ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/353503146

The Influence of Sentiments in Digital Currency Prediction Using Hybrid Sentiment-based Support Vector Machine with Whale Optimization Algorithm (SVMWOA)

Conference Paper · July 2021

DOI: 10.1109/ICOTEN52080.2021.9493454

CITATIONS
4

READS
73

4 authors, including:

Nor Azizah Hitam
International Islamic University Malaysia
9 PUBLICATIONS **79** CITATIONS

SEE PROFILE

Amelia Ritahani Ismail
International Islamic University Malaysia
67 PUBLICATIONS **248** CITATIONS

SEE PROFILE

Omair Ameerbakhsh
Taibah University
17 PUBLICATIONS **53** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project Artificial Intelligence View project

Project Astroglial networks roles in information processing and plasticity View project

All content following this page was uploaded by Nor Azizah Hitam on 08 August 2021.

The user has requested enhancement of the downloaded file.

The Influence of Sentiments in Digital Currency Prediction Using Hybrid Sentiment-based Support Vector Machine with Whale Optimization Algorithm (SVMWOA)

Nor Azizah Hitam
Department of Computer Science
International Islamic University Malaysia (IIUM)
Kuala Lumpur, Malaysia

Amelia Ritahani Ismail
Department of Computer Science
International Islamic University Malaysia (IIUM)
Kuala Lumpur, Malaysia
amelia@iium.edu.my

Ruhaidah Samsudin
Department of Information Systems
Faculty of Computing, Universiti Teknologi Malaysia,
Skudai, Johor, Malaysia
ruhaidah@utm.edu.my

Omair Ameerbakhsh
Department of Information Systems
College of Computer Science and Engineering
Taibah University
oameerbakhsh@taibahu.edu.sa

2021 International Congress of Advanced Technology and Engineering (ICOTEN) | 978-1-6654-1224-7/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICOTEN52080.2021.9493454

Abstract*— Getting an accurate prediction of a digital currency, also known as a cryptocurrency price index, becomes a significant factor in helping investors make the right decision. Failure to predict the movement of the crypto market gives a huge impact on profit loss. The difficult part is that market is dynamic in a way that is driven by many factors including inflation rate, economics, and natural calamities. This creates a chaos in the price of index so does the sentiment of the investor. This study proposes a machine learning model that applies a combination of sentiment-based support vector machine that is optimized by the whale optimization algorithm for predicting the daily price of a digital currency. Support Vector Machine (SVM) technique is used with the Whale Optimization Algorithm (WOA) which is inspired by the swarm optimization algorithms. The proposed Hybrid Sentiment-based Support Vector Machine with a Whale Optimization Algorithm (SVMWOA). will be evaluated and compared based on performance measures. The proposed method is compared with Support Vector Machine Optimized by Genetic Algorithm (SVMGA) and the Support Vector Machine Optimized by Harmony Search (SVMHS). The proposed model is found robust to be used in other fields of study.

Keywords— Cryptocurrency, prediction, hybrid sentiment-based Support Vector Machine (SVMWOA), Support Vector Machine (SVM), Whale Optimization Algorithm (WOA)

INTRODUCTION

Support Vector Machine (SVM) has been applied in many field of studies such as agriculture, finance, biometrics, cheminformatics [1], [2] and is one of the eminent machine learning techniques for solving the problem with classification. The classification model train the model using the training data, and the model will be used to classify the unidentified sample. In SVM, the identified parameters that play a very important part in its performance are known as penalty parameter and kernel parameters.

However, SVM underperforms by a large dataset and suffers when the data set has more noise. Also, traditional models unable to capture non-linear patterns of very complicated problems. Thus, utilizing human sentiments in more sophisticated machine learning models is needed to improve the results. There are many optimization algorithms used in solving the optimization problems including the parameter tuning problems. [3] proposed the Particle Swarm Optimization (PSO) with SVM to tune the parameters to predict the daily price of stocks index while Cuckoo Search Optimizer (CS) is utilized by [4] in the same field of studies. The combination of PSO and External Optimization (EO) is used to improve the integrated structure of numerous boilers in the Thermal Power Plants [5] Moreover, [6] applied a hybrid of PSO and Moth Flame Optimization algorithms (MFO) in solving the well-known benchmark functions.

Whale optimization algorithm (WOA) is a swarm-based metaheuristics algorithm of humpback whales that works based on the bubble net hunting movement method in solving a complicated optimization problem. In WOA, due to its fast convergence speed, very minimal required operator, and having better balancing ability between exploration and the exploitation phases, WOA is widely accepted as a swarm intelligence technique that applies in various engineering fields including financial and trading investment fields. Cryptocurrency is knowns as a very high fluctuates over time and it is very much challenging to forecast its future movement.

