"POLITEHNICA" UNIVERSITY of BUCHAREST

ETTI-B DOCTORAL SCHOOL

Decision No. 610 of November 3, 2020

SUMMARY of PhD THESIS

ARTIFICIAL INTELLIGENCE TECHNIQUES FOR BUSINESS

PhD Student: **Thitimanan Damrongsakmethee**
PhD Supervisor: **Prof. Victor-Emil Neagoe**

DOCTORAL COMMISSION

President	**Prof.dr.ing Gheorghe Brezeanu**	from	**Univ.Politehnica Bucharest**
PhD supervisor	**Prof.dr.ing. Victor-Emil Neagoe**	from	**Univ.Politehnica Bucharest**
Reviewer	**Prof.dr.ing.Alexandru Serbanescu**	from	**Academia Technica Militara "Ferdinand I"**
Reviewer	**Prof.dr.ing Nicu Bizon**	from	**University of Pitesti**
Reviewer	**Prof.dr.ing Mihai Ciuc**	from	**Univ.Politehnica Bucharest**

BUCHAREST 2020

Abstract

The thesis is dedicated to the application of artificial intelligence techniques for business, particularly for financial prediction. This work is focused on the following two research directions: credit scoring and financial time series prediction. Credit scoring is concerned with the prediction of financial risk and informing managerial decision making in the money business. The quality of risk analysis may affect the financial performance of the bank. For credit scoring, one has experimented a set of classifiers (Multilayer Perceptron (MLP), decision trees, Support Vector Machine (SVM), C4.5 algorithm cascaded with AdaBoost) as well as feature selection techniques (Principal Component Analysis (PCA) and ReliefF). On the other side, the financial time series prediction tools are useful to business leaders and organizations to improve decisions regarding the effects of future predicted changes. There have been considered the exchange rate forecasting and the stock market prediction. There have been chosen two neural network techniques for financial time series prediction: Nonlinear Autoregressive Exogenous (NARX) model and Deep Learning Long Short-Term Memory (LSTM) model. The research results have been published; five papers have been included in International Conference Proceedings (WOS indexed) and a sixth paper has been published in an International Journal (Scopus indexed).

Contents

List of tables

List of figures

Chapter 1 Introduction

This thesis is prepared with the aim of studying and developing techniques (algorithms) in the application of machine learning for businesses. The research studied the machine learning classification techniques. It consists of Multilayer Perceptron (MLP) models, Support Vector Machine (SVM), C4.5 decision tree, which is used for classification of the credit scoring. For this classification problems, this research wants to compared the performance of the models and find a good classification model. The research was applied to use machine learning applications in forecasting models for businesses. We used a Nonlinear Autoregressive Exogenous (NARX) model to predict the currency exchange rate. We applied the deep learning technique with Long Short-Term Memory (LSTM) neural network model to forecast the time-series datasets. This thesis designed the machine learning techniques with innovative components for applied to use in the business such as:

1) Application classification techniques of machine learning for credit scoring modeling and applied classification versus feature selections techniques as the Principal Component Analysis (PCA) and ReliefF algorithm for credit scoring.

2) Applying the AdaBoost of the C4.5 decision tree technique of classification models for credit scoring.

3) Implementing the NARX neural network model to predict the currency exchange rates.

4) Applying the advanced deep learning techniques of the LSTM neural network model for the currency exchange rate time-series data.

5) Implementing the prediction technique of deep learning LSTM neural network model for the stock market.

Chapter 2 Background study

In this chapter will summarize the principles, theories and the relevant work used in this research. The content explains the meaning, concepts, and processes of credit scoring and the related researches of the classification techniques of credit scoring. The research explained the basic information about financial forecasting, the basic forecasting of the stock prices market, and the basic principles of data mining applied in this research. Credit Scoring is a fundamental variable that financial institutions use as a tool to measure the behavior of credit applicant's intention to repay debts by identifying those at risk so that financial institutions can prove credit effectively [1]. This section is a collection of 20 research papers selected from a number of rigorous scientific journals, focusing on relevant articles, machine learning to credit scores, and studies focused on improving or proposing new approaches. For this thesis, the details of the credit scoring data tested for the models were based on actual data available from the University of California Irvine Machine (UCI-ML) data collection which is German credit data and Australian credit data. Today, almost every country has new businesses emerging and trying to be leaders in various matters. It is well known that successful businesses are often planned businesses in the future. In general, statistical knowledge is often added to help them make plans in anticipating future events using time-series analysis [2]. Basically, the analysis of time series data is a statistical regulation that can use current data or events, possibly historical data, to make predictions that will happen in the future [3]. For example, the purchasing department may use previous experience data to make predictions or decide which materials to buy in the next three months [4]. Another example: The Electricity Generating Authority of Thailand may decide that the reduction of electricity bills will increase at a rate similar to a decade ago to forecast demand in the next 10 years or 20 years [5]. From the example above, it can find the value of variables to predict in many past periods that data is called "time-series" and it call the method used for this data value as "Time-series analysis". Time-series data is a collection of data collected over a continuous period of time, such as the daily stock market indices when closing the quarterly revenue trading, the dataset of *Gross National Product (GNP)*, the dataset of Gross Domestic Product (GDP) per year and the company daily income information within 1 year [5]. In most cases, time series data can be in the form of annual, quarterly, monthly or daily data, depending on their storage and utilization suitability[6]. In the collection of time-series data, it is compiled about the appropriateness of the data used. Data in business is constantly changing, so organizations must find ways to use it to make decisions, which will affect the changes that will occur in the business. Time series analysis is essential for all businesses to make informed business decisions. It is a technique that can help control current operations and a tool that can help in planning with the same goal in predicting future events [7]. This section is a collection of 25 research papers chosen by various scientific journals by focusing on the study of related articles, using machine learning to prediction on the currency exchange rate and the stock price market prediction, and studies that focus on improving or suggesting new approaches to time-series data. Machine learning techniques were applied in forecasting by using financial data such as foreign exchange rates and stock price market. In conducting the research, researcher chose to use deep learning techniques with Long-Short-Term Memory (Deep LSTM) to estimate the exchange rate in Thai baht (THB) against the US dollar (USD). In addition to this study, the researcher also chose the deep learning LSTM neural network for predicting stock prices market (Dow Jones Industrial index, and S and P 500 indexes).

Chapter 3 Classification models and prediction models / Evaluation classifiers performance and forecasting evaluation measures

3.1 Classification models

3.1.1 Multilayer Perceptron (MLP): The MLP neural network is a form of multi-layered, structured neural network [8]. There are a supervised training process and a process of sending back values called "Backpropagation". The MLP neural network is including three main layers: the input layer, the hidden layer, and the output layer [9]. There are 2 sub-sections for sending back the values in the neural network: forward pass and backward pass. For a forward pass, the data goes through the neural network at the input layer and is passed from one layer to the other until the output layer. In the backward pass, the connection weight values will be adjusted according to the error-correction, which means that the actual response and target response have the error signal. This error signal is sent back into the neural network system in the opposite direction of the connection. The connection weight value will be adjusted until the actual response is close to the target response. The principle of operation of MLP is that each layer of the hidden layer has to calculate a function when it receives a signal from a node in the previous layer, called "*the Activation Function"*.

3.1.2 Support Vector Machine (SVM) is a 2-class linear classifier which has recognized its classification efficiency over other classification methods[10]. The advantage of SVM is that it is efficient in classifying large-dimensional data. In addition to using the kernel function, to convert data to a higher dimension in a feature space that can effectively classify obscure data, the SVM principle is to find the line with the largest margin (Maximum Margin) that can divide the data into two classes [11]. The model presented here is a popular machine learning tool used in various research studies. That can offer a solution to both the classification of the data and the regression analysis.

3.1.3 Decision tree is a very popular model for machine learning and often achieving cutting-edge work regarding black box predictions[12]. There are some popular decision trees such as C4.5, CART, boosted tree and random forest[13]. The probability method allows us to encrypt and to make assumptions about tree structure and shared statistical strength between node parameters. It is also a mechanism for predicting the probability, which is important for applications with uncertain quantitative significance.

3.2 Prediction models

3.2.1 NARX : A Nonlinear Autoregressive Exogenous (NARX) neural network model [14] is a nonlinear autoregressive model with exogenous input attributes for a recurrent dynamic neural network learning with feedback connections. This technique can increase the accuracy results for learning and prediction. The NARX neural network model has two different kinds of structure; *series-parallel architecture (called "open-loop")* and *parallel architecture (called "close-loop")* [15].

3.2.2 Long Short -Term Memory (LSTM): LSTM model approach, created by Hochreiter and Schmidhuber in 1997, solves the problem of missing gradations [16]. LSTM, compared to simpler RNN uses memory cells instead of neurons. A single-stage cell is characterized by a complex architecture consisting of the input, hidden and output states. Inside the block of memory cell there are three exponential and adjustable gateway units: input gate, forgotten gate, and output gate[17].

3.3 Evaluation classifiers performance and forecasting evaluation measures

3.3.1 Evaluation classifiers performance: for measuring the performance of the models designed in this research, the researcher used a rate-of-accuracy variable to evaluate the model's performance. Calculation of the testing accuracy of the model is detailed in the following equation [18] :

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100\,, \tag{3.1}$$

where TP represents the number of applicants for "bad credit" with a predicted value of "bad credit",

TN represents the number of applicants for a "good credit" with a predicted value of "good credit",

FP represents the number of applicants for "good credit" with a predicted value of " bad credit ",

FN represents the number of applicants for "bad credit" with a predicted value of "good credit".

3.3.2 Forecasting evaluation measures: this research has chosen four performance error measures to test the experimental results of the forecast: Mean Squared Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) and the forecast accuracy performance [18] , [19]. These measures use the following variables: y_t is the original value, f_t is the predicted value, $e_t = y_t - f_t$ is the predict error, n is the extent of the testing dataset.

Details of the forecasting measures are shown in table 3.1.

