations of this study.

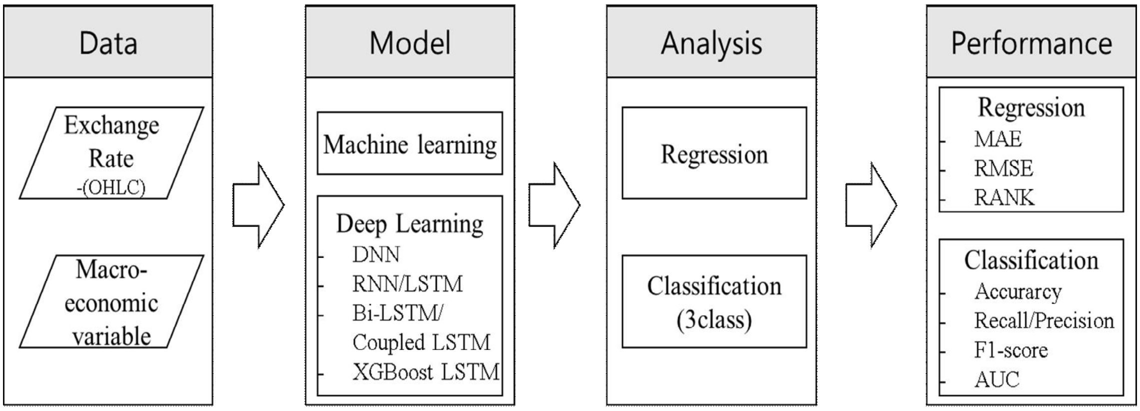

Figure 1. Structural Overview

Chapter 2. Related Works

2.1. Exchange Forecast

Exchange rates have a profound effect on trade between countries and are a critical financial variable. In particular, it affects the prices of various financial and real assets, including stocks, bonds and real estate, and it is a variable that affects the rise and fall of many financial variables. In particular, Korea is vulnerable to foreign exchange risk because it is a small open economy that is highly dependent on exports and is not enough to lead the international market. In these circumstances, research on exchange rates is essential, and research on currency forecasts is critical.

Traditional currency forecasts were made using a Econometrics model. One of the representative models of the Econometrics is the ARCH (Autoregressive Conditional Heteroskedasticity: Engle, 1982) model of Engle. However, the quantitative economic model revealed its limitations due to the volatility of various variables present in the market. Meese and Rogoff (1983) argued that it was pretty difficult to improve the Random Walk Model(Imamura and Sugita, 1980; O'Brien, 1993) above a certain level, taking into account macroeconomic fundamental factors. Since then, Meese and Rogoff (1987) have published studies showing that the statistical model of exchange rate forecasting cannot perform better than the random work model. There has been a 'negative perception' in the academic world for currency forecasting. Later, to overcome the limitations of random work models and Econometrics models, Clarida and Taylor (1997) used the period structure of forward exchange premium, vector error correction model to provide improved predictive power, and Jung (2004) conducted a USD/KRW prediction study in the same way. Kim (2000) conducted a study to predict exchange rates with statelevel input variables and monthly input variables, and had superior performance compared to the existing Econometrics model. However Varian (2014) argued for the need for a powerful tool to handle vast amounts of raw data. Because qualitative factors cannot be used as variables and the size of the available data is relatively small.

2.2. DNN & LSTM

DNN introduced by LeCun *et al*., (1989) was not widely used despite the excellent performance of algorithms due to the disadvantage of excessive time spent on neural network learning at that time (LeCun *et al*., 1989). However, as computing power is rapidly improving, DNN is drawing attention again. Deep Neural Network (DNN) is a model that uses more layers than conventional artificial neural networks by adding hidden layers between the input and output layers and adds characteristics to the data with specific representation and analysis capabilities (Hinton *et al*., 2006). DNN is easy to model nonlinear relationships and has features that are taught by error reversing algorithms (format 1). w indicates weight of each neuron, i is number of neurons in the layer. and j is the number where the neuron is located in the layer.

$$(1)\ \triangle w_{ij}(t+1) = \triangle w_{ij}(t) - \varepsilon\ \partial C/(\partial w_{ij}(t+1))$$

w : weight of each neuron, i : number of neurons in the layer. j : the number where the neuron is located in the layer

Furthermore, studies have recently been actively undertaken to predict stock prices based on models dealing with increased complexity, such as deep learning algorithms. Selvin *et al.,* (2017) predicted stock prices using recursive neural networks (RNNs), long-term memory (LSTM), convolutional neural networks (CNN), and yielded superior performance than traditional time series models such as ARIMA. Chou *et al*., (2018) proposed a MetaFA-LSSVR model based on the LSSVR (Last Squares Support Vector Regression) model, which further improves the complexity and efficiency of the support vector regression (SVR) model, and showed excellent predictive results. Huang (2018) improves prediction accuracy by extracting variables via autoencoder, and proposes a deep neural network model that shows improved prediction accuracy over a single conventional neural network. Research has been actively conducted in recent years to predict stock prices as deep learning algorithms have rapidly developed, enabling highly nonlinear modeling.

Bidirectional LSTM is a recurrent neural network structure that has recently been applied to various time series data processing fields and has shown the best performance. Bidirectional LSTM consists of forward LSTM and backward LSTM. Forward LSTM models the influence of past information on the present, and reverse LSTM models the influence of future information on the present.

BILSTM is known to show higher accuracy than general LSTM because it includes both forward and reverse layers at the same time. Due to these characteristics, BiLSTM is used for forecasting various time series financial variables.

2.3. XGBoost

The XGBoost model is one of the ensemble algorithms applying gradient boosting techniques to tree models that generate more robust classifiers by improving weak classifiers sequentially (Chen & Guestrin, 2016). Expression (2) represents an ensemble model of a tree, with K representing the number of trees and F representing the set of all possible classification and regression trees (CART). f_k corresponds to the weight of each independent tree and each leaf. The final prediction is made by summing and comparing the scores of each leaf.

$$(2)\quad \hat{y}_i = \sum_{k=1}^{K} f_k(x_i), f_k \in F$$

K: number of trees, F: set of the possible CART, f_k : ach independent tree and each leaf

Chen and Guestrin(2016) introduce extreme Gradient Boost (XGBoost) to solve overfitting problems in linear models or tree based models and improve the stability and training speed of datasets. XGBoost is a popular machine learning technique used by more than 20 of the 2015 Machine Learning Challenge winners in Kaggle (using the Chen & Guestrin, 'Boosting' sequential process to generate and utilize the following trees to capture strong dogs). XGBoost is a flexible model that can support these Boosting algorithm based models, classification rankings and regression and customized objectives (Chen & Guestrin, 2016). In this study, the author wants to perform exchange rate prediction by combining XGBoost, known as an excellent ensemble classifier, with an LSTM classifier that shows strength in time series prediction.

