International Journal of Engineering & Technology, 7 (4.15) (2018) 400-404

SPC

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper



FoRex Trading Using Supervised Machine Learning

Thuy Nguyen Thi Thu[1]*, Vuong Dang Xuan[2]

[1]*ThuongMai University, Mai Dich, Cau Giay, Hanoi, Vietnam*
[2]*Sonla Gifted HighSchool, Sonla, Vietnam*
**Corresponding author E-mail: Thuynguyenthithu@tmu.edu.vn*

Abstract

The exchange rate of each money pair can be predicted by using machine learning algorithm during classification process. With the help of supervised machine learning model, the predicted uptrend or downtrend of FoRex rate might help traders to have right decision on FoRex transactions. The installation of machine learning algorithms in the FoRex trading online market can automatically make the transactions of buying/selling. All the transactions in the experiment are performed by using scripts added-on in transaction application. The capital, profits results of use support vector machine (SVM) models are higher than the normal one (without use of SVM).

Keywords: *Classification; Foreign Exchange rate; Supervised Machine Learning; Transaction.*

1. Introduction

Supervised machine learning can apply into many areas in computer science such as decision making, forecasting, and specially is to predict stock price or money exchange rate, and so on. Supervised machine learning means the teaching and monitoring the computer to classify or cluster data from observed data set. With the explosion of data nowadays, particular with business big data, supervised machine learning techniques can support people to look depth and analyze their data in order to extract useful information for their company's purposes. In stock market, the transactions of money exchange can be predicted rise or fall with helping of machine learning and the observed data set collected in the past. The artificial intelligent systems are used with supervised machine learning methods like Logic, Perceptron, Statistics methods (Bayes Networks, Instance based) [2]. The goals are the classification or regression dataset into alternative classes so-called as outputs. These target variable (y) is usually the nominal (discrete) or continuous values [1]. For simplicity, the output of y in the classification problem can be seen as binary values such as 0 or 1.

Support Vector Machines (SVMs), which is one of Supervised Machine Learning techniques, are the techniques that separate data in hyperplane space into two data classes [3]. These techniques create maximized margins and the distance between them to separate data into alternative sides of hyperplane [2]. So, SVMs can be a possible candidate of use to predict Foreign Exchange (Forex) trend (Up or Down) in money exchange rate problems. The predicting FoRex price is performed by using historical data such as Open", "Close", "Low" and "High" in different time ranges of FoRex transactions.

Alternative experiments with different support vector machine models are performed in this paper. These are to show the advantage of using them in FoRex rate prediction. The detail research on SVM in FoRex market can be seen in [27]. The analysis of FoRex rate prediction can be seen in section 3. This shows some specific concepts which contribute to the predicted trend of FoRex rate in the trading market. These specific terms are the basic explanations for the classification output results.

To be more generality, a framework of supervised machine learning (SML) techniques in predicting FoRex rate has been shown in section 4. Based on this framework, a representative SVM model, which is proposed from research experiment in [27], is chosen to install into expert advisor (robotics) in FoRex transaction software to make the actual FoRex transactions with demo account.

2. Related Works

The Foreign Exchange problems can be seen as the ones of predicting FoRex rates (up or down) of each currency pairs. Therefore, they can be seen as classification problems where their outputs (FoRex rates) are the binary values. There are some previous techniques such as Auto Regressive Integrated Moving Average (ARIMA) used for predicting time series data. According to [6], ARIMA, however, is an unvaried model in general. Moreover, the techniques are performed with the assumption of the linear and stationary time series.

According to [7], the noisy date in the FoRex prediction problems made them to be challenged issues in time series forecasting field. Some publications in [7], [8], [6], [9] used RBF, and MLP, and gene to predict the FoRex rates.

The works, for instance, in [6] used Multilayer Perceptron Network (MLP) with the training set, testing set of 1600 and 225 over 1825 instances. The disadvantage in the work is a use of small data set. This might cause the quick regression or mis-predicting output result. Another work of Rumani stock with MLP models can be seen in [10]. The publication works shown the advantages in use supervised learning models in forecast a value of rise or fall for stock price.

There are some ANN and decision tree models which have been used with forecast the uptrend or downtrend in stock market [11-14]. Although these works reported about their abilities to work

Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

with real time data, however, their experiments have reported of lacking of using high frequency data [15].