Tab. 3.1 *The forecasting measures of the Deep LSTM neural network model.*

Forecasting measures	Equations
Mean square error (MSE)	$MSE = \frac{1}{n}\sum_{t=1}^{n} {e_t}^2$
Root mean square error (RMSE)	$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n} {e_t}^2}$
Mean Error Percentage (MAPE)	$MAPE = \frac{1}{n}\sum_{t=1}^{n} \left\lvert\frac{e_t}{y_t}\right\rvert \times 100$
Forecasting accuracy	$\left(1 - \frac{1}{n}\sum_{t=1}^{n} \left\lvert\frac{e_t}{y_t}\right\rvert \times 100\right) \times 100$

Chapter 4 Applied feature selection techniques for credit scorings

This research was interested in selecting, constructing, and comparing test results of decision tree and SVM models, using feature selection techniques (Principle Component Analysis: PCA and ReliefF Algorithm) to make the models more efficient. This chapter explains the types and operations of the feature selection: PCA and ReliefF algorithm. To apply the application in this research we will show the experimental results of the classification models by comparing the results of SVM with kernel functions model and the decision tree method in conjunction with feature selection techniques.

4.1 The details of dataset

In this research, the proposed model has used the financial dataset from the UCI Machine learning database. In table 4.1, shows the details of the characteristics of German credit dataset.

Tab. 4.1 *Details of the characteristics of German credit dataset.*

Dataset name	Attributes number	Good credit	Bad credit	Total	Classes
German credit dataset	25	700	300	1000	1 = good (accepted) 2.= bad (rejected)

4.2 Proposed the model of applied PCA and ReliefF technique on credit scoring

The classification model by applied the technique of feature selection by using the PCA and ReliefF algorithm on financial datasets is shown in figure 4.1.

4.2.1 PCA technique: PCA is a technique for analyzing the components for multivariate data without variable segmentation. This technique is interested only to find the relationships of those variables which will create a new variable consisting of variance or variance of the previous variable. Table 4.2, shows the computation of the resulting variance using the PCA technique involving each component. It shows the optimal number of components in German credit score dataset. Results of the test can be defined as the smallest number of elements with the highest variable value.

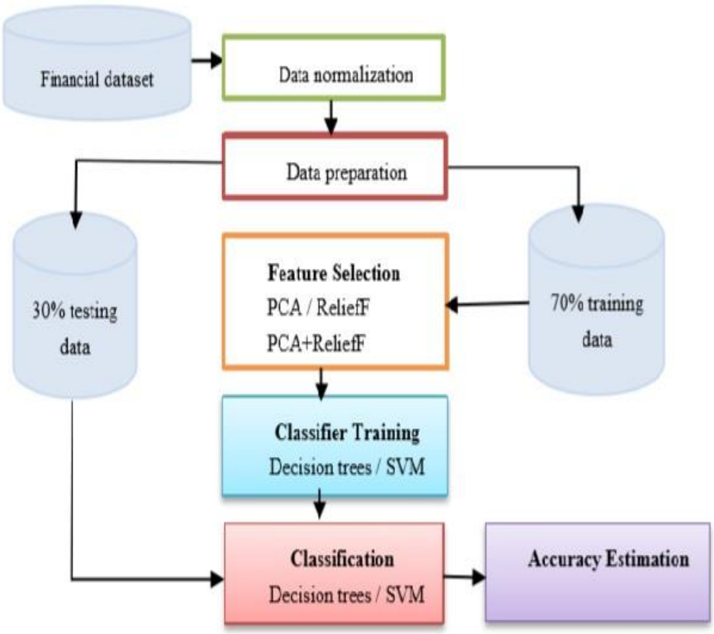

Fig. 4.1 *Proposed the model of applied PCA and ReliefF technique on credit scoring* [20].

4.2.2 ReliefF algorithm: For this work studies, the research applied the ReliefF algorithm to find a candidate feature subset. The variable k was set starting with $k = 10$ (k is the nearest neighbor). The results of the tests for stability and reliability of the ratings and weights of the ReliefF algorithms for the values of k are shown in table 4.3.

Tab. 4.2 *The eigenvalues, variance in PCA transformed space and the cumulated variance.*

Component Index	**1**	**5**	**10**	**15**	**18**	**20**	**23**	**25**
Eigenvalue	2.4602	1.6435	1.0243	0.9334	0.8002	***0.4832***	0.1283	0.1112
% Variance in PCA transformed space	9.84	6.56	4.08	3.73	3.20	***1.93***	0.48	0.44
% Cumulated Variance	9.84	37.92	61.24	80.76	90.80	***95.33***	99.12	100.00

For experiments in table 4.2 above, based on the test results, using the PCA technique, the percentage of cumulative variance was 95.33%, consistent with a selection of m = 20 components. The eigenvalue is 0.4832 and the percent variance in PCA transformed pace is 1.93.

Tab. 4.3 *Ranks and weights of features in German credit dataset.*

Ranks	**Weights**	**Mark**
1	***0.1589***	***Most important feature***
3	0.0771	
7	0.0406	
6	0.0399	
9	0.0379	
24	0.0243	
2	0.0243	
8	0.0234	
10	0.0193	
11	0.0167	
5	0.0167	
12	0.0166	
14	0.0164	
23	0.0141	
13	0.0137	

16	0.0094	
17	0.0080	
18	0.0044	
4	0.0041	
15	0.0021	
22	***0.0004***	**Least important feature**
21	-0.0007	
19	-0.0017	
20	-0.0027	

Table 4.3 shows the ranks of the predictor numbers listed according to their ranking. This research found that the first predictor (with weights value of 0.1589) is the most important and the 22nd predictor (with weights value of 0.0004) is the least important. Fifteen higher weight features were used to classify according to the weight distribution of the characteristic parameters. We also used fifteen features to classify model for ReliefF function.

4.3 The performance of the models

For testing the differences between the models, both feature selection techniques and no feature selection techniques were used. The results are shown in table 4.4

Tab. 4.4 *Test results for various classification models with the feature selection techniques: PCA, ReliefF, PCA+ReliefF and without feature selection technique.*

Classifier model	% accuracy without feature selection	% Accuracy with feature selection		
		PCA	**ReliefF**	**PCA + ReliefF**
SVM-RBF	75.35	76.00	86.16	85.18
Decision tree-Fine tree	74.00	***72.23***	***88.28***	***91.67***
Decision tree-Medium tree	***77.33***	78.00	83.00	81.67
Decision tree-Coarse tree	74.67	76.00	75.33	80.49

From table 4.4, when compared the results of classification models on German credit dataset. The results have shown that the decision tree-Fine tree has the highest accuracy (91.67 %) when compared to the other. By testing all three methods, it can be concluded that combining the PCA technique with ReliefF techniques gives the best classify results.

Chapter 5 Enhancing C4.5 with AdaBoost algorithms and MLP model for credit scoring

In this chapter, algorithms are presented to compare the model performance in classifying credit scoring data using the C4.5 decision tree, C4.5 plus AdaBoost technique and MLP model. In the algorithm design for the C4.5 model, the AdaBoost technique was incorporated into the model to further optimize the credit scoring validation. The mode is consisting of two phases for testing with Decision tree models: 1) classification using C4.5 decision tree model, 2) model design using AdaBoost technique for binary classification (accepted or rejected credit).

5.1 The details of the datasets

In this research, models were designed and tested using credit scoring datasets from the UCI Machine Learning database – see the table 5.1.

5.2 Proposed techniques of C4.5 and MLP

The structure of the with AdaBoost technique is shown in figure 5.1. The MLP model algorithm follows to obtain the most accurate model. In figure 5.2 are shown the neuronal relationships of the MLP model, the transfer function, and the non-linear function.

__Tab. 5.1__ The detail of datasets.

Name of Dataset	Attributes	Status of credit		Total	Label
		Good	Bad		
German credit dataset	25	700	300	1000	1 = good ,2 = bad
Australian credit dataset	15	307	383	690	1= good, 0= bad

In table 5.2, the segmentation rates of the training dataset, the test dataset, and the validation dataset were shown in both datasets.

__Tab. 5.2__ The divided ratio of the datasets.

Name of dataset	Training	Testing	validation	Total
	70%	25%	5%	
German credit dataset	700	250	50	1000
Australian credit dataset	483	172	35	690

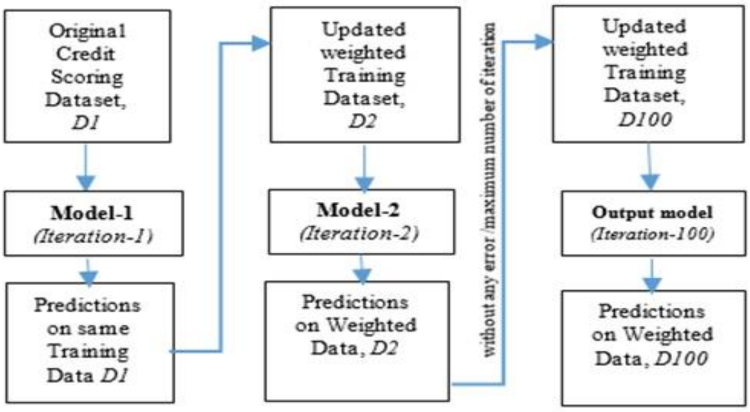

Fig. 5.1 *The proposed technique of C4.5+AdaBoost model for credit scoring.*[21].

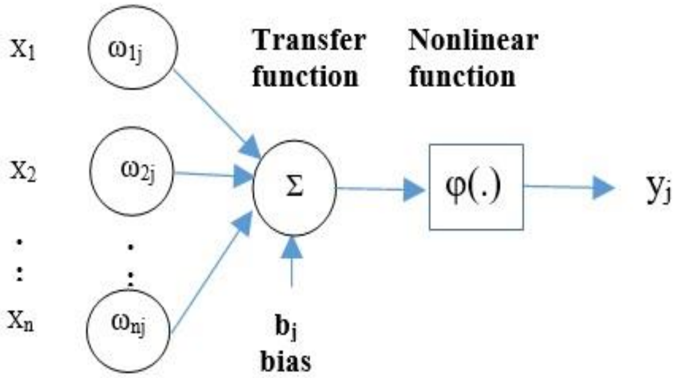

Fig. 5.2 *The application of the MLP model* [21].