In this study, we propose a method of hybridization of SVM and WOA by merging the advantages, both the practicality of SVM and WOA are inherited and the search capability and proficiency are improved.

This paper is organized as follows: Section 2 discussed the introduction and the aim of this paper. Section 3 describes the pieces of literature on SVM, WOA algorithm, and the social media feed. The methodology of this paper is explained Section 4. Data description is explained in Section 5 and followed by the experimental results in Section 6. Finally, conclusions and some future research guidelines are presented in Section 7.

ARTIFICIAL INTELLIGENCE (AI) IN FINANCE, TRADING, AND INVESTMENT

Forecasting market behavior to make the greatest decision which will give much profit becomes one of the most challenging part for stakeholders and investors. Cryptocurrency or a digital currency, decentralized in its nature and is operating based on the peer-to-peer transactions [7]. Investors usually questions about ways of helping them to predict the incoming market price based on the available historical data. To answer this question, predictive model will train sets of selected cryptocurrency data within the specific period. Moreover, research on the cryptocurrency is still very low, as most research is focusing one cryptocurrency rather than broader fields such as the participation

978-1-6654-1224-7/21/$31.00 ©2021 IEEE

Authorized licensed use limited to: International Islamic University Malaysia. Downloaded on August 08,2021 at 12:29:21 UTC from IEEE Xplore. Restrictions apply.

of the government in its regulation, or market improvement including the technology enhancement. Artificial Intelligence (AI) turns to be prominent in Finance starting in the 1980s. It begins with the application of Artificial Neural Network Systems to detect abnormal claims/charges which will then flag into the human investigation. In 1987, the Fraud Prevention Task Force was implemented by the Security Pacific National Bank to overcome the unauthorized usage of debit cards Stewart and Watson (1985). In 2001, the era of robotics comes into existence where the robots by a giant computer from IBM manage to beat humans in simulation of financial trading competition where the New Scientists magazine reported that bots managed to make 7% additional money compared to what the humans made [8]. As such, AI then applied in finance to reduce the financial crimes and fraud by monitoring the users' abnormal changes and irregularities behavioral patterns [9]. AI proved that it is making the markets more efficient thus improves the decision-making in financial sectors. AI systems in trading and investment make fast trading decisions to perform the orders at a greater speed compared to human capabilities and able to produce millions trades per day with no human interventions. This becomes one of the greatest sectors in trading financially. Most large financial institutional investors let the AI systems do the trading. The application of AI in trading and investment combines the historical price with natural language processing that able to accept text from news, feeds from social media, and financial reports. As such, this paper examined the feasibility of using the SVM in financial forecasting.

STATE OF THE ARTS

A. Support Vector Machine (SVM)

In their paper, [2] SVM is applied to calculate the distance between the hyperplanes where N, a linearly dividable training samples , $X = \{x_1, x_2, \ldots, x_N\}$), where x_i is the *ith* of training sample with each sample has d attributes and is one of the two classes $y_i \in \{\pm 1\}$. The line, $\mathbf{w}T\mathbf{x} + b = 0$, represents the decision boundary between the two of the classes, where b denotes the bias, while w denotes a weight vector, and x is denoting the training sample. The hyperplane divides the space into two hyperplanes [2], [10]–[12]. The intention of SVM is to get the values of both w and b to position the hyperplane to be at the largest distance from the closest samples, i.e support vectors, and to develop the two planes, the *H1* and *H2*, where, $H1 \rightarrow \mathbf{w}^T\mathbf{x}_i + b = +1$ for $y_i = +1$ and $H2 \rightarrow \mathbf{w}^T\mathbf{x}_i + b = -1$ for $y_i = -1$, as $\mathbf{w}^T\mathbf{x}_i + b \geq +1$ to be positive class and $\mathbf{w}^T\mathbf{x}_i + b \leq -1$ to be negative class, and the other two equations can be grouped as follows, $y_i(\mathbf{w}^T\mathbf{x}_i + b) - 1 \geq 0\ \forall i = 1, 2, \ldots, N$.