The nonlinear regression problems, especially in predicting things, can be solved with SVMs models. There are some publications which have been successful in use nonlinear regression predictions such as in 16], [17] [23], [24], [25]. However, there are limited publications used SVMs models for financial time-series forecasting [19], [20] although the SVMs models can perform with nonlinear problems and they have been reported as the advantage techniques. Therefore, continue to the research in [27], the representative supervised machine learning method of SVM is selected to estimate the uptrend or downtrend of each currency rates. It is because of SVMs can avoid some issues such as over-fitting problems (like ANN in [1]), or no parameters to be harmony except the upper bound C [21].

3. Research Background

3.1. Supervised Support Vector Machine

The Support Vector Machines (SVMs) can be divided in to linear or nonlinear SVMs. These are based on linear or nonlinear hyperplane lines which separate the training data set in hyperplane space. The linear form in linear hyperplane in explained as following:

$H: w.x + b = 0$

Where $\mathrm{w} = (\mathrm{w}_1, \mathrm{w}_2, \ldots, \mathrm{w}_\mathrm{d}) \in \mathbb{R}^\mathrm{d}$ and $\mathrm{x} = (\mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_\mathrm{d}) \in \mathbb{R}^\mathrm{d}$ then $\mathrm{w}.\mathrm{x} = \sum_{\mathrm{i}=1}^{\mathrm{d}} \mathrm{w}_\mathrm{i}.\mathrm{x}_\mathrm{i}$

Assume that $y_i \in \{+1, -1\}$, $i = \overline{1,n}$. y_i is labelled as output class. Therefore, outputs can be formed as:

$$y_i = \begin{cases} 1 & if \quad w.x_i + b \geq 0 \\ -1 & if \quad w.x_i + b < 0 \end{cases}$$

The hyperplane H need separate the data onto alternative classes. So, SVMs have to build two hyperplane as: $H_1: w.x + b = 1$; and $H_2: w.x + b = -1$ with no existing x_i in H_1 and H_2, & margin of H_1 and H_2 is max.

In some cases, the SVM need to use one special margin so-called as soft margin to separate the data set into two classes including some mis-classified training instances [4]. This is because SVM can not find any separating hyperplane which can separate exactly the data set.

It seems to be difficult to solve with nonlinear training dataset. The solution is that the data is mapped onto a higher dimensional space so-called as feature space in where the linear hyperplane can separate the data set in to two classes.

By using kernel function, any new points in feature space are represented for the data in the given data set [5].

Assume that $E(x, y)$ is predicting error rate for each pair of (x, y), selection of α_i, α_jwill be produced in two heuristic loops so that $|E_i - E_j|$ will max.

3.2. Foreign Exchange Rate Prediction

The FoRex market has its own characteristics of high benefit. In there, the transaction can be performed online via the internet services. This means that the investors can request their transactions at any time and from anywhere in the world. The buying and selling can be done with any currencies pairs such as EUR/USD, EUR/JPY, or GBP/EUR etc. in order to make profit.

A term of PIP (Price Interest Point) is defined in order to measure the profit or loss in the FoRex. A PIP is the smallest price change that a given exchange rate can make. Every money pair has a corresponding value (a quote). For example, a quote of 1.2776 is the exchange rate of every EUR is USD 1.2776 (it costs the trader of 1.2776 to buy a single EUR). If the rate is lightly moved from 1.2776 to 1.2777, it shows 1 PIP increase.

The PIP is calculated as:

$$PIP_{value} = PIP/rate \times Amount\ being\ purchased$$

For example, if a trader purchase 10,000 units with the rate of EUR/USD of 1.2776. Therefore, $\mathrm{PIP}_{\mathrm{value}}$= 0.0001/1.2776 x 10,000 = 0.7827 EUR.

The concept of the transaction trend describes the direction of the market moving as uptrend or downtrend. In the uptrend, a resistance level means there is a change in supply and demand (more sellers than buyers) whereas in a downtrend, there is a concept of support level, in which there are more buyers than sellers. When the trend is ended, there are two possible outcomes of opposite starting trend or in range market (stop at certain resistance or support levels). The task for the Expert Advisor (in the experiment in later section) is to determine the trend, calculate the time of lasting and decide if the timing is right to place a good transaction of Bid/Ask.