2.4. Sliding Window

This work uses sliding window techniques, a model that flexibly sets learning intervals and prediction intervals to verify the presented model's suitability and to confirm error rates. Sliding window techniques are characterized by creating input variables and accumulating input variables and output values (predicted values) by interval. In this study, the sliding window technique is applied to the exchange rate data from 2010 to 2019 to allow the initial five days of exchange rate data to predict the closing price of the exchange rate on the following day. The sliding window technique is a suitable technique for analyzing data with time series. The data for learning periods are not utilized as one-off, data for years 1 to 5 are used for predictions for year 6 and data for years 2 to 5 are used for predictions for year 7 to minimize loss of input data and maintain the time series of input variables. Therefore, sliding window techniques are known as useful methods for financial time series analysis and prediction (Jang *et al.*, 1993; Cheong & Oh, 2014).

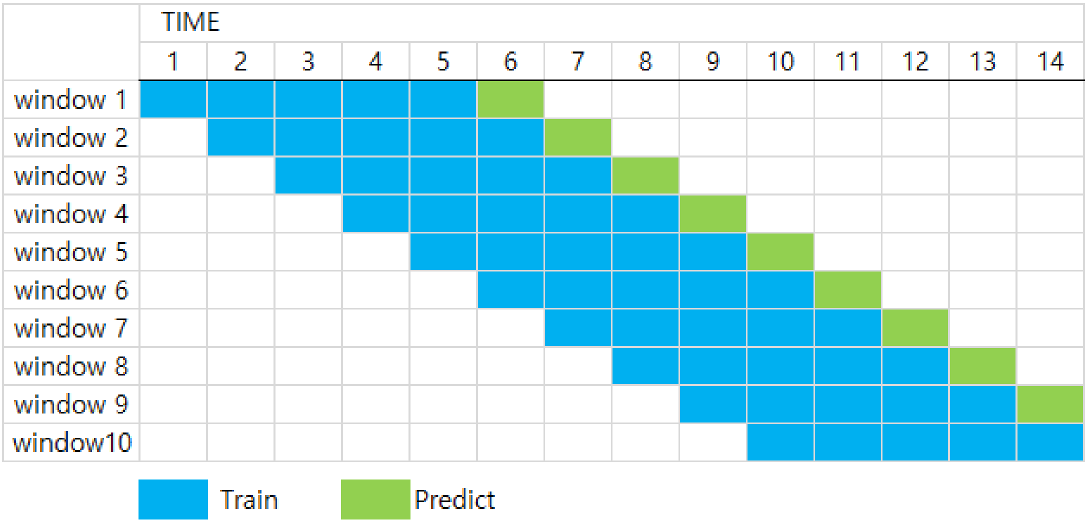

Figure 2. Sliding Window Techniques

This work aims to conduct a high-level exchange rate prediction study by applying a sliding window to secure enough learning periods with five years of learning and one year of experiments. The expected effect of sliding window techniques is believed to overcome the limitations of relatively small exchange rate data size.

Chapter 3. Research Questions

Due to the development of AI technology, recent research on predicting financial variables using AI has been actively conducted in academia. The DNN and LSTM techniques used in this work have been widely used. DNNs have been used in classification based financial market prediction studies (Dixon *et al.*, 2017) . In addition, LSTM is widely used when the time series characteristics of financial variables are salient. Fischer *et al.*, (2018) for the analysis of financial time series characteristics are a representative study. In addition, Cao (2020) conducted a USD/CNY exchange rate prediction study using deep LSTM, while Cocianu and Avramescu (2020) conducted a study that demonstrates the superiority of the model by combining LSTM with the NARX model. Influenced by these studies, financial variable prediction studies using deep learning techniques have also been conducted in Korea, such as representatively stock price prediction studies using LSTM and cyclic neural network techniques (Shin *et al.*, 2017).

However, existing research focused on stock price forecasting and rare exchange rate forecasting studies using LSTM have been conducted. In addition, there are a number of studies that have shown that LSTM outperforms traditional deep neural networks in stock price prediction, but the studies that improve the performance of prediction models in combination with other ensemble algorithms are rare. AI technologies are proved useful in predicting and classifying financial variables, and currency forecasting research is clearly important in the export and import oriented Korean financial environment. However, it is true that Korean exchange rate forecasting research has not been conducted much and that new methods have not been applied enough. In particular, the rapid growth of data, algorithms, and computing power as the three main components of AI activation has led to the enhanced perception of AI techniques in the field of financial variable analysis. Still, rare research has focused on exchange rate prediction. Therefore, this work seeks to conduct a KRW/USD exchange rate prediction study focusing on LSTM and XGBoost, deep learning techniques applied with AI technology. To this end, we propose two Research Questions :

RQ1: Is there a significant difference between deep learning models and general machine learning models in predicting exchange rates?

RQ2: Does XGBoost-LSTM, the deep learning ensemble technique proposed in this work, perform better than simple machine learning or deep learning methods?

Chapter 4. Method

The LSTM (Long Short-Term Memory) is a particular type of circular neural network model proposed to overcome the long-term dependency problem of the RNN (Recurrent Neural Network). The RNN uses hyperbolic tangent (Tanh), an activation function that converts input values into nonlinear relationships, to produce output. Because the output value is determined between -1 and 1, the number becomes smaller as the operation is repeated. Repetition of these procedures will result in a "Gradient Vanishing Problem" in which past learning results are lost. When a dependence problem occurs, there is a problem that prevents learning from a certain stage. To overcome past weight problems, LSTM uses cells with small memory inside the neural network. The cell consists of the input gate, the forget gate, and the output gate, and determines the input, output and the forget gate. As this process repeats, the slope of the past extinguishes and there is nothing impossible to learn. LSTM basically calculates the output value in the same way as the circulatory neural network, but appropriately utilizes the input, forget and output gate in the cell of the concealed layer to regulate the information flow. As a result, circulatory neural networks using LSTM cells do not cause slope loss, even for data that require long-term time series learning, such as financial time series data.

This study was conducted from January 1, 1990 to April 30, 2021, utilizing data on the market price, high price, low price, closing price, and trading volume of the daily USD/KRW exchange rate. Thirty years of exchange rate data were secured to learn long-term memory models. The USD/KRW exchange rate data for a total of 9452 trading days were used, and the ratio of the data used for learning and forecasting was set to 8:2 to conduct the study.

The following is the exchange rate data layout used in the study.

Figure 3. Raw Data Layout (Exchange Rate)

Macroeconomic variables used in this study are foreign exchange reserves, current account balances, and interest rates on government bonds. For foreign exchange reserves, vast foreign exchange reserves including dollars, gold, and special drawing rights were used. For the current account, the current account data published by the Bank of Korea was used. Since foreign exchange reserves and current account are monthly data, preprocessing was performed by learning the two models separately from the deep learning model and merging them. Next, six types of KTB(Korean Treasury-Bond futures) interest rates were used. The six types of KTB are 1-year, 3-year, 5-year, 10-year, 20-year, and 30-year KTBs. The average of KTB interest rates has been applied since each KTB existed.