Thus, the distance from *H1* and *H2* to the hyperplane or its decision boundary is represented by d_1 and d_2, correspondingly, where $d_1 = d_2 = \frac{1}{\|w\|}$ and the summation of these distances denote the margin of SVM. In SVM, the margin width requires to be maximized as below:

$$\min \frac{1}{2} \|w\|^2 \qquad (1)$$
$$s.t.\, yi(\boldsymbol{w}^T\boldsymbol{x}_i + b) - 1 \geq 0\ \forall i = 1,2,\ldots,N$$

Equation (1) can be formulated as follows:

$$\min L_P = \frac{\|\boldsymbol{w}\|^2}{2} - \sum_i \alpha_i \left(yi(\boldsymbol{w}^T\boldsymbol{x}_i + b) - 1\right) \qquad (2)$$
$$= \frac{\|\boldsymbol{w}\|^2}{2} - \sum_i \alpha_i y_i \left(yi(\boldsymbol{w}^T\boldsymbol{x}_i + b) + \sum_{i=1}^{N} \alpha_i\right)$$

where $\alpha_i \geq 0$, $i = 1, 2, \ldots, N$ is the Lagrange multipliers. The dual form for SVM can be inscribed as follows:

$$\max L_D = \sum_{i=1}^{N} \alpha_i - \sum_{i=1}^{N} \alpha_i \alpha_j\, y_i y_j\, \boldsymbol{x}_i^T \boldsymbol{x}_j \qquad (3)$$
$$\text{s.t. } \alpha i \geq 0, \sum_{i=1}^{N} \alpha_i\, y_i = 1,2,\ldots,N$$

where L_D is the dual form of L_P.

A new sample $\mathbf{x}_0$ is classified by evaluating $y_0 = \text{sgn}(\mathbf{w}^T\mathbf{x}_0 + b)$ and if y_0 is positive; thus, new sample will be grouped in the positive class; otherwise, it will be grouped in the negative class. Result will be misclassified when the data is non-separable. Hence, the limitations of linear SVM must be satisfied. Also, a nonlinear separable data can be recognized using kernel functions as below:

$$min \frac{1}{2} \|w\|^2 + C \sum_{i=j}^{N} \in_i \qquad (4)$$
$$\text{s.t. } y_i (\mathbf{w}^T\varphi(\mathbf{x}_i) + b) - 1 + \in_i \geq 0\ \forall_i = 1, 2, \ldots, N$$

where $\in_i$ is the distance of the i^{th} training sample and its corresponding margin hyperplane which should be minimized, C is the regularization or penalty parameter which controls the trade-off between the margin size as well as the slack variable penalty, and φ is a nonlinear function which the data can be separated linearly.

There are different kernel functions such as Radial Basis Function (RBF) $K(x_i, x_j) = \exp(-\|x_i - x_j\|^2/2\sigma^2)$, Linear Kernel, $K(x_i, x_j) = \langle x_i x_j \rangle$, and polynomial kernel of degree d, $K(x_i, x_j) = (\langle x_i | x_j \rangle, + c)^d$, are used in [13]. However, this study applies the RBF kernel to obtain the optimal solution.

B. Whale Optimization Algorithm (WOA)

This study proposed a hybrid sentiment-based Support Vector Machine (SVM) optimized by Whale Optimization Algorithm (SVMWOA). In this hybrid algorithm, parameters of kernel function were optimized by the WOA. The techniques and proposed solutions by the SVMWOA are presented in the next subsections.

WOA is a swarm technique introduced by [14], [15] that impersonates the behavior of the humpback whales while hunting for their food of krill herds and small fish which is close to the water surface. The whales behavior is known as the bubble-net feeding method where they make bubbles individually for about 12meters down in a spiral shape everywhere near the prey, hunting them by circling the prey and then swim up to the surface (see Figure 2). The whale optimizer hunting behavior is clarified by three steps, encircling of the prey, bubble-net attacks, and lastly is the prey searching.

WOA, behaviorally as a mathematical model towards getting the optimization method. After position N number of whales (particles) in a solution space, the whales then simulate their movement towards possible preys (possible values for optimum solutions).

Encircling prey mechanism is performed within the solution space, by considering 2D, 3D, or hyper-dimensional (quadratic, cubic, or hyper-cubic) vector positions to simulate the encircling mechanism [16], [17]. The movement of a whale towards the prey is completed in a spiral shape through the phase of bubble-net attack-feeding [14], [15], [17]. The equation below expresses the spiral movement [16]:

$$\vec{X}(t+1) = \vec{D'}.e^{bl}.cos(2\pi l) + \vec{X^*}(t) \qquad (5)$$

Authorized licensed use limited to: International Islamic University Malaysia. Downloaded on August 08,2021 at 12:29:21 UTC from IEEE Xplore. Restrictions apply.

where $\overrightarrow{D'}$, denotes the distance of the associated whale besides the conforming prey (currently is the best solution), b represents the spiral shape, while *l* denotes a random number of [-1, 1]. The step for searching of prey comprises the movements like the ones made based on position of vectors explained for the phase of encircling prey. All phases of the WOA are accomplished according to a terminating criterion where in this case, we call it as number of iterations.