Fig.1: Illustration of Uptrend or Downtrend

To predict target of exchange rate as uptrend or downtrend, the following parameters are defined.

- Target Profit (TP) shows the target profit (positive number).
- Stop Loss (SL) indicates the acceptable loss (positive number).
- Time Trend Detection (T) indicates the time starting from present stage (positive number).

Hence, current rate (P), the trend of rate can be as:

- up if the rate of (P+TP) is increased, and there is a decreasing of an amount of (P-SL). All these are calculated in the time T.
- Low when the (P-TP) is decreased, and an amount of P+SL is increased. All these are calculated in the time T.

4. Supervised Machine Learning Prediction Framework

The supervised machine learning prediction frame work can be seen in Figure 2, which is based on data mining process.

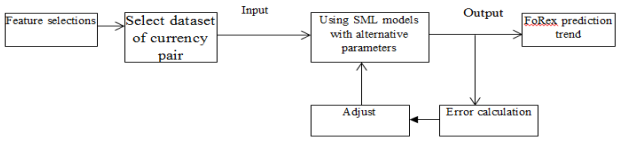


Fig.2: Supervised machine learning prediction framework

The detail as follows:

Step 1: Feature Selection. Collected data set might include unnecessary features for predicting process. Therefore, these features are removed.

Step 2: Select dataset of currency pair. In this step, a defined pair of money eg. EUR vs USD is chosen. Also in this step, the data is taken from alternative candles, time frames etc. in order to be easily used in installation with Robotics.

Step 3: Processing data with supervised machine learning (SML) models (here is support vector machine models). Dataset is divided into training, testing and validation sets. All data in mapped on the range of [0,1].

Step 4: Evaluation models with assessment metrics of mean square errors (MSE).

Step 3,4 is repeated many time in order to choose a model with minimum MSE.

The output is a certain definition of trend (up or down).

5. Experiment

5.1. Data analysis and Model configuration

The experimental data is collected from 01/01/2013 to 30/09/2016 by using MetaTrader 4 [22]. For general use, the pair of EUR/USD has been chosen. Two data sets are created with the one of training of D (including the instances in a time of 01/01/2013 to 31/12/2015), and one of testing D' (including data in a period 1/01/2016 to 30/09/2016).

The features are chosen basing on the profits. In this paper, the profit is taken of bout 10 – 15 Pips in each transaction. These selections might effect to the experimental results. The input features for samples exchange rates "Open", "Close", "Low" and "High". These rates are taken in alternative candles in the window time (time frame). There are 1 minute, 5 minutes, 15 minutes,1 hour, 1 day according to M1, M5, M15, H1, D1 respectively. Other features are taken from Custom Indicators, Bollinger Bands, RSI and MA.

To continue with the experiment in [27], 196 features are chosen (in Table 1). The addition more features (increase from 137 to 196) for the classification is for expecting to have better results in the classification process and having better transactions. This means data is taken from an extension of candlesticks and timeframe. This is to concentrate more detail in specific historical transaction data.

Table 1: Feature selections

Feature#	*Data*	*No*	*Explanations*
#1-20	O , H , L , C on M1	20	04 value X 5 nearest candles
#21-40	O , H , L , C on M5	20	04 value X 5 nearest candles
#41-56	O , H , L , C on M15	16	04 value X 4 nearest candles
#57-72	O , H , L , C on H1	16	04 value X 4 nearest candles
#73-92	O, H, L, C on D1	20	4 values x 5 nearest candles
#93-97	Time data	5	Time attributes
#98-105	RSI(7) on M5, M15	8	2 values x 4 Timeframe
#106-113	RSI(14) on M5, M15	8	2 values x 4 Timeframe
#114-145	MA(9), MA(12), MA(100), MA(200)	32	8 values x 4 Timeframe
#146-165	Custom Indicator	20	8 PAX + 12 MKC
#166-174	Bollinger Bands	9	3 Lines x 3 Timeframe
#175-180	Average True Range	6	2 values x 3 Timeframe
#181-196	Highest, Lowest M1, M5	16	4 values x 4 Timeframe

Model configurations

The SVMs models are chosen in the experiments. The models parameters are used the same as using in [27]. Therefore, the model parameters are:

- RBF: $k(x,y) = e^{-\frac{\|x-y\|^2}{2\sigma}} = e^{-\gamma\|x-y\|^2}$ with $\gamma \in [0;5]$
- Polynomial: $k(x,y) = (x.y + \theta)^d, \theta \in \mathbb{R}, d \in \mathbb{N}^*$ with $d = 2,3,4$ and $\theta \in [0;1]$.
- The parameter of C has been used with $C \in [1;10]$.