5.3 The results of the proposed model

Table 5.3 shows the experimental results of the credit score classification of C4.5 decision tree model and C4.5 decision tree followed by AdaBoost technique.

__Tab. 5.3__ Results of a credit scoring model with the C4.5 decision tree model and C4.5 decision tree followed by AdaBoost technique.

Name of dataset	Proposed model	Percentage of accuracy
German credit dataset	C4.5	72.68
	C4.5+AdaBoost	***78.67***
Australian credit dataset	C4.5	83.42
	C4.5+AdaBoost	***89.00***

The results indicated in table 5.3 suggest that the C4.5 decision tree technique with AdaBoost technique (by 100 iterations bagging schedule) get the highest predictions value in German credit scoring dataset and dig up the Australia rating data with details concluded that: The classification accuracy in German credit with the C4.5 decision tree model followed by the AdaBoost technique was 78.67% compared to the C4.5 model, leading to an accuracy of just 72.68%. For the Australian credit, the research concludes that C4.5 model, followed by the AdaBoost technique, led to an accuracy of 89.00% compared to the C4.5 decision tree model that achieved accuracy of just 83.42%. The results of MLP model credit scoring are shown in table 5.4.

Tab. 5.4 *The best experimental results of MLP credit scoring model (when set H=6 hidden layers).*

Name of Dataset	**input neuron numbers**	**Optimum number of neurons for the hidden layer $h_1=\ldots=h_6$**	**Optimum number of training epochs**	**Percentage of Accuracy**
German credit dataset	24	18	8	81.20
Australian credit dataset	14	41	12	90.85

The results of the MLP model shown in table 5.4 can be summarized as follows: 1) The MLP model test results of the German credit scoring data set showed that the best accuracy was 81.20%, while the Australian credit scoring data set classifying credit scoring with the MLP model led to a classification accuracy of up to 90.85%. 2) To obtain a MLP model that fits that credit rating classification scheme, we need to determine the optimal number of neurons for the hidden layers and the optimal number of training epochs procedures to achieve the best classification efficiency. Table 5.5 shows a comparison of the best results of tests performed on both models.

Tab. 5.5 *Performance comparisons results of the C4.5 model (C4.5 and C4.5 followed by AdaBoost technique) and MLP model.*

Model	**German Credit (*% correct*)**	**Australian Credit (*% correct*)**
C4.5	72.68	83.42
C4.5 +AdaBoost	78.67	89.00
MLP	***81.20***	***90.85***

By comparing the best performance of the tested models in table 5.5, the results of the two datasets could be evaluated. It can be concluded that the MLP model was more effective in classifying credit scoring data than the C4.5 decision tree model (both C4.5 and C4.5 using AdaBoost technique). The MLP model had the highest accuracy (81.20%) in German credit. In Australian credit, the MLP model was also able to show the highest accuracy (90.85%).

Chapter 6 Using A Nonlinear Autoregressive Exogenous (NARX) neural network model on currency exchange rate forecasting

In this chapter we can find some techniques for predicting time-series data using NARX neural network model to predict the currency exchange rates. In this study, we chose to use a data set of the daily exchange rate of Thailand Baht against the US dollar (THB:USD) with information taken from the Bank of Thailand by selecting all historical data for a total of 10 years. There were used data on some of the most important economic indicators to be used as an input to test the neural network model, namely inflation rate, GDP growth rate, interest rate, trade balance and account balance for input data in NARX neural network model

6.1 The details of dataset

In this study we used the dataset obtained from historical foreign exchange rate of Bank of Thailand for ten years (since 2009 to 2018), which contains a lot of important information for the Thai baht versus USD exchange rate. The total number of time steps in the dataset was 2373 [22]. The dataset was split as follows: 70% - training data, 25% - testing data and 5% - validation data. The details of splitting the exchange rate time-series data are shown in table 6.1.

Tab. 6.1 *The amount of the dataset and the proportion of the dataset.*

Data	Amount	Ratio (Percentage)
Training set	1661	70
Testing set	593	25
Validation set	119	5
Total	2373	100

6.2 Proposed NARX neural network architecture

Designing NARX neural network model [23] with series-parallel architecture is made by the following analytical input output in equation 6.1.

$$\hat{y}(t+1) = F\left(y_{(t)}, y_{(t-1)}, \dots, y_{(t-d_y)}, x_{(t)}, x_{(t-1)}, \dots, x_{(t-d_x)}\right), \qquad (6.1)$$

where $F(\cdot)$ is the mapping function of NARX neural network;

$\hat{y}(t+1)$ is a predicted output of NARX neural network at the moment $(t+1)$;

$y(t), y(t-1), \dots, y(t-d_y)$ are the true previous values of time-series output values of NARX neural network;

$x(t+1), x(t), x(t-1), \dots, x(t-d_x)$ are the input of NARX neural network;

d_x represents the number of input delays and d_y represents the number of output delays.

In equation 6.1, in the open-loop for the series-parallel architecture, future value of time-series $\hat{y}(t+1)$ is predicted from the present and the previous values of $x(t)$ and the true previous values of time-series $y(t)$. As we already mentioned, the proposed method of prediction of time-series uses NARX neural network. The proposed NARX neural network model structure is shown in figure 6.1.

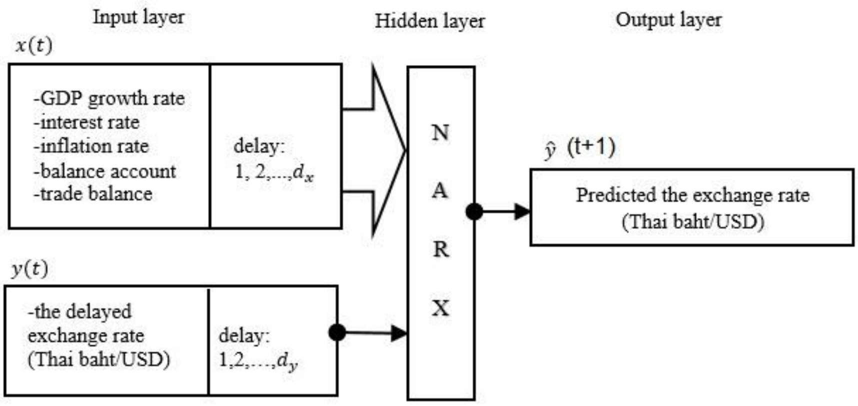

Fig. 6.1 *The architecture of proposed NARX neural network* [24].

We used historical data to train the model of a best NARX neural network in open-loop. The model was used to calculate an exchange rate as the input for the next step in close-loop. We assume that there are 3 input neurons and 1 output neuron. The number of input delays is varying from one to three in order to consider the lagged values of both independent and dependent variables. The number of neurons in the hidden layer simulation was found to be 10. In figure 6.2 we can see the model of NARX neural network configuration.

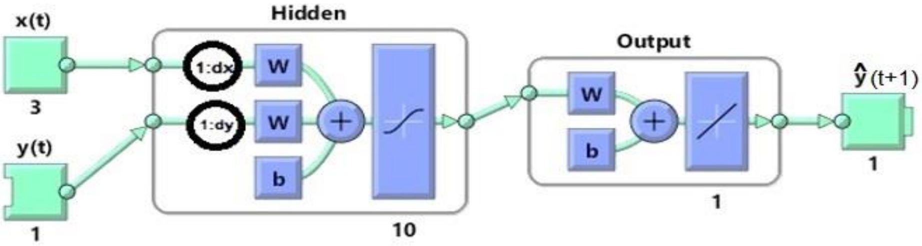

Fig. 6.2 *NARX neural network configuration* [24].

Table 6.2 shows the details of the parameters for the best performance with the smallest input set of NARX neural network.

Tab. 6.2 *The best results of NARX neural network architecture.*

Parameter	**Data/ Technique used**
Number of input neuron (s)	3
Number of output neuron (s)	1
Number of hidden neuron (s)	10
Time steps of training data, testing data and validation data	1661 : 593: 119
The number of input delays (d)	3
Training parameter	Error : MSE, MAPE, Learning algorithm : The scaled conjugate gradient (trainscg)
Training Algorithm	Feed-forward neural network

6.3 The results of proposed model

In order to achieve the research objectives, design for testing the model was divided into 3 parts:

Part 1 - Finding the important variables of input data: we will find the important variables of input data from five economic indicators: inflation rate, GDP growth rate, interest rate, trade balance and account balance. To complete the structure of the NARX neural network we need to study the determination of important parameters for the neurons number in each layer. Table 6.3 shows the results obtained when determining the number of neurons in a network, for different layers of NARX neural network model.

Tab. 6.3 *The experimental results (setting the number of hidden neuron (H) to 10, while number of input delays(d_x) and number of output delays(d_y) are 4) (H = 10, $d_x = d_y$ = 4).*

Number of Input	**Input Significance**	**Performances**		
		MSE	***MAPE***	***R***
2	interest rate, GDP growth rate	0.011	6.019	0.997
3	inflation rate, interest rate, GDP growth rate	**0.008**	**4.263**	**0.998**
4	inflation rate, interest rate, GDP growth rate, balance account	0.009	4.475	0.997
5	inflation rate, interest rate, GDP growth rate, trade balance, balance account,	0.013	6.441	0.992

Table 6.3 shows how the research has trained the neural network to fit the input and target with 4 sets of input significance. The research found that the best structure of the proposed NARX neural network model was obtained by using three input neurons in the input layer: inflation rate, interest rate and GPD growth. There are 10 neurons in the hidden layer and there is 1 neuron in the output layer. This test has established that the input and output delays (d) have a value from 1 to 4. The best performances of MSE, MAPE and R obtained with this structure are 0.008, 4.263% and 0.998.