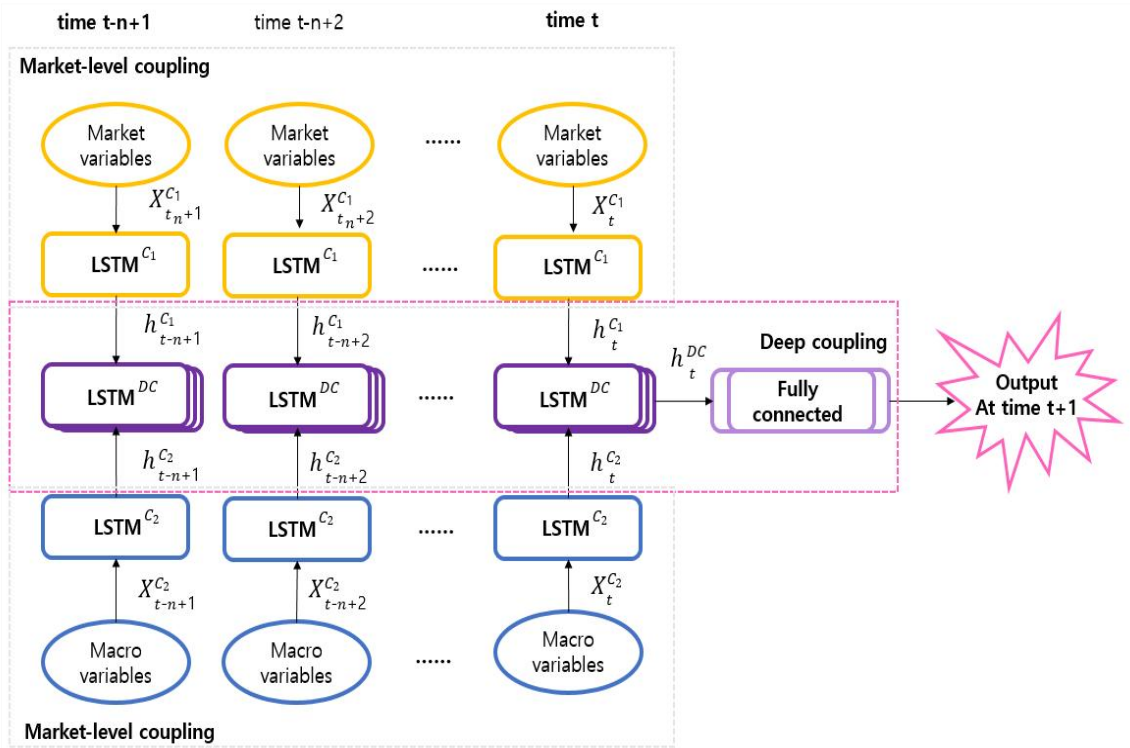

Figure 4. Architecture of the DC-LSTM approach (Compiled from [25] Cao, W *et al.*, 2020)

The RMSE (Root Mean Squared Error) and MAE(Mean Abosloute Error) were used. RMSE is a measure used to address the difference between the actual and predicted values of the data. It is particularly suitable for expressing the precision and accuracy of the model (formula 3). MAE is a method of averaging them by applying absolute values to the error of real and predicted values (formula 4)

$$(3)\ \ \mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(real_t - predict_t)^2}$$

$$(4)\ \ \mathrm{MAE} = \frac{1}{n}\sum_{t=1}^{n}\left|(real_t - predict_t)\right|$$

4.1. DNN Model

The DNN model consists of 20 nodes of the input layer, 10 nodes of the middle layer, and 1 node of the output layer. Four input nodes were used to match the number of input data (open price, high price, low price and closing price), and the number of intermediate layers (hidden layers) was determined using trial and error. Adam optimization algorithm, which has a fast margin of error, was used as an optimizer and was designed as a model to predict the closing price of tomorrow by utilizing the variables of the day. Finally, when the target variable is set to a regression model, we apply the sigmoid function as an activation function. In addition, when the target variable was set to the exchange rate rise, fall, or maintain, the softmax function was used in the classification model.

Table 1. DNN Model

Layout	Structure
Input Layer	20 nodes
Hidden Layer	10nodes, 3 Layer
Output Layer	1 node
Activation Function	ReLu(Rectified Linear Unit)
Optimizer	Adam Optimizer

Output Value	1. Regression model: Exchange rate 2. Classification model: Exchange rate Trend (rise, maintain, fall)

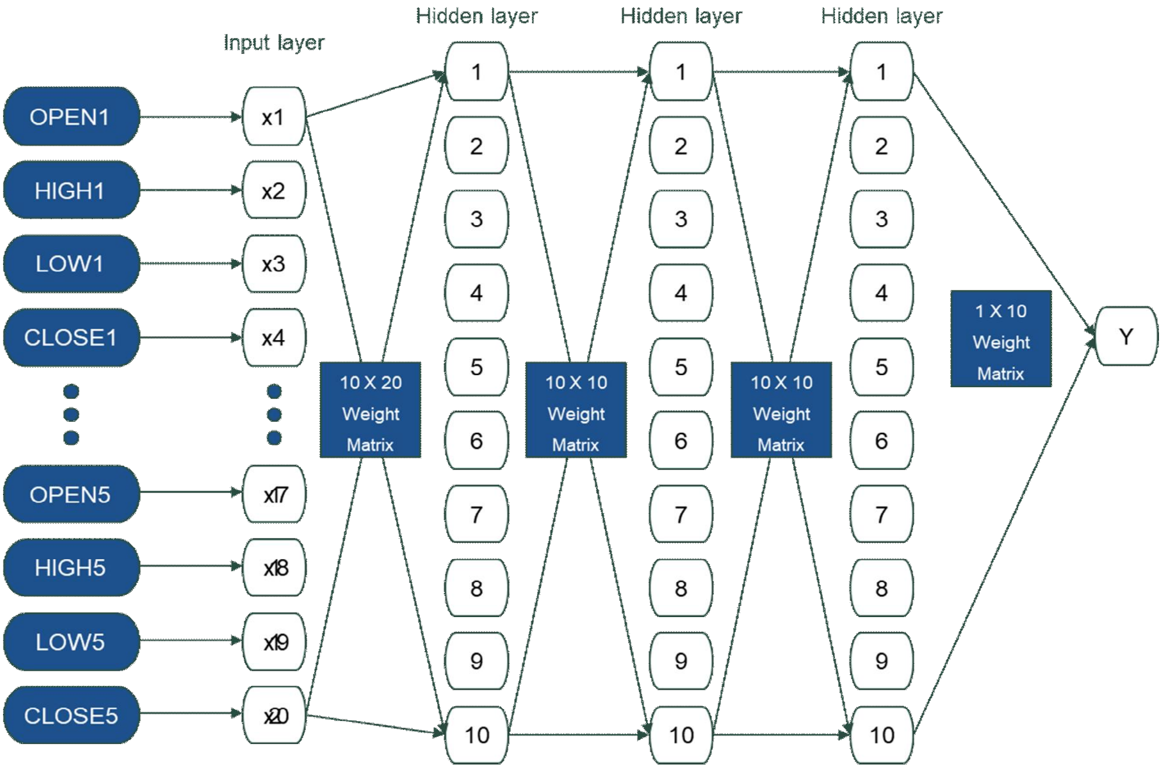

Figure 5. DNN Layout

4.2. RNN & LSTM Model

The RNN model estimated the exchange rate for the next market day by using 20 price forecasting factors at a time by building five cells on the input node from Monday to Friday. Also, The LSTM model estimated the exchange rate for the next market day by using 20 forecast factors at a time by building five cells on the input nodes Monday through Friday. Likewise, The RNN and LSTM models applied the sigmoid function as an activation function in the regression model. And the softmax function was used in the classification model.