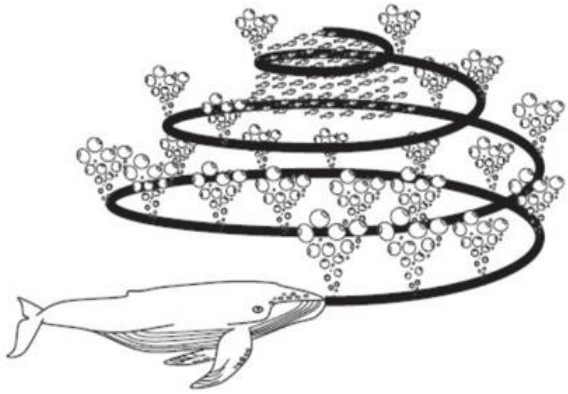

Figure 2. Bubble net feeding behavior of humpback whales in WOA [14], [16]

C. *Social Media Feed*

Modern data mining procedures brings us to the establishment of sentiment analysis, an algorithm developed to analyze the sentiment dataset collected from social media. Social media becomes a necessity to the society, growing and no one can imagine how much we can perceive the utilization of this kind of technology, and to what extend the user would go in deceiving or preying on other users who are susceptible [18]. A major field for the use of sentiment analysis in this study has been cryptocurrency market forecasting which undergoing intense studies at present [19]. A lot of information is communicated over various social media such as Twitter. This becomes a valuable information and proven to provide a significant perception about the communal opinion on a certain subject matter. The correct forecasting about the movement of cryptocurrency market price are influenced by many factors and public opinions are questionably included.

D. *Root Mean Squared Error (RMSE)*

The Root Mean Squared Error is signified by RMSE. RMSE is used to measure performance of the model in this study. The RMSE can be calculated in Equation 6 [20].

$$\text{RMSE} = \sqrt{\left(\frac{\sum (T_i - O_i)^2}{N}\right)} \quad (6)$$

Where N represents the number of all data, O_i represents the predicted value of the model, while T_i is the value, i=1,2, ……,N.

METHODOLOGY

Twitter Sentiment Data Fetching through API

Sentiment Analysis can be achieved by either machine learning or lexicon-based approach. In this thesis, we applied a Lexicon Based Approach and considered this approach as a practical approach to analyzing tweet text without training or using machine learning algorithms. For implementation purposes, Python 3.7 is used. For this study, we collected a sentiment dataset from Twitter's random stream (using the 'getOldTweets3' package for the period from January 2017 through January 2020 for all selected cryptocurrencies under studies. However, not all tweets are included in the sample of the study. We extracted 1000 tweets on 'bitcoin', 'dash', 'ethereum', 'litecoin', 'ripple', 'AUDUSD', 'EURUSD', 'USDCAD', per day from Twitter as shown in Figure 6.0. The approach is to ensure whether the tweet included a cryptocurrency reference or not illustrated on a related works that also refers to the similar situation. By using the Natural language Processing (NLP) pipeline using Python Regular Expression (Regex) library, we cleaned redundant information and stop words, remove punctuations, convert words, handle emojis followed by counting the number of vectorization to convert tweets data into pieces of tokens. This step is very important to prepare the dataset for the next step which is to perform sentiment analysis by TextBlob approach. TextBlob is a Python library that reuses NLTK corpora, to get the sentiment polarity and subjectivity. Figure 3 below shows the installation command to install the TextBlob and to download the NLTK corpora from the Python library.

```
pip install -U textblob
python -m textblob.download_corpora
```

Figure 3. Installation command to install TextBlob

The sentiment analysis was incorporated using Python's TextBlob library where sentiment computation and polarity classification is performed. This approach categorizes the polarity of the dataset in positive, neutral, and negative categories into '1','0', and '-1'. Lastly, results from the dataset retrieved were then visualized. Figure 3 shows the flowchart of sentiment classification with TextBlob.