Table 2: Experimental Model Configurations

Models	Poly 1	Poly 2	Poly 3	GsRBF 1	GsRBF 2	GSRBF 3
C	1.0	1.0	1.0	2.0	1.0	2.0
Kernel	Poly	Poly	Poly	RBF	RBF	RBF
Parameters	Power 2	Power 3	Power 3	γ=5.0	γ=2.0	γ=2.0
No of vectors	2500	3000	3972	3000	4500	5000

5.2. Classification results

Data set D data can be separated to two parts D_{pos} and D_{neg} positive and negative output respectively. Which can be further split to $D_{pos}^1, D_{pos}^2, \dots, D_{pos}^k$ and $D_{neg}^1, D_{neg}^2, \dots, D_{neg}^k$ sub-sets with each sub-sets has $\frac{1}{k}|D_{pos/neg}|$ samples. D_Test set is taken with each sub-set. D_Train includes k-1 remaining samples in the sub-sets. In here, k is 5.

Following Table 3 & 4 show experimental results with 6 alternative models. The performance between alternative models.is evaluated and compared by using the Accuracy Exactness Rate, Precision of Positive, Negative, Micro-Average and Macro-Average individually.

Table 3: Results of RBF Configuration Model

Kernel: Gaussian RBF Features: 196		**GsRBF1**		**GsRBF2**		**GSRBF3**	
		D_Train	D_Test	D_Train	D_Test	D_Train	D_Test
Accurate		84.22%	55.80 %	83.14%	58.14 %	83.81%	58.17 %
Precision	**Positive**	84.05%	55.59 %	82.59%	57.55 %	83.12%	57.56 %
	Negative	84.39%	55.98 %	83.71%	58.70 %	84.52%	58.76 %
	Micro-Average	84.22%	55.80 %	83.14%	58.14 %	83.81%	58.17 %
	Macro-Average	84.22%	55.79 %	83.15%	58.13 %	83.82%	58.16 %

Table 4: Poly Configuration Models Results

Kernel: Polynomial Features:196		**Poly1**		**Poly2**		**Poly3**	
		D_Train	D_Test	D_Train	D_Test	D_Train	D_Test
Accurate		86.15%	74.20%	86.15%	74.35%	86.25%	74.32%
Precision	**Positive**	86.73%	80.29%	86.58%	80.18%	86.80%	80.21%
	Negative	85.59%	70.36%	85.73%	70.62%	85.72%	70.56%
	Micro-Average	86.15%	74.20%	86.15%	74.35%	86.25%	74.32%
	Macro-Average	86.16%	75.32%	86.16%	75.40%	86.26%	75.38%

5.3. Expert Advisor experimental results

The Poly2 SVM model is chosen for installation into Robotics (Expert Advisor- EA) in MetaTrader 4 [30]. The script (written in

EA) has performed with a collection of observed data set. The information incorporates some specialized markers and exchanges of purchase or offer FoRex match of USD and EUR.

For correlation with the utilization directed machine learning model in Robotics, the typical exchange (test of Robotics without utilizing backing of SVM) additionally has been performed.

The results can be seen in Table 5. The bold numbers in the table show the comparison between the transactions use or not use SVM.

Table 5: Results Indicators'.

Index	*Indicators*	*No_SVM*		*SVM*	
1	Net Profit		**328.73**		**462.5**
2	Gross Profits		1238.6		713.93
3	Gross Loss		-909.87		-251.44
4	Profit rate		**1.36**		**2.84**
5	Growth		**32.9%**		**46.3%**
6	Average Growth (in a month)		**3.21%**		**4.31%**
7	Strict Drawdown		155.07		17.49
8	Fair Drawdown	436.23	**34.05 %**	315.62	**24.31 %**
9	Volume		188		94
10	Profit Transactions	139	73.9%	76	80.9%
11	Loss Transactions	49	26.1%	18	19.2%
12	Ask Profit Transactions	86	95.4%	49	100.0 %
13	Bid Profit Transactions	102	55.9%	45	60.0%

5.4. Discussion

Alternative models results with different kernel functions (Gaussian RBF, Polynomial) can be seen in Table 3 and Table 4. There are big gap results of test set compared to the results of training set (see in table 3, about 57% and 83.8 % respectively). The advantage use of SVM in the transaction Robotics can be seen in Table 5. The indicators' results show that the net profit of using SVM in add-on robot is higher than the one without using SVM. For instances, the number of samples for the script using support vector machine is half of number of SVM (94 & 188) and its profit rate can achieve of 2.84 compared to 1.36.