Table 2. RNN Model

Layout	Structure
Input Layer	(5,4) array
Hidden Layer	Single Layer RNN
Output Layer	1 node
Activation Function	Tanh
Optimizer	Adam Optimizer
Output Value	1. Regression model: Exchange rate 2. Classification model: Exchange rate Trend (rise, maintain, fall)

Since the RNN model is a deep learning model with a cyclic structure, it is known that the time series prediction performance is superior to that of the existing DNN model. As mentioned above, one RNN layer contains 512 neurons, and the Tanh activation function is applied to each RNN cell. The RNN model first produces a regression result, which means the exchange rate itself. Next, the classification results show the trend of exchange rates (Figure 6).

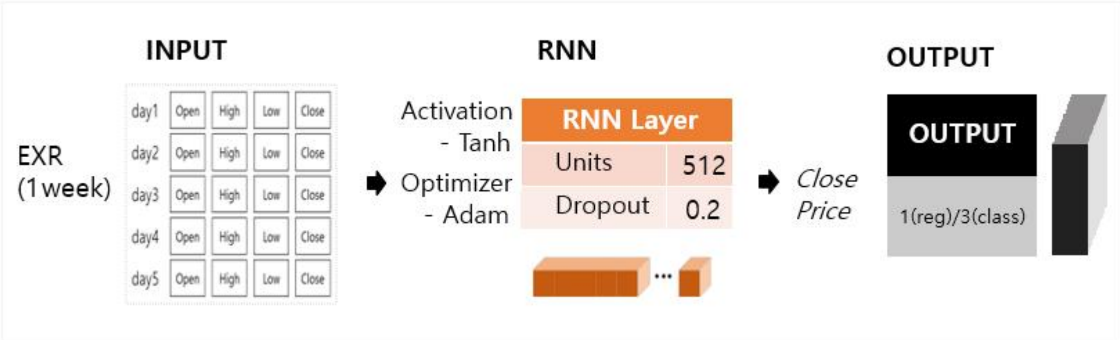

Figure 6. RNN Layout

LSTM is an advanced model of RNN, and has the characteristic of managing input data in long-term memory gates and short-term memory gates. There are already many studies on the time series analysis performance of LSTM. Three types of activation functions were applied for performance comparison according to LSTM and each activation function. The LSTM model also produces a regression result first, which means the exchange rate itself. Next, we will present the classification results showing the trend of exchange rates.(Figure 7).

Table 3. LSTM Model

Layout	Structure
Input Layer	(5,4) array
Hidden Layer	Single Layer LSTM
Output Layer	1 node
Activation Function	Tanh / ReLU /Sigmoid
Optimizer	Adam Optimizer
Output Value	1. Regression model: Exchange rate 2. Classification model: Exchange rate Trend (rise, maintain, fall)

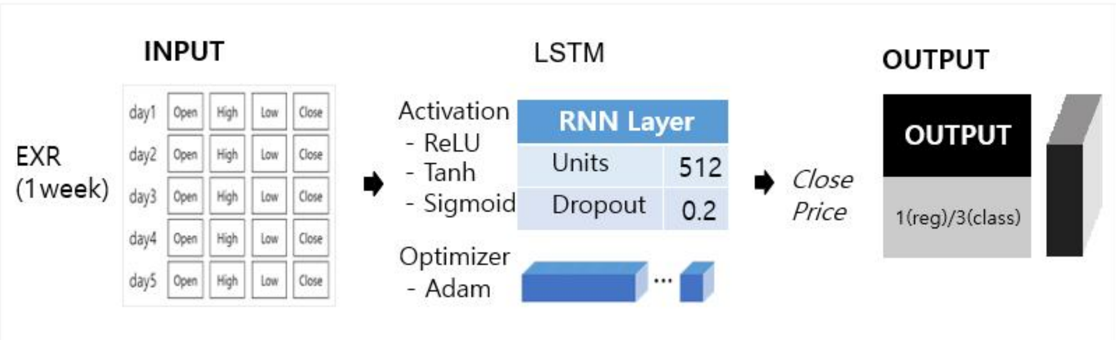

Figure 7. LSTM Layout

4.3. Bidirectional LSTM Model

The bidirectional LSTM model is a model in which a reverse algorithm is added by supplementing the unidirectional LSTM model. The interactive LSTM model is widely used for time series analysis and text analysis. In this study, a two-way LSTM layer and one LSTM layer were used to build a two-way LSTM model. As for the activation function, ReLU is applied, and like other deep learning models, it produces two types of outputs.

Table 4. Bi-LSTM Model

Layout	Structure
Input Layer	(5,4) array
Hidden Layer	Double Layer Bi-LSTM + Single Layer LSTM
Output Layer	1 node
Activation Function	ReLU
Optimizer	Adam Optimizer
Output Value	1. Regression model: Exchange rate 2. Classification model: Exchange rate Trend (rise, maintain, fall)