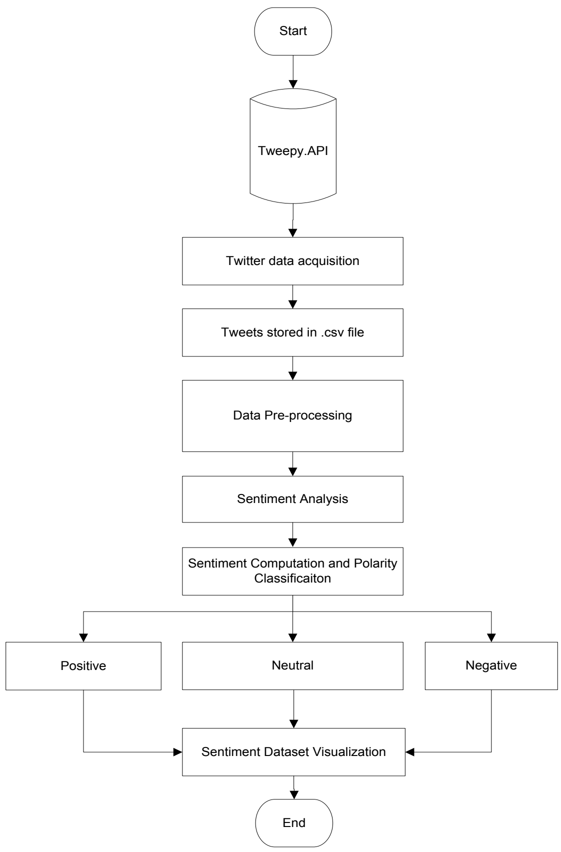

Figure 4. Flowchart of Sentiment Classification with TextBlob

The sentiment analysis process is further explained through step 1- 5, from the retrieval of the Tweets towards the polarized data

Authorized licensed use limited to: International Islamic University Malaysia. Downloaded on August 08,2021 at 12:29:21 UTC from IEEE Xplore. Restrictions apply.

of the sentiment. Figures 4-10 and the Algorithm 1 pseudo-code of sentiment analysis using the TextBlob serve as a guideline of the proposed method of Twitter sentiment analysis.

Step 1:

Tweets retrieval using Python library called Tweepy.API and it takes as input various parameters (tweet.id, permalink, date, hashtags and location).

Step 2:

Collected Tweets are saved in .csv file.

Step 3:

Pre-processing of the tweets is performed using Natural Language Processing (Python's Regular Expression library and Vectorization count).

Step 4:

TextBlob Sentiment Analysis is performed by using NLTK from Python Library to classify tweets either positive or negative. Figure 5 shows the TextBlob Approach step.

```
#        print(daystweets)
        totalPolarity=0.0
        for text in daystweets['text']:
            text = preprocess_tweet(text)
            s=TextBlob(text)
#            print(s.sentiment)
            totalPolarity = totalPolarity+ float(s.sentiment.polarity)
```

Figure 5. TextBlob Approach Step

Step 5:

Sentiment data visualization as depicted in Figure 8 shows the polarized data, that act as a final stage of TextBlob sentiment analysis where it gives the sentiment score and polarity classification as a result.

Algorithm 1: Pseudocode of Sentiment Analysis with TextBlob Approach in Twitter Data

Input: Fetching Tweets from Twitter
Output: Polarised data
Method:

```
function TwitterAPI_setup()
    set TweetCriteria
    set argument 'bitcoin'
end function
write tweets into .csv file
Create a dataframe tweet using pandas dataframe
Loading Python Regular Expression (RE) library
Read the dataset
#Preprocess tweet
    #Remove punctuation
    #Replace URLs
    #Replace #hashtag
    #Remove RT
    #Replace emojis
    #Replace multiple spaces with single space
return processed_tweet
#Sentiment computation and Polarity classification
totalPolarity=0.0
    for text in daystweets['text']:
        text = preprocess_tweet(text)
        s=TextBlob(text)
#print(s.sentiment)
Calculate sentiment polarity of summarized text
        totalPolarity = totalPolarity+
float(s.sentiment.polarity)
return polarized data
```

E. Transposing the Twitter Data

To perform further analysis, the Twitter data needed to transpose to match with the cryptocurrency price index dataset. This means that the tweets need to be collapsed into a total of 1126 rows, one for each day for the period from January 2017 through January 2020. To transpose the Twitter data, we grouped the texts of the tweets by day to capture broader sentiment expressed by the investors and traders. Once the tweets texts were grouped daily, we performed the sentiment analysis and fetched the average daily score of all the 'compound' scores. As discussed in the previous section of this thesis, we are using the TextBlob package for sentiment analysis, by using TextBlob and NLTK we obtain three elementary scores. They are positive, negative, and neutral where each score compound ranging between -1 and 1 which indicates the overall intensity of the sentiment expressed in the text based. Further, by taking the average of all the compound scores of each tweet in a day, we obtained the daily average sentiment score to align our Twitter dataset to 1126 rows which represented the overall score for each day. This will further be used to assign polarity and subjectivity of the data. TextBlob gives us the sentiment score that is how positive/negative a word is while a subjectivity score gives us how subjective or opinionated a word is. We carried out the accuracy test and the result show a 77% accuracy rate, and a sentiment score (how positive/negative a word is), and a subjectivity score (how subjective, or opinionated a word is).