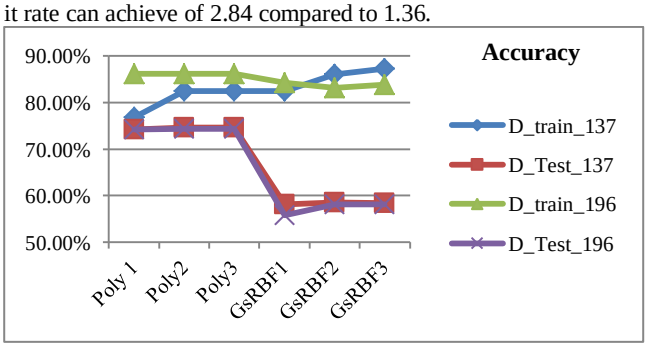


Fig 3: Comparison between 196 features-models and 137 features-models in [27].

The accuracy rates of models using with 137 features in [27] are used to compare to the ones using with 196 features in this research (see in Figure 3). In general, there are not big gap of rates between these models. This means the increase of features does not have better improvement in the accuracy rate.

For more comparison, the fair drawdown values, which shows the rate of capital reduction, taken from [27] and this research are represented on the chart (Figure 4).

It is clear from Figure 4 that the Robotics using SVM with 196 features can prevent the capital reduction compared to the normal Robotics' transactions. In other words, more features are used in SVM, the profit loss can be better control.

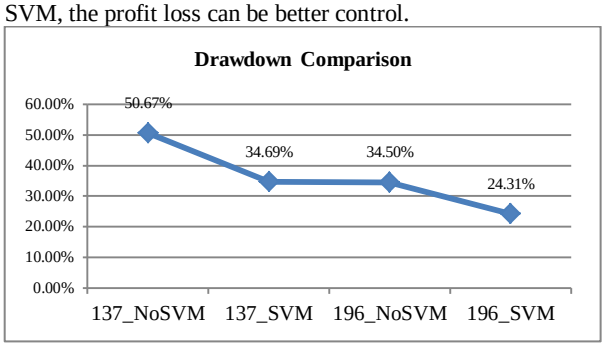


Fig. 4: Drawdown rate comparison between 137 features model [27] and 196 features model.

6. Conclusion

The support vector machine can help to forecast FoRex trend of up or down. This is because the outputs of FoRex rate are received as binary values (uptrend or downtrend). The FoRex problems now can be seen as classification ones which can be solved by supervised learning particular with SVMs models. The experiment has shown the power of SVMs in performing a huge, complex and time series data in order to help the investors with their transactions decisions.

The experiment has been performed with alternative kernel functions (Gaussian and Polynomial). Elective models with elective parameters additionally have been per-framed to win the higher outcomes. The consequences of utilizing Polynomial Kernel demonstrated the best outcomes contrasted with models utilized Gaussian when all is said in done.

The picked model to introduce in Robotic exchange has a place with the best Poly show in above analyses. Its exchange results can acquire a larger number of benefits than the conventional one (without use of SVM). Therefore, an increase number of features (from 137 to 196) compared to [27] has shown the better profit lost models can get (reduce the drawdown rate). This means that although there has been no different in classification accuracy rates (Figure 3), but the more concentrating on specific data the better controlling of loss we can get.

Acknowledgement

Many thanks to ThuongMai University and research group for supporting the research.