Part 2 - Finding the appropriate number of hidden neurons : we find the appropriate number of hidden neurons for testing the NARX neural network model by tested 2 groups: group 1, where hidden neuron had the values of 1, 5, 10, 15, 20, 25 and 30 and group 2, where the hidden neuron is set to 40, 50, 60, 70, 80, 90, and 100. We selected the best experimental results and trained the network to fit the inputs (see table 6.3), then split the initial value of hidden neuron in 2 sets: a first set of hidden neurons which vary from 1 to 30, with an increment of 5, and a set of hidden neurons which vary from 40 to 100, with an increment of 10. The research also set the number of input and output delays (d) from 1 to 4. The experimental results are shown in table 6.4.

Tab. 6.4 *The best experimental results when setting the value N = 3 (inflation rate, interest rate, GDP growth rate) depending on the number of hidden neurons and a ratio of input output delay of 1: 4 ($N = 3$, $d_x = d_y = 4$).*

Set 1							
Hidden neuron	***1***	***5***	***10***	***15***	***20***	***25***	***30***
MSE	0.018	0.025	**0.012**	0.013	0.019	0.021	0.026
MAPE	3.396	4.010	**3.026**	3.141	3.401	4.001	4.024
R	0.997	0.995	**0.998**	0.997	0.996	0.996	0.995
Set 2							
Hidden neuron	***40***	***50***	***60***	***70***	***80***	***90***	***100***
MSE	0.026	0.077	0.083	0.018	0.140	0.086	0.327
MAPE	4.031	4.716	4.823	3.346	5.267	4.842	6.019
R	0.995	0.988	0.988	0.996	0.978	0.986	0.949

From table 6.4 we can see that the best structure of proposed NARX neural network model is obtained by using 3 neurons in the input layer (inflation rate, interest rate and GDP growth rate) with 10 neurons in the hidden layer and 1 neuron in the output layer. The best performances of MSE, MAPE and R obtained using this structure are 0.012, 3.026% and 0.998.

Part 3 - Finding the appropriate input and output delays : we tested the NARX neural network model to find the appropriate input and output delays with the input and output delays being 1, 3, 4, 5, 10, 15, 20, 25, and 30. The research was designed to train the network again to fit the input and the target for d_x and d_y. This task needs to arrange the inputs variable along with the inflation rate, interest rate and GDP growth rate (n = 3). The b research configured the value number of the hidden neuron with 10 neurons in the hidden layer and 1 neuron in the output layer. It chose a number of input delays d_x and d_y of 1, 2, 3, 4, 5, 10, 15, 20, 25 and 30 to find the best results for the number of input delays. The results are shown in table 6.5.

Tab. 6.5 *The experimental performance results for N=3 inputs depending on the number of hidden neurons (H=10).*

Test 1	$d_x = d_y$	***1***	***2***	***3***	***4***	***5***	***10***	***15***	***20***	***25***	***30***
	MSE	0.018	0.023	0.019	0.017	0.013	0.011	0.016	0.020	0.018	0.022
	MAPE	3.728	3.353	3.812	3.689	3.538	3.243	3.625	3.878	3.728	3.916
	R	0.997	0.966	0.996	0.997	0.998	0.998	0.997	0.997	0.997	0.996
Test 2	$d_x = d_y$	***1***	***2***	***3***	***4***	***5***	***10***	***15***	***20***	***25***	***30***
	MSE	0.019	0.011	**0.006**	0.010	0.027	0.027	0.012	0.062	0.011	0.020
	MAPE	3.762	3.532	**3.001**	3.318	4.526	4.523	3.738	5.216	3.158	3.892
	R	0.996	0.998	**0.998**	0.998	0.995	0.995	0.997	0.990	0.998	0.996
Test 3	$d_x = d_y$	***1***	***2***	***3***	***4***	***5***	***10***	***15***	***20***	***25***	***30***
	MSE	0.037	0.012	0.011	0.008	0.014	0.023	0.030	0.014	0.069	0.017
	MAPE	4.524	3.623	3.562	3.018	3.745	3.923	4.126	3.748	5.518	3.845
	R	0.994	0.998	0.998	0.998	0.997	0.996	0.995	0.997	0.989	0.997
Test 4	$d_x = d_y$	***1***	***2***	***3***	***4***	***5***	***10***	***15***	***20***	***25***	***30***
	MSE	0.011	0.011	0.013	0.010	0.018	0.026	0.020	0.073	0.012	0.012
	MAPE	3.509	3.502	3.625	3.453	3.723	3.992	3.845	5.827	3.572	3.570
	R	0.998	0.998	0.997	0.998	0.997	0.995	0.996	0.988	0.998	0.998
Test 5	$d_x = d_y$	***1***	***2***	***3***	***4***	***5***	***10***	***15***	***20***	***25***	***30***
	MSE	0.027	0.017	0.012	0.020	0.017	0.021	0.051	0.018	0.016	0.023
	MAPE	3.992	3.855	3.552	3.946	3.853	3.962	4.346	3.915	3.623	3.986
	R	0.995	0.997	0.998	0.996	0.997	0.997	0.992	0.997	0.997	0.996

We used the results from table 6.5 for training the network to fit the input and the target for the delay d_x and d_y. Five tests were made (Test 1…Test 5). The best results of the experiments of proposed NARX neural network model were obtained by 3 neurons which were used in the input layer using inflation rate, interest rate and GDP growth rate variables. There were 10 neurons (H = 10) in the hidden layer and 1 neuron in the output layer. The best number of input delays d_x and the best number of output delays d_y is 3 (d_x= d_y=3). The best efficiency value of MSE, MAPE and R value are 0.006, 3.001% and 0.998. The details are shown in table 6.6.

Tab. 6.6 *The best experimental performances results (hidden neuron H=10, the input delays d_x = output delays d_y = 3).*

Number of Input	Input Significance	Performances		
		MSE	***MAPE***	***R***
3	GDP growth rate, interest rate, inflation rate	0.006	3.001	0.998

In figure 6.3, the research presents an association of predicted and actual exchange rate for all test, training and validation samples. It also shows the magnitude of error expressed as the difference between predicted and actual values. Even if the graphic strongly suggests that NARX neural network fits properly in predicting exchange rate, we still need to justify this claim from the quantitative point of view. In figure 6.4, the research has shown an error histogram with 20 bins of NARX neural network for the steps of training, testing and validation. This error shows that data fitting errors are scattered out within a large range around zero. In this situation, we can see that most errors can be found in the interval -0.3891… 0.3561 with the learning point having an error of 0.01218. Training with the aid of the scaled conjugate gradient algorithm with feed-forward neural network converged after less than 74 epochs. In the figure 6.5, we can remark that the model has a good stability and no overshoot. Regression of error between actual and forecast data is presented in figure 6.6.

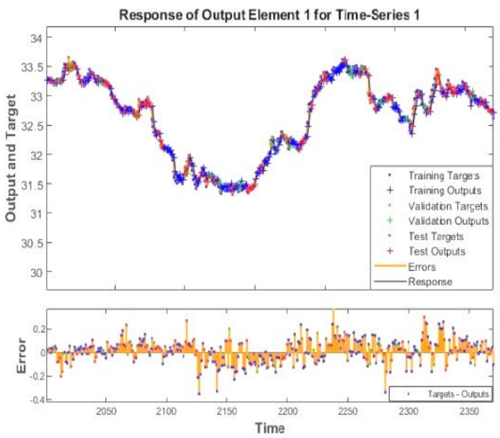

Fig. 6.3 *Time-series response for NARX neural network* [24].

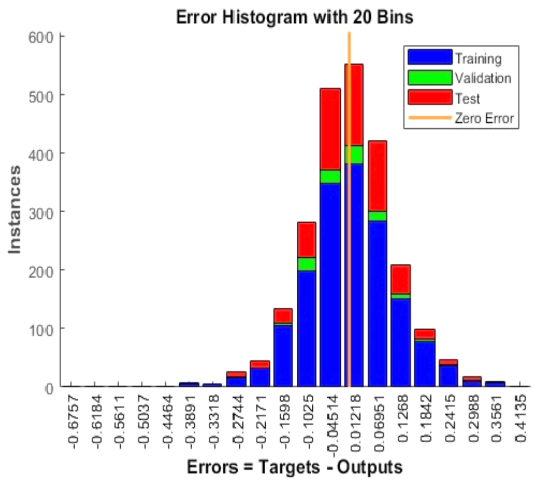

Fig. 6.4 *An error histogram of NARX neural network* [24].

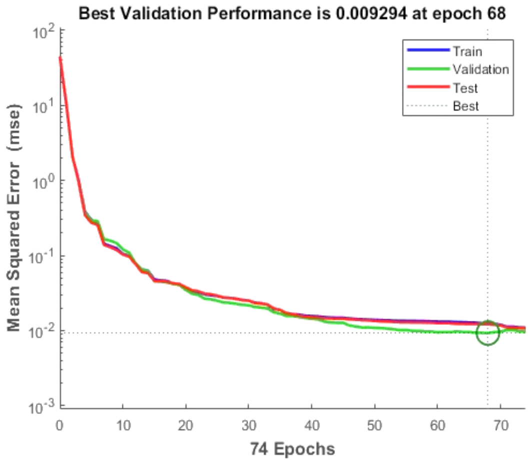

Fig. 6.5 *The best validation performance of NARX neural network model* [24].