F. Integration of Sentiment and Cryptocurrency Data

Knowing that there could be more than one tweet on each day about each market, so the data extracted from sentiment prediction in previous is accumulated daily. It means, if there are more positive tweets than negative, we can conclude that the crypto sentiment is positive on that day, and investors can buy the share. The crypto market data and sentiment data are integrated at this stage through matching dates. At the end of the preprocessing phase, there will be three variables in the dataset file i.e. date, price, and sentiment data respectively. The integrated data set is then further investigated with the proposed Hy-Senti SVMWOA).

G. SVM Optimized by WOA

In this study, the WOA is applied to optimize the SVM parameters, namely C and ε by searching for the optimal values that minimize the root mean square error (RMSE) of an SVM in the training phase. The optimization performance of SVMWOA is tested by the objective function and two algorithms for comparison with SVMWOA. The WOA flowchart is briefly described in Figure 11.

DATA DESCRIPTION

This paper applies four (4) years daily prices from 2017 through 2020 for all data models and is prepared from daily open, high, low, close (OHLC) of a daily trading for all five types of cryptocurrency.

Table 1: Description of Variable

Variable	Description
Open Price	First price of a daily trading
Close Price	Last transaction of a daily trading
High Price	Highest price of a daily trading
Low Price	Lowest price of a daily trading

Authorized licensed use limited to: International Islamic University Malaysia. Downloaded on August 08,2021 at 12:29:21 UTC from IEEE Xplore. Restrictions apply.

The collected datasets from the cryptocurrency market are summarized in Table 2.

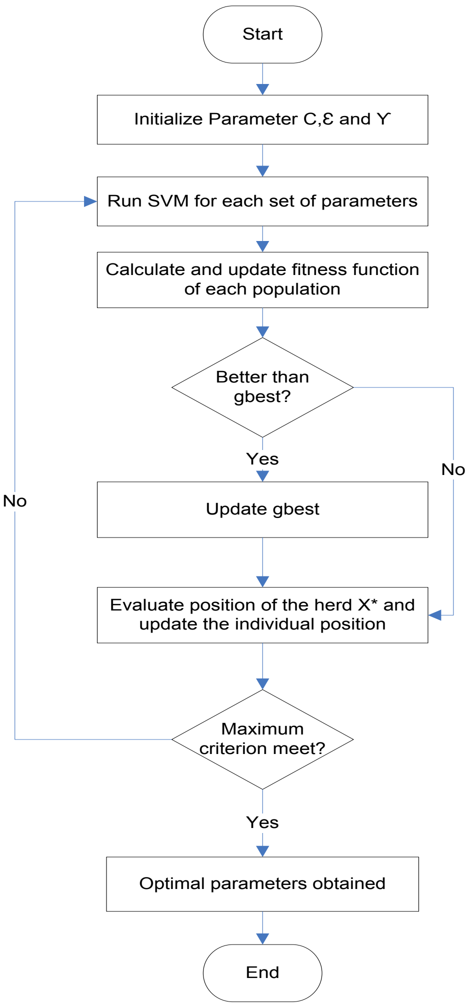

Figure 6. Flowchart of Support Vector machine optimized by whale optimization algorithm (WOA)

Table 2. Cryptocurrency market price data

Domain	Training Data: 1st January 2017 through 20th June 2019 Testing Data: 21st June 2019 – 31st January 2020 **Sentiment data were extracted for the same range of dates*	
	#Observations	Sources
Bitcoin	1126	https://coinmarketcap.com/
Dash		
Ethereum		
Litecoin		
Ripple		

To compare the results of Hy-Senti-SVMWOA model and to benchmark models, an SVM method has been applied which is the feature selection and parameter optimization. An SVM method with the Genetic Algorithm (named SVMGA) and an SVM with the Harmony Search (named SVMHS). More information on the SVMGA and SVMHS can be referred to [21]–[24].