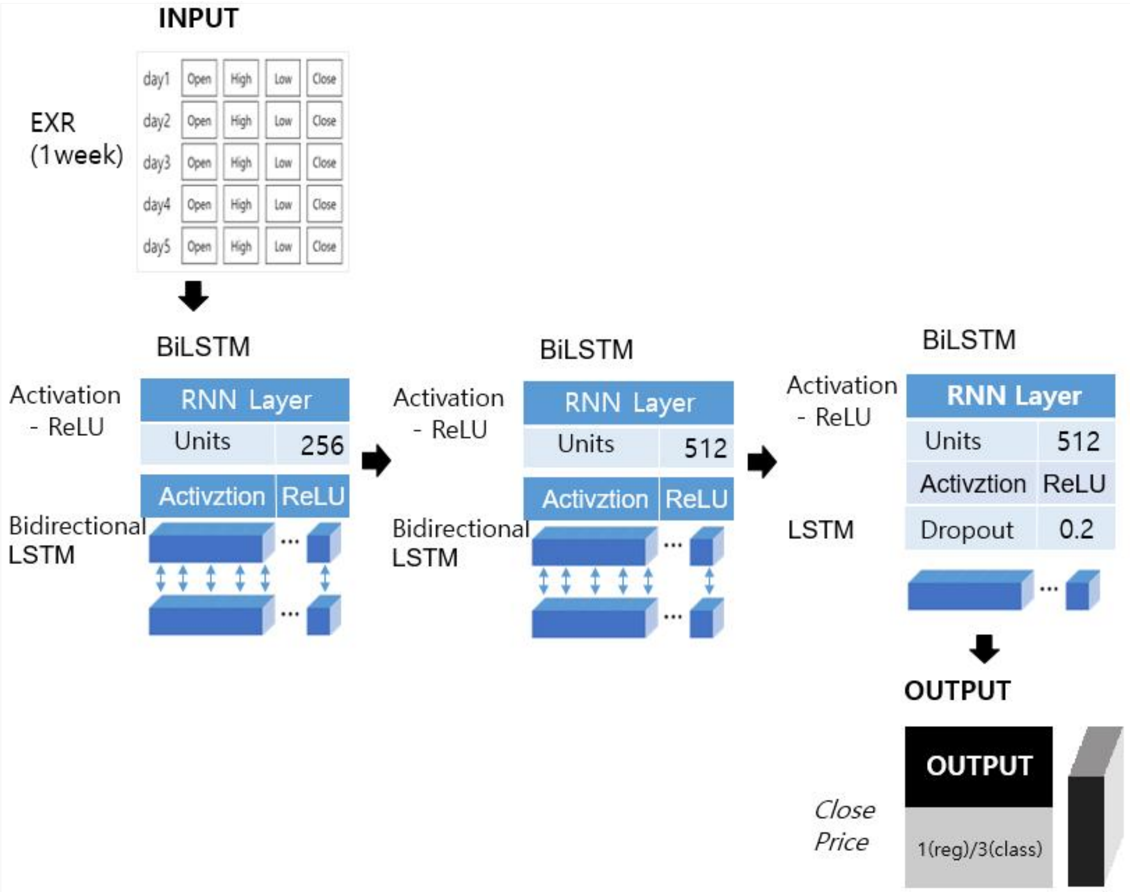

Figure 8. Bi-LSTM Layout

4.4. Multi-channel LSTM Model

This LSTM model consists of putting inputs into each of the two LSTM layers and combining them into a single model using a concatenate layer. One LSTM model learns weekly data. Therefore, it has input values in the form of (5,4) consisting of five trading days. Next, other LSTM models have four weeks of exchange rate data as input values for monthly learning.

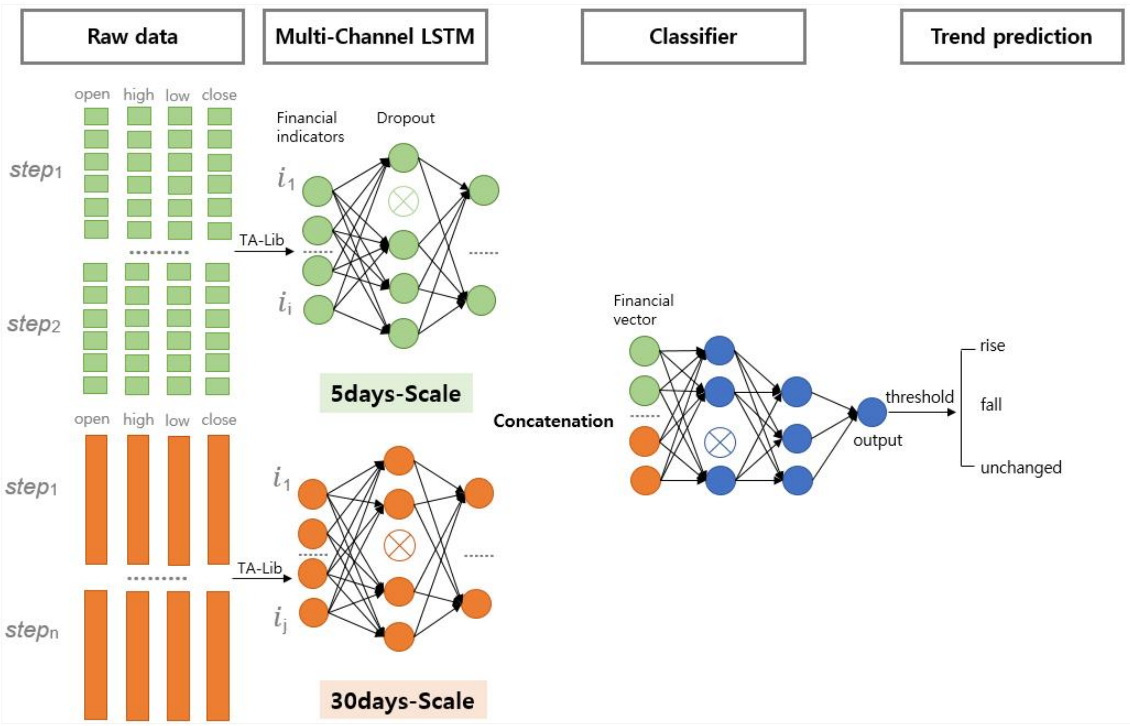

Figure 9. Architecture of the Multi-channel LSTM Approach
(Compiled from [26] Wei, W., & Li, P. 2019)

Thus, the cell of the final input layer consists of (20,4). Later, the activation function was ReLU, with a total of three layers of LSTM used as hidden layers. Outputs were used just like the above written deep learning models.

Table 5. Multi-channel LSTM Model

Layout	Structure
Input Layer	Exchange rate (EXR): (5,4) array + (20,4) array
Hidden Layer	Triple Layer LSTM
Output Layer	1 node
Activation Function	ReLU
Optimizer	Adam Optimizer
Output Value	1. Regression model: Exchange rate 2. Classification model: Exchange rate Trend (rise, maintain, fall)