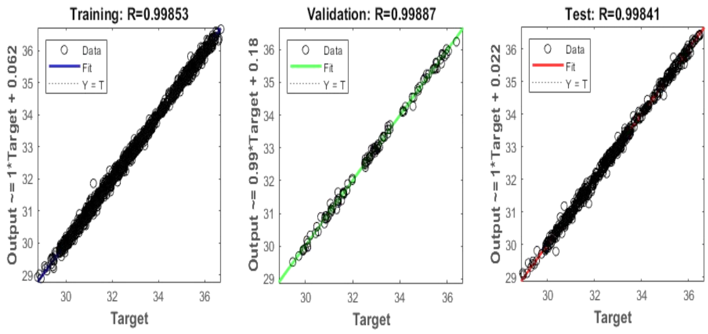

Fig. 6.6 *An error regression of NARX neural network model.* [24].

Chapter 7 Deep Learning Long Short- Term Memory (DLSTM) neural network model for currency exchange rate prediction

This chapter demonstrates the application of deep learning model to forecast the exchange rate data of Baht per US Dollar. Research has studied and tested the model using the same data like in the previous chapter. The main objective of this chapter is to continue the results from the previous chapter by developing a deep learning model with Long Short-Term Memory (DLSTM) model to perform time-series predictions. Experimental results obtained were compared with previously predicted test results by NARX neural network method.

7.1 The details of currency exchange rates dataset

Modeling a DLSTM neural network in this research used historical exchange rate for Thai Baht versus US dollars from the Bank of Thailand, by selecting data for a total of 10 years with information from April 10, 2009 to April 10, 2019 [25]. Table 7.1 shows that the currency exchange rate dataset has a total of 2442 rows. The tests were performed in order to compare test results from different data sets in series A and series B. In series A the details of the data were divided with a ratio of 80 % for the training data and 20 % for the testing data. In series B the dataset was divided with a ratio of 70 % for the training set and 30 % for the testing set.

Tab. 7.1 *Details of the divided ratio of the currency exchange rates of Thai Baht versus USD.*

The set of the data	The training ratio: test (%)	The amount of training : test	Total
Series A	80 : 20	1953 : 488	2442
Series B	70 : 30	1709 : 732	2442

7.2 Proposed DLSTM neural network model

The structure of the proposed DLSTM neural network model is shown in figure 7.1.

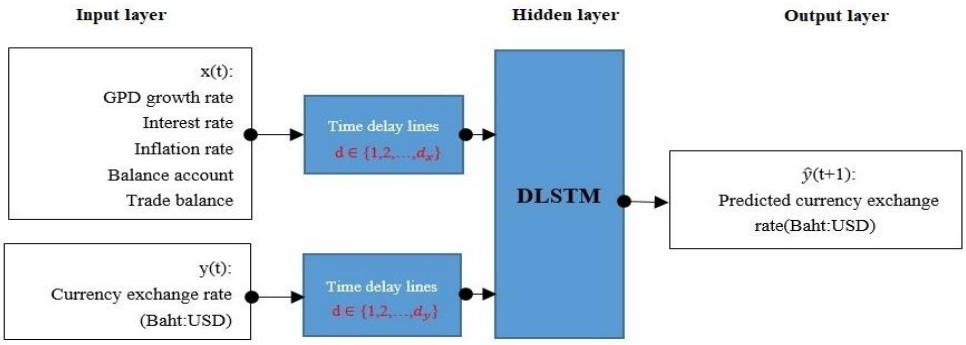

Fig. 7.1 *The structure of the proposed DLSTM neural network model* [26].

For the resulting equation using the DLSTM neural network model forecasting calculation is given by:

$$\hat{y}\ (t+1) = F(y(t), y(t-1), \dots, y(t-d_y), x(t), x(t-1), \dots x(t-d_x)) \qquad (7.1)$$

where $F(\cdot)$ is function of DLSTM neural network model;

$\hat{y}(t+1)$ are the predicted values from the present and the previous values of $x(t)$ and the true previous values of the time-series $y(t)$ at $(t+1)$ time;

$y(t), y(t-1), \dots, y(t-d_y)$ are the true previous values of the time-series output;

$x(t), x(t-1), \dots, x(t-d_x)$ are the input of the DLSTM model; d_x is the maximum delay for the input and d_y is the maximum delay for the output.

DLSTM neural network model architecture has 10 layers, as shown in table 7.2. The framework of DLSTM model is shown in figure 7.2.

Tab. 7.2 *Details of DLSTM neural network model layers.*

Layers	Descriptions
Input Layer	The sequence input with *m* neurons including: the GDP growth rate, interest rate, inflation rate, balance account and trade balance {m = $5d_x + d_y$}.
LSTM layers	3 main layers of the LSTM corresponding to the following numbers of neurons for each of LSTM layers: 50, 100 and 150.
Dropout layers	3 dropout layers corresponding to a good probability default with the probability of 50%.
Relu layer	the Relu layer performs a threshold operation to each element of the input data, where any value less than zero is set to zero.
Fully Connected layer	1 fully connected layer with N neurons, the full connected layer acts independently on each time step.
Regression Output layer	1 neuron output for the prediction of the currency exchange rate (THB:USD).

7.3 The experimental results of DLSTM neural network model

Table 7.3 shows the results of training DLSTM neural network model to fit the inputs and targets in series A. The dataset is divided into 80 percent for training data and 20 percent for testing data. The results showed that the best structure of DLSTM neural network model consists of 3 main layers of DLSTM neural network, with the number of neurons for each LSTM layer being 50, 100 and 150. The maximum number of epochs and the maximum number of iterations of the network is 250. The best number of input and output delays (d_x , d_y) is 3 measurements. The best performance of MAPE and MSE values was 0.3122 and 0.0036. For experimental results on dataset series B shown the dataset was divided into 70% for training data and 30% for testing data. The best structure of DLSTM neural network model consists of 3 main layers of DLSTM neural network, with a number of neurons for each LSTM layer of 50, 100 and 150. The maximum number of epochs and the maximum number of iterations of the network is 250. The best number of input and output delays (d_x , d_y) is 2. DLSTM neural network shows that the best performance of MAPE and MSE values was 0.28440 and 00276.

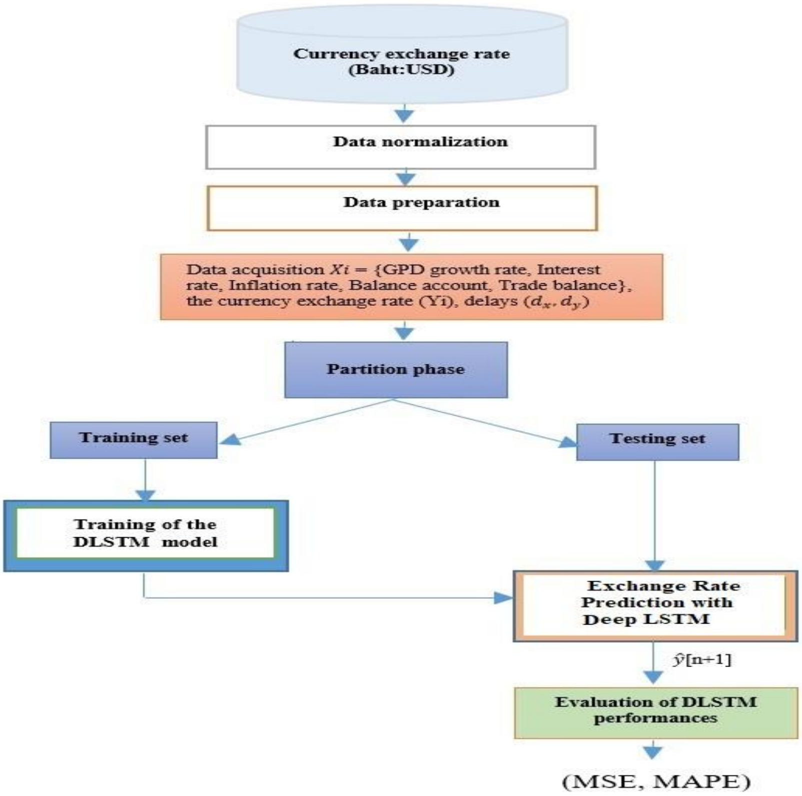

Fig. 7.2 *The framework of DLSTM neural network model architecture* [26].

Tab. 7.3 *The experimental results on series A and Series B as a function of number of iterations (DLSTM neural network model has 3 main LSTM layers with a corresponding number of neurons of 50, 100 and 150).*

Delays	Epochs/ Iterations	Series A (70:30)		Series B (80:20)	
		MSE	MAPE	MSE	MAPE
$d_x = d_y$ =1	150	0.00563	0.3488	0.00532	0.3391
	200	0.00455	0.3194	0.00507	0.3405
	250	0.00433	0.3231	0.00377	0.3087
$d_x = d_y$ =2	150	0.00560	0.3472	0.00792	0.3994
	200	0.00414	0.3174	0.00379	0.3032
	250	0.00418	0.3150	***0.00276***	***0.2844***
$d_x = d_y$ =3	150	0.00534	0.4328	0.00543	0.3418
	200	0.00497	0.3343	0.00516	0.3326
	250	***0.00369***	***0.3122***	0.00333	0.2943

Figure 7.3 shows a comparison between observed and predicted values on data series A. Figure 7.4. shows a comparison between observed and predicted values on data series B.

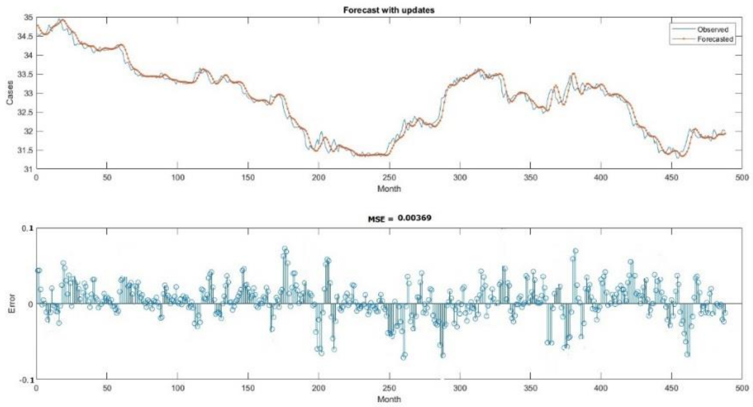

Fig. 7.3 *Comparison between observed and forecasted results and MSE value on series A* [26].