Experimental Result

Performance Measures and Comparison

A. Comparative performance of all models for cryptocurrency data

Figure 7(a) through Figure 7(e) shows a comparative performance of all models for the predicted and the actual value of Bitcoin, Dash, Ethereum, Litecoin and Ripple against time for all the Optimized SVM models used in this study respectively. It shows that the Optimized SVM models have forecasted higher than the actual values indicating that the models have overpredict. However, approximately between September 2019 through January 2020, SVMGA and SVMHS capture some patterns of the data, with SVMWOA produced the most accurate prediction. Among the three models, the model which deviated by a large margin from the actual value is SVMHS except for Ripple dataset where SVMGA produces a poor prediction result. Furthermore, throughout the graph, it is apparent that SVMHS failed to capture the pattern of the data.

Figure 7(a). Actual and predicted values for individual SVM models using Bitcoin data

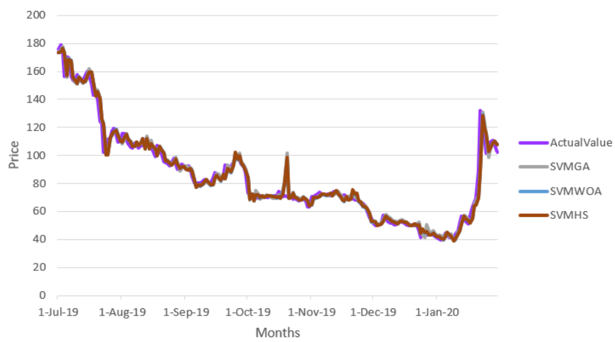

Figure 7(b). Actual and predicted values for individual SVM models using Dash data

Authorized licensed use limited to: International Islamic University Malaysia. Downloaded on August 08,2021 at 12:29:21 UTC from IEEE Xplore. Restrictions apply.

Figure 7(c). Actual and predicted values for individual SVM models using Ethereum data

Figure 7(d). Actual and predicted values for individual SVM models using Litecoin data

Figure 7(e). Actual and predicted values for individual SVM models using Ripple data

Moreover, Table 3 shows the results of three methods above in five datasets that signify the Hy-Senti-SVMWOA method outperforms the other two in for all datasets. This is followed by SVMGA. SVMHS is the most underperforms method among the three selected methods in this study. In another word, the Hy-Senti-SVMWOA is considered a stable model with its capability to make a prediction which is close to the actual value.

Table 3. Performance measures for all dataset

Dataset	Classifier	MAE	RMSE
Bitcoin	SVMGA	2.440089e+02	3.686461e+02
	SVMHS	2.509928e+02	3.800013e+02
	Hy-Senti-SVMWOA	**2.425547e+02**	**3.673133e+02**
Dash	SVMGA	3.40474147	5.71052476
	SVMHS	3.17122197	5.55401446
	Hy-Senti-SVMWOA	**3.19372009**	**5.52816917**
Ethereum	SVMGA	5.69603047	8.62071142
	SVMHS	5.75933447	8.70238983
	Hy-Senti-SVMWOA	**5.60349903**	**8.51713825**
Litecoin	SVMGA	2.35203835	3.71321937
	SVMHS	2.36493258	3.70164420
	Hy-Senti-SVMWOA	**2.34609659**	**3.67800421**
Ripple	SVMGA	0.0087290314	0.0136820163
	SVMHS	0.0101003339	0.0154136034
	Hy-Senti-SVMWOA	**0.0083897818**	**0.0123276216**

The result of MAE shows that there is some variation in the magnitude of the errors. Results describe that the average difference between the forecast and the actual values for Bitcoin, Dash, Ethereum, Litecoin, and Ripple are 2.425547e+02, 3.19372009, 5.60349903, 2.34609659, and 0.0083897818 respectively. Ripple shows to be producing the smallest MAE among other data sets. Result shows that Overall, the RMSE-MAE difference for all methods is not large enough to indicate the presence of very large errors. This confirms that the proposed Hybrid Sentiment based-SVMWOA model meets the objective of the study where the sentiment does give a very significant impact on the time series forecasting for the digital currency.

CONCLUSION

In this concluding section, the study aims at getting the accurate prediction of a digital currency, or also known as cryptocurrency which is also a significant factor to help the investors in their decision making in order to gain profit in every single trade made. As the market is dynamic and drives by many factors including inflation rate, economics, and natural calamities, thus creates volatility in the cryptocurrency price. Many issues have been touched during the study in particular data acquisition, sentiment data extraction, preprocessing, adapting the model to the classification of the future price, including the optimization of the algorithms. Results from analysis performed in the next section, the computational intelligence technique by developing a Hybrid Sentiment Based Support Vector Machine Optimized by Whale Optimization Algorithm (Hy-Senti-SVMWOA), the result shows similarity in its value which is comparable with the other datasets for the particular period of study. Finally, promising results were obtained in one out of the three comparative studies, comparable to the literature in the field. It shows that the sentiment including SVM parameters optimization plays a vital part in producing successful and accurate forecasting results. Future work will focus on the enhancement of the model accurateness including the adjustment of other specific parameter selections and be applied in the other field of studies.