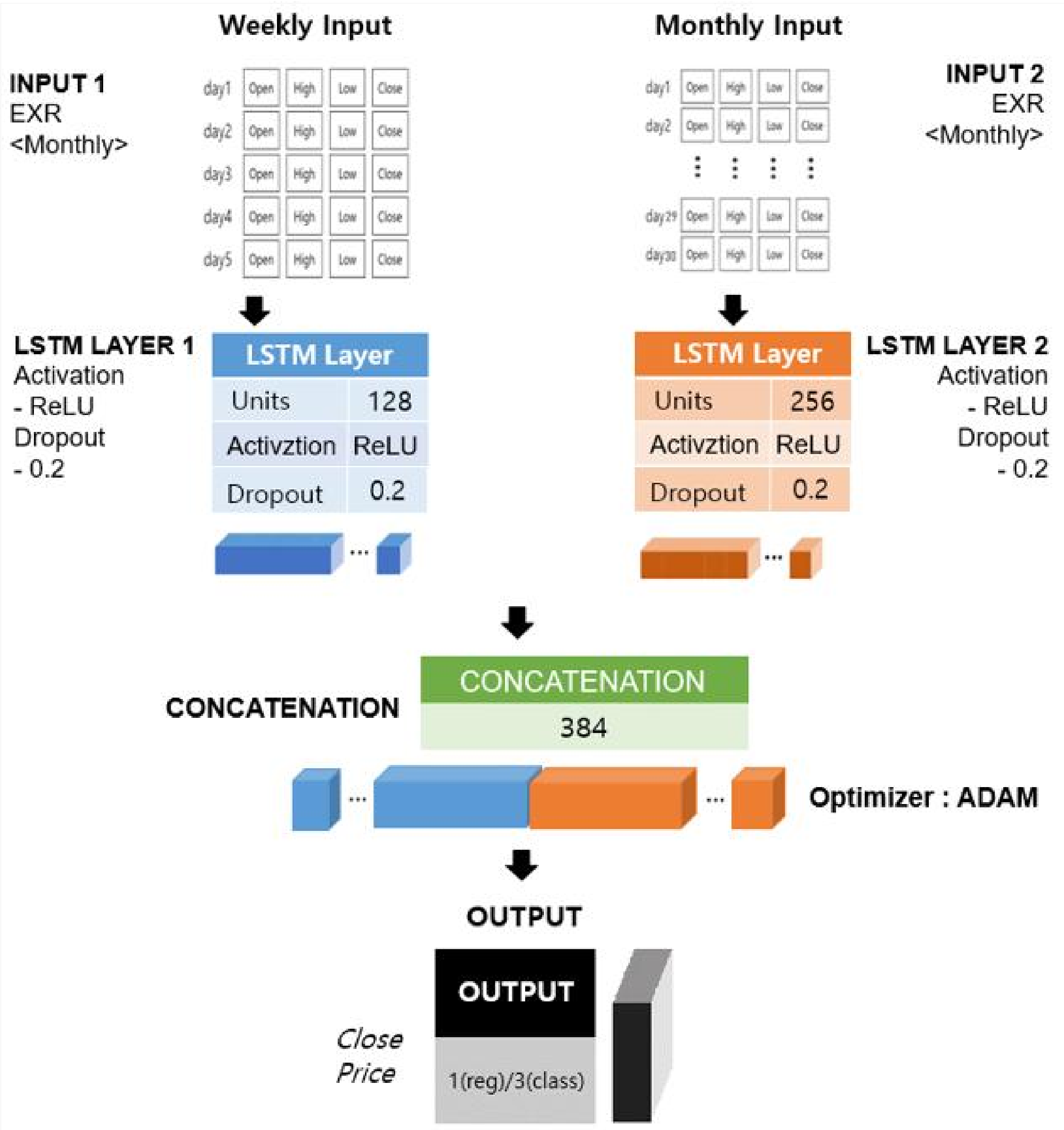

Figure 10. Multi-channel LSTM Layout

4.5. Ensemble XGBoost-LSTM Model

Finally, the XGBoost-LSTM model utilizes two types of exchange rate data borrowed from Multi-channel LSTM and additional macroeconomic variables as input variables. The macroeconomic variables used were current account, foreign exchange reserves, and market interest rates based on government bonds. Before injecting macroeconomic variables, the VIF was verified to be free from multicollinearity problems, and the analysis was conducted after confirming that there were no such problems. VIF with exchange rate and bond interest rate is 1.259043. VIF with exchange rate and current account is 2.881084. VIF with exchange rate and foreign exchange reserves is 2.933990.

The XGboost combined model utilizes macroeconomic(weekly and monthly) exchange rates as input data from the three LSTM models, respectively. Among macroeconomic data, which are not provided as a set, were transformed to a set and applied to the model.

Table 6. XGBoost-LSTM Model

Layout	Structure
Input Layer	Exchange rate (EXR): (5,4) array + (20,4) array Exchange rate (EXR):
Hidden Layer	Triple Layer LSTM
Output Layer	XGBoost
Activation Function	ReLU

Optimizer	Adam Optimizer
Output Value	1. Regression model: Exchange rate 2. Classification model: Exchange rate Trend (rise, maintain, fall)

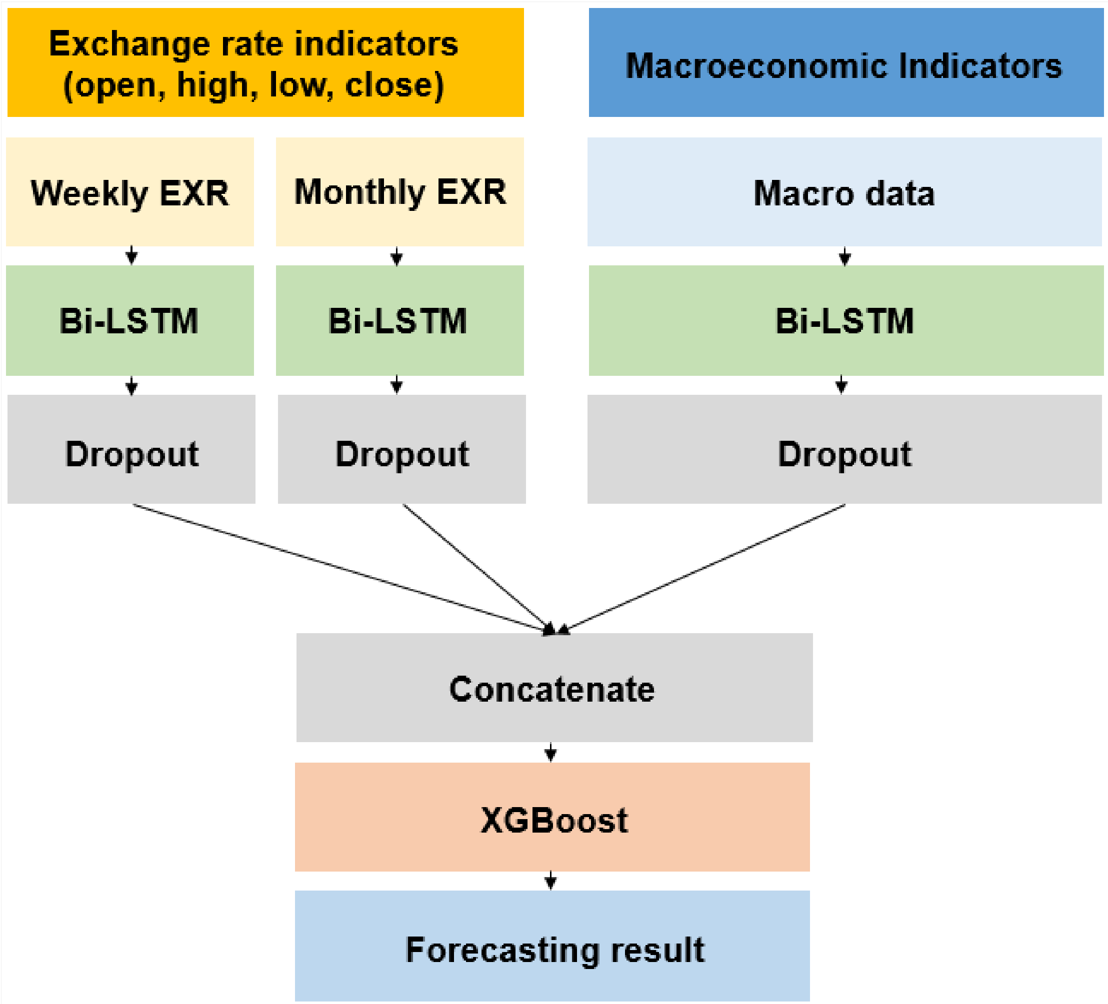

Figure 11. Proposed XGBoost-LSTM Layout

Chapter 5. Results

5.1. Regression Results

This study conducted an analysis by setting the total learning period (1990/03/01 to 2021/04/30), of which 80 percent(6,162 days) as learning data and 20 percent(1,540 days) data as experimental data. The forecasts were made using the machine learning algorithms (Logistic regression, SVM, Random Forest), Deep Learning algorithms(DNN, RNN, LSTM) and Ensemble models. The average square root error (RMSE) was measured by comparing the estimates for each model with actual data.