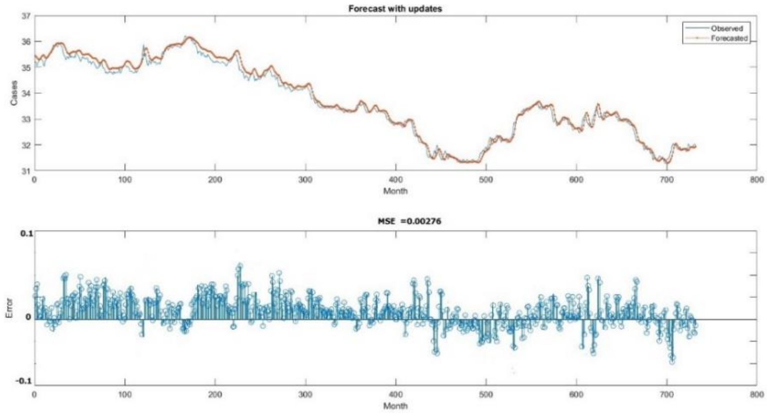

Fig. 7.4 *Comparison between observed and forecasted results and MSE value on series B* [26].

7.4 Comparison between the results of DLSTM neural network and NARX neural network model

The primary purpose of testing the DLSTM neural network model is selecting the best experimental results from table 7.3. and comparing it with the experimental results of the NARX neural network model tested in chapter 6 for forecasting the currency exchange rate of Thai Baht against the US dollar. The compared results are shown in table 7.4.

Tab. 7.4 *Comparison between the results of proposed DLSTM neural network model and NARX neural network model.*

Models	**Performance Measures**		**References**
	MSE	MAPE	
DLSTM	***0.0027***	***0.2844***	the proposed model
NARX	0.0060	3.0010	[24]

In table 7.4, the best experimental results of DLSTM neural network model were compared with the previously tested results of NARX neural network model for forecasting the currency exchange rate of Thai Baht versus US dollar by comparing MSE and MAPE values. DLSTM neural network model had the least MSE of 0.0027 and the least MAPE of 0.2844.

Chapter 8 Deep learning Long Short-Term Memory (Deep LSTM) neural network model for stock prices prediction.

This chapter shows the application of Deep Learning with Long Short-Term Memory (Deep LSTM) neural network model for estimating a stock price market. The research has studied and tested the model by using the stock prices data from the yahoo's financial website. The datasets for this study are Dow Jones Industrial index and S and P 500 index. Both stock price time-series data are the most commonly used benchmarks to determine the state of an economy in the stock market. It also measures the effectiveness of Deep LSTM model and tests the efficiency of Deep LSTM neural network model by comparing the best results of Deep LSTM neural network model with other results of Deep Learning models that has previously conducted researches in the same databases.

8.1 The details of dataset

The stock prices datasets are Dow Jones Industrial and S and P 500 indexes from the yahoo's financial website. Both datasets are from February 2010 to February 2020. The details of the datasets as shown in table 8.1.

Tab. 8.1 *Details of the stock prices datasets.*

Stock prices Dataset	Feature	Training set	Testing set
		80%	***20%***
Dow Jones Industrial	open, low, high, close, adj close, volume	February 2010 – January 2018 (2012 days)	February 2018 – February 2020(499 days)
S and P 500 Indexes	open, low, high, close, adj close, volume	February 2010 – January 2018 (2012 days)	February 2018 – February 2020(499 days)

8.2 Proposed Deep LSTM neural network

This research presents a deep LSTM neural network model to predict daily stock market prices. It starts by loading the stock market datasets, followed by data standard size and data normalization. The stock prices datasets were divided into training data and testing data. The input data are used to assemble open, high, low, adj close and close and to forecast daily closing price data. The model is evaluated for error performances with RMSE, MSE, MAPE, Accuracy. The details are shown in figure 8.1.

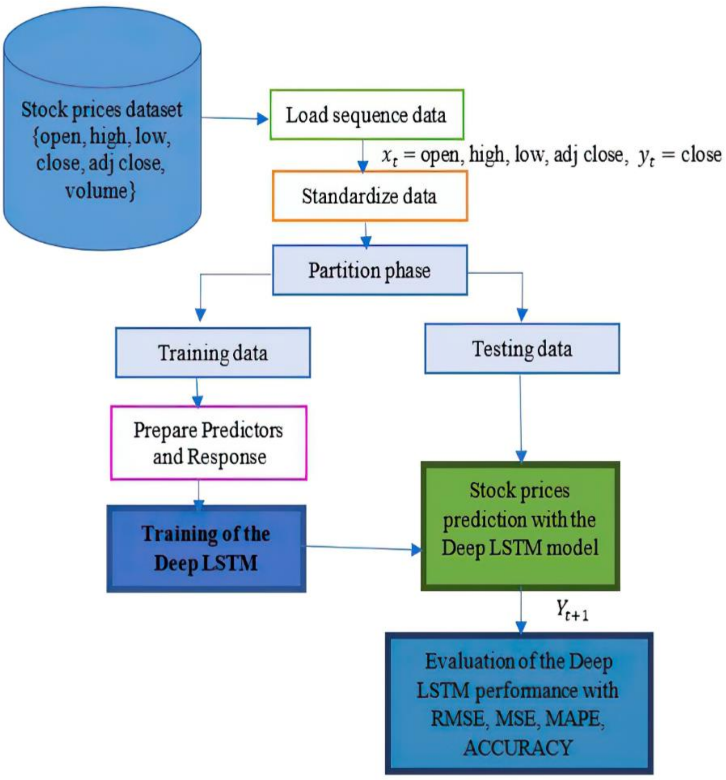

Fig. 8.1 *Framework of Deep LSTM neural network model* [27].

Figure 8.2 shows the architecture of Deep LSTM model with feedback neural network connection. The input data is a vector with 5 dimensions (x_t) representing five features: open, high, low, adj close, and close (y_t). The model predicting the close prices of the next day (Y_{t+1}) is given by

$$Y_{t+1} = \left\{F\left(y_{(t)}, y_{(t-1)}, \dots, y_{(t-d_y)}, x_{(t)}, x_{(t-1)}, \dots, x_{(t-d_x)}\right)\right\}, \quad (8.1)$$

where F(.) is a function of Deep LSTM neural network,

$y_{(t)}, y_{(t-1)}, \dots, y_{(t-d_y)}$ are the true values of the time-series output, $x_{(t)}, x_{(t-1)}, \dots, x_{(t-d_x)}$ are the input vectors of Deep LSTM neural network model,

Y_{t+1} is the predicted value of the time-series $y_{(t)}$ at $(t+1)$ time using as inputs the present and previous values of $x_{(t)}$ as well as the true present and previous values of $y_{(t)}$,

d_x is the maximun input delay,

d_y is the maximun output delay.

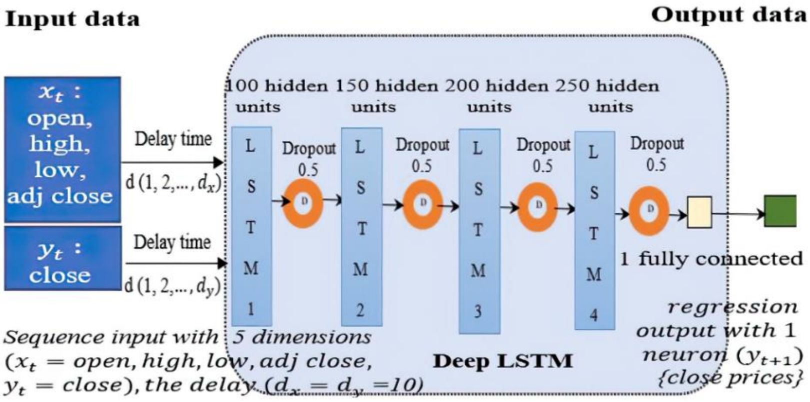

Fig. 8.2 *Architecture of Deep LSTM neural network of the stock price prediction model [27].*

The model Deep LSTM neural network has 11 layers. The details are shown in table 8.2. To fit the effectiveness of Deep LSTM neural network model we created a set of options of parameters for training the networks. The details are shown in table 8.3.

Tab. 8.2 *The details of the Deep LS TM neural network layers.*

Layers	Descriptions
Input layer	The sequence input with a vector of 5 dimensions array (x_t) consists of five features: open, high, low, adj close, and close (y_t). This model will predict the close prices of the next day (Y_{t+1}).
LSTM	Four main layers of LSTM corresponding to the number of neurons for each of LSTM layers in a set (100, 150, 200, 250). The number of epochs in a set is 10, 20, 30, 40, 50.
Dropout layer	Four dropout layers corresponding to a reasonable probability default of 50%.
Fully connected layer	A fully connected layer with a neuron.
Output layer	Activating function of the output layer uses linear as an approximation prediction of the true value with one neuron output for prediction of the close prices $Y_{(t+1)}$.

Tab. 8.3 The parameters of the Deep LSTM neural network model.

Parameters	Values
Optimizer	Adam optimizer
Number of input neuron	5-Dimension of array {open, high, low, adj close, close}
Number of output neuron	1Dimension of array {close price}
Number of iterations	16 (Batch Size is 128)
Number of the time step delay	The delay of $d_x = d_y = 10$.
The number of epochs	10, 20, 30, 40, 50
Number of neurons for LSTM cell of each layer	100, 150, 200, 250
Learning rate /Learning rate schedule	0.001 / Piecewise

8.3 Experimental results of Deep LSTM neural network model

For testing Deep LSTM neural network model, two databases were used: Dow Jones Industrial and S and P 500 index. The model test used the same parameters. Table 8.4 shows the test results of the models for both databases, for a given number of neurons of 100, 150, 200 and 250 and a number of epochs / iterations of 10, 20, 30, 40 and 50 respectively.