References

[1] Witten IH, Frank E, Hall MA (2011), Data mining: practical machine learning tools and techniques, *Morgan Kaufmann Publishers*.
[2] Kotsiantis SB (2007), Supervised Machine Learning: A Review of Classification Techniques, *Informatica: 31*, 249-268.
[3] Vapnik V (1995), The Nature of Statistical Learning Theory, Springer Verlag.
[4] Veropoulos K, Campbell C, Cristianini N (1999), Controlling the Sensitivity of Support Vector Machines, *In Proceedings of the International Joint Conference on Artificial Intelligence*, (IJCAI99).
[5] Scholkopf C, Burges JC, Smola AJ (1999), Advances in Kernel Methods, MIT Press.
[6] Baasher Aa, Mohamed WF (2010), FOREX Trend Classification using Machine Learning Techniques, *Arab Academy for Science and Technology*, 41-47.
[7] Deboeck GJ (1994), Trading on the Edge: Neural, Genetic and Fuzzy Systems for Chaotic Financial Markets, *New York Wiley*.
[8] Box GEP, Jenkins GM (1970), Time Series Analysis: Forecasting and Control, *Holden- Day, San Francosco, CA.*

[9] Punniyamoorthy M, Thoppan JJ (2013), ANN-GA based model for stock market surveillance, *Journal of Financial Crime*, vol. 20, no. 1, 52-66.
[10] Nemeș Dm, Butoi A (2013), Data Mining on Romanian Stock Market Using Neural Networks for Price Prediction, *Informatica Economică*, vol. 17, no. 3,125-136.
[11] Huang W, Nakamori Y, Wang SY (2004), Forecasting foreign exchange rates with artificial neural networks, a review. *International Journal of Information Technology & Decision Making*, Vol 3, No 1, 145-165.
[12] Huang CL, Tsai CY (2010), A hybrid SOFM-SVR with a filter based feature selection for stock market forecasting, *Expert Systems with Applications*, 36 (2, Part 1), 1529 – 1539.
[13] Tsai CF, Hsiao YC (2010), Combining multiple feature selection methods for stock prediction: Union, intersection, and multi-intersection approaches, *Decision Support Systems*, 50(1), 258 – 269.
[14] Shen W, Guo X, Wu C, Wu D (2011), Forecasting stock indices using radial basis function neural networks optimized by artificial fish swarm algorithm, *Knowledge-Based Systems*, 24(3), 378–385.
[15] Reboredo M, Rubio JJ, Nonlinearity in forecasting of high-frequency stock returns, *Computational Economics*, 40(3), 245–264.
[16] Müller KR, Smola A, Rätsch G, Schölkopf B, Kohlmorgen J, and Vapnik V, (1997), Predicting time series with support vector machines, *In proceedings of the IEEE workshop on neural networks for signal processing 7*, 511 – 519.
[17] Mukherjee S, Osuna E, Giroso F (1997), Nonlinear prediction of chaotic time series using SVM. *In proceedings of the IEEE workshop on neural networks for signal processing 7*, 511 – 519.
[18] Kim K (2003), Financial time series forecasting using Support Vector Machines, *Neurocomputing 55*, 307- 319.
[19] Tay EFH, Cao L (2001), Application of support vector machines in Financial time series forecasting, *Omega 29*, 309–317.
[20] Liu L, Wang W (2008), Exchange Rates Forecastiing with Least Squares SVM, *International Conference on Computer Science and Software Engineering*, 2008.
[21] Drucker H, Wu D, Vapnik VN, Support vector machines for spam categorization, *IEEE Trans. Neural Networks 10* (5),1048–1054
[22] Metatrader 4, (2016). Website: http://www.metaquotes.net/
[23] Kamruzzaman J, Sarker RA, Ahmad I (2003), SVM Based Models for Predicting Foreign Currency Exchange Rates, *Third IEEE International Conference on Data Mining*, ICDM 2003, 557 – 560.
[24] Sidehabi SW, Indrabayu, Tandungan S (2016), Statistical and Machine Learning approach in forex prediction based on empirical data, *IEEE International Conference on Computational Intelligence and Cybernetics, CYBERNETICSCOM 2016*, 63 – 68.
[25] Sharma DK, Hota HS, Handa R (2017), Prediction of foreign exchange rate using regression techniques" *Review of Business and Technology Research*, Vol. 14, No. 1, 29-33.
[26] Ramakrishnan SK, Butt S, Chohan MA, Ahmad H (2017), Forecasting Malaysian exchange rate using machine learning techniques based on commodities prices, *IEEE International Conference on Research and Innovation in Information Systems*, ICRIIS 2017.
[27] Thuy NTT, Vuong DX (2018), Using Support Vector Machine in FoRex Predicting, IEEE *International Conference on Innovative research and development*, ICIRD 2018.