First of all, machine learning algorithms that were used as benchmarking models showed differences in performance with deep learning classifiers. Among the machine learning algorithms, the best regression performance was SVM. The RMSE of the SVM is the smallest, 0.0104.

Next, among deep learning algorithms, the LSTM model combined with ReLu function performs best. In particular, it is impressive that the difference among regular DNN, RNN, and LSTM is noticeable. These results are due to the time series nature of exchange rates, and to it's long learning period. There is a possibility that the DNN and RNN models have a gradient vanishing problem.

Table 7. Regression Performance

Category	Classifier	RMSE	MAE	RANK
Machine Learning	Logistic Regression	0.0107	0.0064	9
	Support Vector Machine	0.0104	0.0061	7
	Random Foreset	0.0114	0.0065	10
Deep Learning	DNN	0.0099	0.0063	8
	RNN(Tanh)	0.0067	0.0054	5
	LSTM(Sigmoid)	0.0068	0.0053	6
	LSTM(Tanh)	0.0062	0.0052	4
	LSTM(ReLu)	0.0061	0.0049	3
Ensemble Deep Learning	Multi-channel LSTM	0.0060	0.0045	2
	XGBoost-LSTM	0.0055	0.0042	1

Moreover, the results are further improved in ensemble models. In particular, the error was reduced in the last model combined with XGboost. The XGBoost-LSTM model has the lowest error rate of the total comparison target model, suggesting that the combination of the two algorithms has synergy with each other.

The results of the MAE were also found to be quite similar to RMSE. The most miniature error model was found to be XGBoost-LSTM, an ensemble deep learning classifier. The smallest of the deep learning classifiers is the deep learning model with the ReLu function, and the best predictive classifier among the machine learning classifiers is SVM.

5.2. Classification Results

Table 8 shows the predictive performance of classifiers. First of all, support vector machines (SVMs) perform the best among machine learning classifiers. The prediction rate of SVM was 42.76% and the F1 score was 41.74. Next, the deep learning classifier, LSTM, appears to show 52.75% accuracy and F1 score 52, demonstrating a clearly advanced performance over conventional machine learning classifiers. The Multi-channel LSTM model, an ensemble classifier, improved accuracy by 54.54 percent and the F1 score by 1.5 points. Notably, the ensemble LSTM model combined with XGboost also improves classification performance. However, although Precision is down 0.8 points compared to Multi-channel LSTM, the author observed that the model's overall performance has improved with remarkable improvement in Recall.

At the bottom of the table, the model with macroeconomic variables is shown. The input of these variables demonstrates an improvement in classification performance compared to the use of existing exchange rate data only. Through this, the author shows that time series changes in macroeconomic variables can contribute to improving the predictive performance of the model. Unusually, however, the results of LSTM compared with Multi-channel LSTM have been shown to be even better. These results are likely to have caused overfitting between modeling, or the learning of macroeconomic variables rather led to bias. This point needs to be reviewed closely later. Next, the best predictive performance across all models is observed in the XGBoost-LSTM model. Similarly, however, F1 scores and Precision are observed to be lower than the

LSTM model, which requires further review. The highest AUC model is the XGBoost-LSTM model. The figure is 0.77, which is as high as 0.10 compared to the most miniature LR model. As AUC is a number that reflects sensitivity and specificity, it is easy to observe how consistent the model's performance is.

Table 8. Classification Performance

DATA	Metric	Logistic Regression	Support Vector Machine	Random Foreset	LSTM	Multi-channel LSTM	XGBoost-LSTM
EXR	Accuracy	42.68	42.76	41.75	52.75	54.54	55.32
	F1	41.44	41.74	40.61	52.00	53.50	55.14
	Precision	51.20	48.20	48.07	57	59.8	59.00
	Recall	34.80	36.80	35.15	47.8	48.4	53.40
	AUC	0.67	0.68	0.65	0.72	0.74	0.75
EXR +MACRO	Accuracy				57.53	57.17	57.70
	F1				56.57	56.31	56.51
	Precision				62.60	61.40	60.00
	Recall				51.60	52.00	53.40
	AUC				0.76	0.76	0.77

The predictive graph (fig.12) analysis shows that XGBoost-LSTM has a higher visual prediction rate than DNN/RNN. In particular, the error was even greater when the graphs in the early parts of the forecast were checked. XGBoost-LSTM model performs close predictions with actual observations. The model also performs high level predictions with the smallest error (RMSE) and with actual data. The RMSE analysis showed that among XGBoost-LSTM models, models using the Relu activation function were excellent.

Figure 12. DNN/RNN(Above), XGBOOST-LSTM(Below) Predicted Data

Chapter 6. Discussion and Conclusion

This study conducted an exchange rate prediction study using LSTM and ensemble-LSTM among artificial neural network techniques. In particular, the model was trained by segmenting sections with sliding window techniques. Deep learning models showed better predictive performance than machine learning models, while ensemble deep learning models showed better performance than deep learning models. LSTM is known to have strengths in time series analysis and has shown good performance in practice. In addition to the exchange rate data, the improvement in prediction performance was observed by adding macroeconomic data with time series features, and the exchange rate prediction performance was further improved. Therefore, the research question supported that the model's performance would be improved if indicators reflecting market volatility in addition to technical indicators such as market prices, low prices, high prices, and closing prices were included, to improve the predictive performance of exchange rate data.

Furthermore, this study has a theoretical contribution in that it analyzed the model by combining ensemble classifiers. In particular, coupling LSTM and XGBoost-LSTM models can be further expected to improve quality of prediction Next, existing studies mainly dealt with exchange rate variables. In this study, there is also a contribution in the model's predictive performance by dealing with macroeconomic variables and exchange rate variables together.

However, it is meaningful that this study was the exchange rate prediction study that applies the enhanced LSTM model. And this study proved that various quantitative and qualitative variables reflecting exchange rate volatility should be considered to increase the accuracy of the exchange rate prediction. The Korea exchange rate system has been changed to a free floating exchange rate system since the 1997 IMF. As a result, the absolute amount of accumulated exchange rate data is insufficient. In 2008, 10 years after the financial crisis (IMF), the subprime mortgage crisis again ignited a big change to the foreign exchange market. In this study, there is a limitation that the number of samples is insufficient because daily data since 1990 have been utilized as samples to minimize uncertainty and noise. There is a possibility that this instability may have affected the exchange rate forecast because the international situation is chaotic in the aftermath of Corona which began in 2020. In addition, there is a limitation that a rather simple structure model was used because the model was designed using only technical indicators of exchange rates. Therefore, future studies should include further consideration of quantitative and qualitative variables in exchange rate prediction. Furthermore, the author believes that the research can be conducted by applying state-of-the-art artificial intelligence techniques by extending the model of the present simple structure by obtaining more variables.