***Tab. 8.4** Experimental results of Deep LSTM neural network model with 4 main LSTM layers corresponding to the number of neurons (100,150,200 and 250).*

Historical stock price data name	Epochs/ iterations	Forecast measures			
		RMSE	*MSE*	*MAPE*	*%Accuracy*
Dow Jones Industrial					
	10	***9.2103***	***8.4830***	***0.1637***	***83.62***
	20	9.1710	8.4107	0.1641	83.58
	30	9.4367	8.9052	0.1701	82.98
	40	9.5171	9.0575	0.1716	82.83
	50	12.6853	1.6091	0.2512	74.87
S and P500 Indexes					
	10	9.3511	8.7443	0.1672	83.27
	20	***8.4165***	***7.1673***	***0.1413***	***85.86***
	30	8.8523	7.8673	0.1559	84.40
	40	8.7039	7.5758	0.1534	84.65
	50	9.2103	8.4830	0.1637	83.62

From table 8.4. we can see that Dow Jones Industrial dataset have shown the best results of RMSE, MSE, MAPE and Forecast accuracy for Deep LSTM: 9.2103, 8.4830, 0.1637 and 83.62. On the other hand, in S and P 500 index dataset, the best results of RMSE, MSE, MAPE and Forecast accuracy for Deep LSTM are 8.4165, 7.1673, 0.1413 and 85.86, respectively. Figure 8.3 and figure 8.4 show the best results of Deep LSTM model of the original and the predicted stock closing prices on the testing dataset.

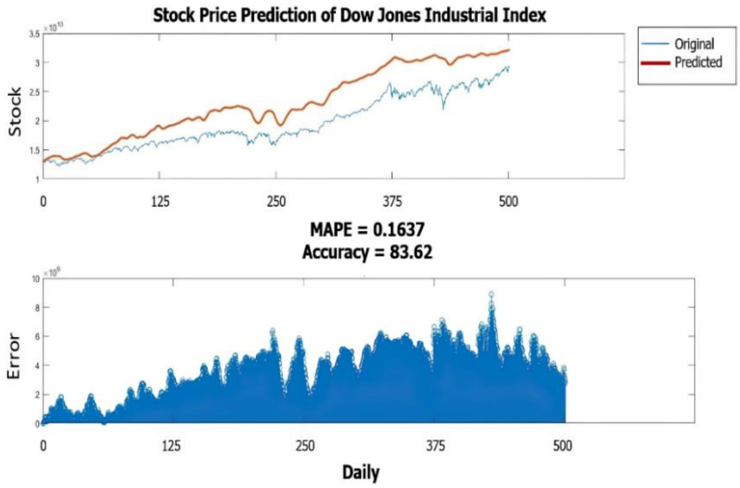

Fig. 8.3 Comparison results of original data and predicted closing prices on testing dataset of Dow Jones Industrial with 83.62% of forecast accuracy[27].

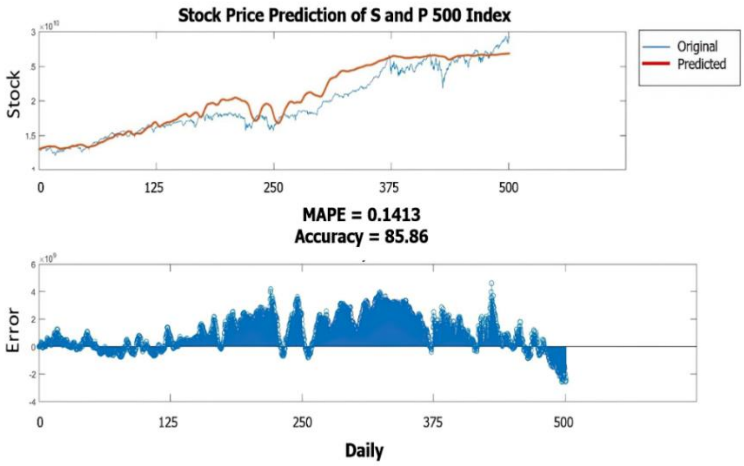

Fig. 8.4 Comparison results of original data and predicted closing prices on testing dataset of S and P 500 index with 85.86% of forecast accuracy[27].

8.4 Comparison results of Deep LSTM neural network model with other models

We compared the best results from proposed Deep LSTM neural network model with all the other research and studies conducted [28], [29] and [30], the predictions being made on the same database. Table 8.5 shows the comparison of the best values of the proposed Deep LSTM neural network model with other research and studies for the Dow Jones Industrial database and table 8.6 shows the comparison of the best values of the proposed Deep LSTM neural network model with other studies conducted for the S and P 500 indexes.

Tab. 8.5 *Results of the proposed Deep LSTM neural network model compared with other models in Dow Jones Industrial.*

Model	Performance Measure	References
	% Accuracy	
Deep LSTM neural network	83.62	Proposed model
Recurrent Embedding Kernel (REK)	71.89	[28]
Events Embedding and Technical Indicators on CNN+LSTM	69.86	[29]

From the results shown in table 8.5., we found that the experimental results of the proposed Deep LSTM neural network model on Dow Jones Industrial leads to the best accuracy of 83.62% when compared to other benchmarks models; REK [28] leads to an accuracy of 71.89% and CNN+LSTM [29] leads to 69.86%, respectively.

Tab. 8.6 *The comparison results of the proposed Deep LSTM neural network model with other models in S and P 500 indexes.*

Model	**Performance Measure**	**References**
	% Accuracy	
Deep LSTM neural network	85.86	Proposed model
Variational Autoencoder LSTM (VAE-LSTM)	84.30	[30]
Recurrent Embedding Kernel (REK)	71.43	[28]
Events Embedding and Technical Indicators on CNN+LSTM	62.02	[29]

The experimental results on S and P 500 index given in table 8.6, show that the proposed Deep LSTM neural network model has the best accuracy results of 85.86% when compared with accuracy results of other reference models. VAE-LSTM model [30] leads to an accuracy of 84.30%, REK model [28] leads to 71.43% and CNN+LSTM model [29] leads to 62.02%, respectively.

Chapter 9 Conclusions and future work

9.1 Conclusions

This thesis consists of nine chapters, each dealing with applied classification techniques and deep learning techniques in conjunction with other techniques in business data analysis:

Chapter 1, this chapter shows the purpose of the article, scope, objectives, and a brief summary of the content of the article.

Chapter 2, this chapter presented the principles and theories, including related work used in this research. The content explaining the meaning of the concepts of the credit scoring, basic information about forecasting of the time-series data, and the data mining principles and summarizes all the collected related studies papers.

Chapter 3, this chapter presented the classification techniques and prediction modes for two types of datasets: credit scoring dataset and time-series dataset. The classification techniques show the details of models using MLP, SVM, C4.5 decision tree. The predictive models used for forecasting financial datasets are NARX neural network model and LSTM. We also explained the performance evaluation of the classify techniques and the forecasting measures.

Chapter 4, this chapter describes the application by applying the feature selection techniques used to classify the credit scoring datasets. This chapter describes the types and operations of feature selections such as PCA and ReliefF algorithm.

Chapter 5, this chapter presented the application of decision tree C4.5 with the AdaBoost technique and MLP to classify in credit scoring data. This chapter has shown the experimental results of each classification technique and comparison the results of the decision tree C4.5 with AdaBoost and MLP techniques.

Chapter 6, this chapter implemented the time-series prediction model with the NARX neural network to predict the currency exchange rates. For this study, the model was designed and tested using daily Thai Baht / US dollar exchange rate. And this model used the main economic indicators and macroeconomic variables, along with the GDP growth rate, interest rate, inflation rate, balance account, and trade balance for input data of the NARX neural network model.

Chapter 7, this chapter has shown the application of the deep learning technique to forecast the currency exchange rate with LSTM neural network model.

Chapter 8, this chapter illustrates the application of deep learning with the LSTM neural network to predict the stock market price. The model used stock price data from Yahoo's financial website. We have chosen two sets of indices Dow Jones Industrial (DJI) and S and P 500 index, both of which are commonly used benchmarks to determine the state of the economy. Overall, to measure the efficiency of the LSTM neural network model, we have compared the best results of the LSTM neural network model with the results of the deep learning that have been researched in the same databases.

9.2 Future work

The implementation and design of machine learning models for all business applications related to the classification and time-series analysis of this thesis are given below. The author hopes that these models can be adapted to accommodate a variety of operations involving data as well. In the future to increase the efficiency of testing models. We can develop the models presented using new techniques:

- As for the credit score classification, the testing of the model found that the feature selection techniques are well suited for the implementation of data classification. For the best experiment results of the model we can apply the deep learning models combined with single classification techniques or a combination of differentiation techniques for better efficiency.
- As we know, there are many kernel functions that can be applied in classify techniques. To develop our SVM models, we can compare our proposed SVM-RBF classification results with other kernel functions such as the Polynomial, linear or sigmoid kernel.
- From various studies related to credit scoring and analysis of time-series data, it was found that the application of the classification model and financial risk forecasting techniques was of paramount importance. These two models are now an important task as financial institutions face more intense competition and challenges. Additionally, all test results suggest that complex models may not always be applicable to real-world situations. Therefore, the widely used evaluation model is used to adapt to various techniques. A key task of future research is to achieve a greater prediction accuracy.

9.3 List of original contributions

[1] **Damrongsakmethee, T.**, and Neagoe, V.-E. (2017). Data Mining and Machine Learning for Financial Analysis. *Indian Journal of Science and Technology*, *10*(39), pp. 1–7. (Scopus, Q4). **DOI number :** 10.17485/ijst/2017/v10i39/119861. **Indian Journal of Science and Technology:** Scopus, Q4.