References

[1] Ahn. J. J., Byun, H. W., Oh, K. J. and Kim, T. Y. (2012). Using ridge regression with genetic algorithm to enhance real estate appraisal forecasting. Expert Systems with Applications, 39, 8369-8379.

[2] Alahi, A., Goel, K., Ramanathan, V., Robicquet, A., Fei-Fei, L., & Savarese, S. (2016). Social lstm: Human trajectory prediction in crowded spaces. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 961-971).

[3] Lee, C. I., Chang, C. H., & Hwang, F. N. (2019, November). Currency Exchange Rate Prediction with Long Short-Term Memory Networks Based on Attention and News Sentiment Analysis. In 2019 International Conference on Technologies and Applications of Artificial Intelligence (TAAI) (pp. 1-6). IEEE.

[4] Chen, K., Zhou, Y., & Dai, F. (2015, October). A LSTM-based method for stock returns prediction: A case study of China stock market. In 2015 IEEE international conference on big data (big data) (pp. 2823-2824). IEEE.

[5] Cocianu, C., & Avramescu, M. (2020). The use of LSTM neural networks to implement the NARX model. A case study of EUR-USD exchange rates. Informatica Economica, 24(1), 5-14.

[6] Diebold, F. X., & Nason, J. A. (1990). Nonparametric exchange rate prediction?. Journal of international Economics, 28(3-4), 315-332.

[7] Dixon, M., Klabjan, D., & Bang, J. H. (2017). Classification based financial markets prediction using deep neural networks. Algorithmic Finance, 6(3-4), 67-77.

[8] 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.

[9] Fu, R., Zhang, Z., & Li, L. (2016, November). Using LSTM and GRU neural

network methods for traffic flow prediction. In 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC) (pp. 324-328). IEEE.

[10] Galeshchuk, S. (2016). Neural networks performance in exchange rate prediction. Neurocomputing, 172, 446-452.

[11] Gers, F. A., Schmidhuber, J., & Cummins, F. (1999). Learning to forget: Continual prediction with LSTM.

[12] Gers, F. A., Schmidhuber, J., & Cummins, F. (1999). Learning to forget: Continual prediction with LSTM.

[13] Gers, F. A., Schraudolph, N. N., & Schmidhuber, J. (2002). Learning precise timing with LSTM recurrent networks. Journal of machine learning research, 3(Aug), 115-143.

[14] Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., & Schmidhuber, J. (2016). LSTM: A search space odyssey. IEEE transactions on neural networks and learning systems, 28(10), 2222-2232.

[15] Huang, Z., Xu, W., & Yu, K. (2015). Bidirectional LSTM CRF models for sequence tagging. arXiv preprint arXiv:1508.01991.

[16] Ince, H., & Trafalis, T. B. (2006). A hybrid model for exchange rate prediction. Decision Support Systems, 42(2), 1054-1062.

[17] Jeong, G., & Kim, H. Y. (2019). Improving financial trading decisions using deep Q-learning: Predicting the number of shares, action strategies, and transfer learning. Expert Systems with Applications, 117, 125-138.

[18] Jeong. K. H. (2017). Quantile causality from dollar exchange rate to international oil price. Journal of the Korean Data & Information Science Society, 28, 361-369.

[19] Kim, S. H. (2000). Establishment of Optimal Artificial Neural Network Model and Exchange Rate Prediction Performance Analysis. Journal of Money and Finance, 14, 57-85.

[20] Kim, T. Y., Oh, K. J., Sohn, I. and Hwang, C. (2004). Usefulness of artificial neural networks for early warning ystem of economic crisis. Expert Systems with Applications, 26, 583-590.

[21] Lee, J. Y. and Kim, H. J. (2014). Identification of major risk factors association with respiratory diseases by data mining. Journal of the Korean Data & Information Science Society, 25, 373-384.
[22] Meese. R. A. and Rogoff. K. (1983). Empirical exchange rate models of the seventies: Do they fit out of sample?. Journal of Imternational Economics, 14, 3-24.
[23] Refenes, A. N., Azema-Barac, M., Chen, L., & Karoussos, S. A. (1993). Currency exchange rate prediction and neural network design strategies. Neural Computing & Applications, 1(1), 46-58.
[24] Shin, Y. G. (2009). Study on the causality between call rate and exchange rate under global economic crisis. Journal of the Korean Data and Information Science Society, 20(4), 655-660.
[25] Cao, W., Zhu, W., Wang, W., Demazeau, Y., & Zhang, C. (2020). A deep coupled LSTM approach for USD/CNY exchange rate forecasting. IEEE Intelligent Systems, 35(2), 43-53.
[26] Wei, W., & Li, P. (2019, November). Multi-channel LSTM with different time scales for foreign exchange rate prediction. In Proceedings of the international conference on Advanced Information Science and System (pp. 1-7).
[27] Weigend, A. S., Rumelhart, D. E., & Huberman, B. A. (1991, November). Generalization by weight elimination applied to currency exchange rate prediction. In [Proceedings] 1991 IEEE International Joint Conference on Neural Networks (pp. 2374-2379). IEEE.
[28] Zhang, G. P., & Berardi, V. L. (2001). Time series forecasting with neural network ensembles: an application for exchange rate prediction. Journal of the operational research society, 52(6), 652-664.'

논문요약

MC(Multi-Channel & Coupled)-LSTM를 이용한 원/달러 환율 예측에 관한 연구

김대겸
데이터사이언스융합학과
성균관대학교

환율은 국가 간 무역에 지대한 영향을 미치는 중요한 변수이다. 한국과 같이 수출의존도가 높고, 특히 외환 위험이 금융 및 실물자산 평가에도 상당한 영향을 미치는 소규모 개방경제에서는 환율 리스크 관리가 매우 중요하다. 그러나 금융·경제 분야 변동성 예측에 널리 활용됐던 기존 금융 시계열 모델의 한계가 드러나면서 환율에 대한 관심은 많이 줄어들었다. 하지만 인공지능 기술이 발전하면서 시계열 분석에 강점을 보이는 딥러닝 모델을 활용한 연구가 활발해지고 있다. 그럼에도, 환율에 대한 연구는 여전히 부족하다. 본 연구는 기존 국내 연구에 드물게 이용되는 앙상블 LSTM 모형인 XGBoost-LSTM을 제안하여 환율 예측 연구를 수행한다. 나아가 결과를 일반 머신러닝 분류기, 딥러닝 분류기와 비교하여 제안한 모델의 우수성을 입증한다. 본 연구에서 제시된 환율 예측 모델은 환율을 입력 변수로 사용했으며 추가적으로 외환보유고, 시장금리(국채기준)와 경상수지가 사용되었다. 먼저, 회귀 모형으로 환율을 예측하고, 분류 모형을 사용하여 환율의 추세를 상승, 유지, 하락으로 예측하였다. 회귀 모형과 분류 모형의 실증 분석은 1990년 이후 30년간의 데이터와 슬라이딩 윈도우 기법을 사용하여 수행되었다.

키워드: 환율 예측, 머신러닝, 딥러닝, 장단기 메모리(LSTM), XGBoost