[2] **Damrongsakmethee, T.**, and Neagoe, V.-E. (2019). Principal Component Analysis and ReliefF Cascaded with Decision Tree for Credit Scoring. *Computer Science On-Line Conference (CSOC2019): Artificial Intelligence Methods in Intelligent Algorithms*, (24-27 April 2019), Springer, vol.985, pp. 85–95. (Springer, ISI). **DOI number:** 10.1007/978-3-030-19810-7_9 **Accession number WOS:** 000503762800009.

[3] **Damrongsakmethee, T**.,and Neagoe, V.-E. (2019). C4.5 Decision Tree Enhanced with AdaBoost Versus Multilayer Perceptron for Credit Scoring Modeling. *Comoutational Methods in Systems and Software (CoMeSySo2019), Computational Statistics and Mathenatical Modeling Methods in Intelligent systems*, (3-5 October 2019), Springer, vol.1047, pp, 216-226. **DOI number:** 10.1007/978-3-030-31362-3_22. **Springer :** It will be indexed in Web of Science.

[4] **Damrongsakmethee, T.,** and Neagoe, V.-E. (2019). A Neural NARX Approach for Exchange Rate Forecasting. *11th International Conference on Electronics, Computers and Artificial Intelligence (ECAI2019)*, Pitesti, Romania (27-29 June 2019), IEEE, pp. 1-6. (IEEE Xplore, ISI) **DOI number:** 10.1109/ECAI46879.2019.9042094. **Accession number WOS:** 000569985400117.

[5] **Damrongsakmethee, T.,** and Neagoe, V.-E. (2020). A Deep Learning Model with Long-Short Term Memory (DLSTM) for Prediction of Currency Exchange Rates. *Computer Science On-line Conference (CSOC2020): Artificial Intelligence and Bioinspired Computational Methods,* (23-26 April 2020), vol.1225, Springer, Cham. pp.484-498.
DOI number: 10.1007/978-3-030-51971-1_40. **Springer :** It will be indexed in Web of Science.

[6] **Damrongsakmethee, T.,** and Neagoe, V.-E. (2020). Stock Market Prediction Using a Deep Learning Approach. *12th International Conference on Electronics, Computers and Artificial Intelligence (ECAI2020),* Bucharest, Romania (25-27 June 2020),IEEE,pp.1-6.
DOI number: 10.1109/ECAI50035.2020.9223142. **IEEE Xplore:** It will be indexed in Web of Science.

Bibliography

[1] J. Marquez, "An Introduction to Credit Scoring For Small and Medium Size Enterprises .," no. February, pp. 1–47, 2008.

[2] A. Tealab, "Time Series Forecasting using Artificial Neural Networks Methodologies : A Systematic Review," *Futur. Comput. Informatics J.*, vol. 3, no. 2, pp. 334–340, 2018, doi: 10.1016/j.fcij.2018.10.003.

[3] A. Tarsauliya, R. Kala, R. Tiwari, and A. Shukla, "Financial Time Series Forecast Using Neural Network Ensembles," *Int. Conf. Swarm Intell.*, vol. 6728, no. February 2015, pp. 480–488, 2011, doi: 10.1007/978-3-642-21515-5.

[4] A. Dingli and K. S. Fournier, "Financial Time Series Forecasting – A Machine Learning Approach," *Mach. Learn. Appl. An Int. J.*, vol. 4, no. 1, pp. 1–27, 2017, doi: 10.5121/mlaij.2017.4302.

[5] K. Chapagain, S. Kittipiyakul, and P. Kulthanavit, "Short-term electricity demand forecasting: Impact analysis of temperature for Thailand," *Energies*, vol. 13, no. 10, pp. 1–29, 2020, doi: 10.3390/en13102498.

[6] T. Damrongsakmethee and V.-E. Neagoe, "Data Mining and Machine Learning for Financial Analysis," *Indian J. Sci. Technol.*, vol. 10, no. 39, pp. 1–7, 2017, doi: 10.17485/ijst/2017/v10i39/119861.

[7] V. Caggese, "Forecasting Exchange Rates : a Comparative Analysis Faculty of Economics," *Int. J. Bus. Soc. Sci.*, vol. 3, no. 10, pp. 145–156, 2012.

[8] Y.-A.-K. Al-hroot, "Bankruptcy Prediction Using Multilayer Perceptron Neural Networks In Jordan," *Eur. Sci. J.*, vol. 12, no. 4, pp. 425–435, 2016, doi: 10.19044/esj.2016.v12n4p425.

[9] P. Marius-Constantin, V. E. Balas, L. Perescu-Popescu, and N. Mastorakis, "Multilayer perceptron and neural networks," *WSEAS Trans. Circuits Syst.*, vol. 8, no. 7, pp. 579–588, 2009.

[10] J. Han, M. Kamber, and J. Pei, *Data Mining Concepts and Techniques*. Elsevier Inc., 2012.

[11] A.-D. Pietersma, "Feature Space Learning in Support Vector Machines through Dual Objective optimization Auke-Dirk Pietersma," 2010.

[12] J. R. Quinlan, "Induction of Decision Trees," *Mach. Learn.*, pp. 81–106, 1986.

[13] J. R. Quinlan, "Improved Use of Continuous Attributes in C4 . 5," *J. Artif. Intell. Res.*, vol. 4, no. March, pp. 77–90, 1996.

[14] Z. Boussaada, O. Curea, A. Remaci, H. Camblong, and N. M. Bellaaj, "A nonlinear autoregressive exogenous (NARX) neural network model for the prediction of the daily direct solar radiation," *Energies*, vol. 11, no. 3, 2018, doi: 10.3390/en11030620.

[15] J. Buitrago and S. Asfour, “Short-Term Forecasting of Electric Loads Using Nonlinear Autoregressive Artificial Neural Networks with Exogenous Vector Inputs,” pp. 1–24, 2017, doi: 10.3390/en10010040.
[16] D. Shah, W. Campbell, and F. H. Zulkernine, “A Comparative Study of LSTM and DNN for Stock Market Forecasting,” in *2018 IEEE International Conference on Big Data (Big Data)*, 2018, pp. 4148–4155, doi: 10.1109/BigData.2018.8622462.
[17] W. S. Jhang, S.-E. Gao, C.-M. Wang, and M.-C. Hsieh, “Share Price Trend Prediction Using Attention with LSTM Structure,” in *20th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, (SNPD 2019)*, 2019, pp. 208–211, doi: 10.1109/SNPD.2019.8935806.
[18] J. S. Armstrong, “Evaluating Forecasting Methods,” in *Principles of Forecasting: A Handbook for Researchers and Practitioners*, 2001, pp. 1–26.
[19] M. Shcherbakov, A. Brebels, and A. Tyukov, “A Survey of Forecast Error Measures,” *World Appl. Sci. J. 24*, no. March 2016, pp. 171–176, 2013, doi: 10.5829/idosi.wasj.2013.24.itmies.80032.
[20] T. Damrongsakmethee and V.-E. Neagoe, “Principal Component Analysis and ReliefF Cascaded with Decision Tree for Credit Scoring,” *Comput. Sci. On-line Conf. Artif. Intell. Methods Intell. Algorithms*, vol. 2, no. May 2019, pp. 85–95, 2019, doi: https://doi.org/10.1007/978-3-030-19810-7_9.
[21] T. Damrongsakmethee and V.-E. Neagoe, “c4.5 Decision Tree Enhaned with AdaBoost Versus Multilayer Perceptron for Credit Scoring Modeling,” in *3rd Comoutational Methods in Systems and Software*, 2019, pp. 216–226.
[22] “Exchange rate,” *Bank of Thailand*, 2019. https://www.bot.or.th/english/_layouts/application /exchangerate/exchangerate.aspx (accessed Apr. 09, 2019).
[23] C. L. Cocianu and M.-S. Avramescu, “New Approaches of NARX-Based Forecasting Model. A Case Study on CHF-RON Exchange Rate,” *Inform. Econ.*, vol. 22, no. 2/2018, pp. 5–13, 2018, doi: 10.12948/issn14531305/22.2.2018.01.
[24] T. Damrongsakmethee and V.-E. Neagoe, “A Neural NARX Approach for Exchange Rate Forecasting,” in *11th Edition Electronics, Computers and Artificial Intelligence (ECAI), Pitesti, “Romania,”* 2019, pp. 1–6.
[25] “Historical Foreign Exchange Rates,” *Bank of Thailand*, 2019. https://www.bot.or.th/english /_layouts/application/exchangerate/exchangerate.aspx. (accessed Sep. 05, 2019).
[26] T. Damrongsakmethee and V. E. Neagoe, “A Deep Learning Model with Long Short-Term Memory (DLSTM) for Prediction of Currency Exchange Rates,” in *Computer Science On-line Conference (CSOC 2020): Artificial Intelligence and Bioinspired Computational Methods*, 2020, pp. 484–498.
[27] T. Damrongsakmethee and V. E. Neagoe, “Stock Market Prediction Using a Deep Learning Approach,” in *12th International Conference on Electronics, Computers and Artificial Intelligence (ECAI2020)*, 2020, pp. 1–6.
[28] L. Le and Y. Xie, “Recurrent Embedding Kernel for Predicting Stock Daily Direction,” in *5th IEEE/ACM International Conference on Big Data Computing, Applications and Technologies (BDCAT 2019)*, 2019, pp. 160–166, doi: 10.1109/BDCAT.2018.00027.
[29] P. Oncharoen and P. Vateekul, “Deep Learning for Stock Market Prediction Using Event Embedding and Technical Indicators,” in *5th International Conference on Advanced Informatics: Concepts Theory and Applications (ICAICTA 2018)*, 2018, pp. 19–24, doi: 10.1109/ICAICTA.2018.8541310.
[30] C. Gu, H. Hsieh, C. Wu, R. Chang, and J. Ho, “A Fund Selection Robo-Advisor with Deep-learning Driven Market Prediction.,” in *IEEE International Conference on Systems, Man and Cybernetics (SMC 2019)*, 2019, pp. 2845–2850.