REFERENCES

[1] N. A. Hitam, A. R. Ismail, and F. Saeed, "An Optimized Support Vector Machine (SVM) based on Particle Swarm Optimization (PSO) for Cryptocurrency Forecasting." 2019.

[2] A. Tharwat, T. Gabel, and A. E. Hassanien, "Parameter optimization of support vector machine using dragonfly algorithm," *Adv. Intell. Syst. Comput.*, vol. 639, no. January, pp. 309–319, 2018.

[3] G. Xie, "The Optimization of Share Price Prediction Model Based on Support Vector Machine," *Control. Autom. Syst. Eng. (CASE)*, pp. 1–4, 2011.

[4] M. K. N. Devi, D. V. M. Bhaskaran, and G.Prem Kumar, "Cuckoo Optimized SVM for Stock Market Prediction," 2015.

[5] C.-L. Yu, Y.-Z. Lu, and J. Chu, "An ``Load Forecasting - dispatching'' Integration System for Multiple Boilers

Authorized licensed use limited to: International Islamic University Malaysia. Downloaded on August 08,2021 at 12:29:21 UTC from IEEE Xplore. Restrictions apply.

in Thermal Power Plants," *2011 IEEE Power Eng. Autom. Conf.*, no. 60721062, pp. 5–10, 2011.
[6] Z. Yang, K. Shi, A. Wu, M. Qiu, and Y. Hu, "A Hybird Method Based on Particle Swarm Optimization and Moth-Flame Optimization," *Proc. - 2019 11th Int. Conf. Intell. Human-Machine Syst. Cybern. IHMSC 2019*, vol. 2, pp. 207–210, 2019.
[7] I. Madan, S. Saluja, and A. Zhao, "Automated Bitcoin Trading via Machine Learning Algorithms," *URL http//cs229. stanford. edu/proj2014/Isaac\% 20Madan*, vol. 20, pp. 1–5, 2015.
[8] Http://news.bbc.co.uk/, "BBC News | BUSINESS | Robots beat humans in trading battle." 2001.
[9] D. Schutzer, "CTO Corner: Artificial Intelligence Use in Financial Services - Financial Services Roundtable," *Fsroundtable.Ccom*, vol. April. pp. 1–6, 2015.
[10] Y. Kara, M. Acar, and Ö. Kaan, "Expert Systems with Applications Predicting direction of stock price index movement using artificial neural networks and support vector machines : The sample of the Istanbul Stock Exchange," *Expert Syst. Appl.*, vol. 38, no. 5, pp. 5311–5319, 2011.
[11] H. Nor Azizah and I. Amelia Ritahani, "An Optimized Support Vector Machine (SVM) based on Particle Swarm Optimization (PSO) for Cryptocurrency Forecasting," in *The 16th International Learning and Technology Conference*, 2018, vol. 91, pp. 399–404.
[12] H. Nor Azizah, I. Amelia Ritahani, and F. Saeed, "An Optimized Support Vector Machine (SVM) based on Particle Swarm Optimization (PSO) for Cryptocurrency Forecasting," *16th Int. Learn. Technol. Conf.*, vol. 91, pp. 399–404, 2019.
[13] I. Mierswa, "Evolutionary learning with kernels," no. February, p. 1553, 2006.
[14] S. Mirjalili and A. Lewis, "The Whale Optimization Algorithm," *Adv. Eng. Softw.*, vol. 95, pp. 51–67, 2016.
[15] D. Mokeddem and S. Mirjalili, "Improved Whale Optimization Algorithm applied to design PID plus second-order derivative controller for automatic voltage regulator system," *J. Chinese Inst. Eng. Trans. Chinese Inst. Eng. A*, vol. 00, no. 00, pp. 1–12, 2020.
[16] U. Kose, "A Hybrid SVM-WOA Approach for Intelligent Fault Diagnosis Applications," *Proc. - 2019 Innov. Intell. Syst. Appl. Conf. ASYU 2019*, pp. 1–6, 2019.
[17] R. Masadeh, A. Alzaqebah, and A. Sharieh, "Whale optimization algorithm for solving the maximum flow problem," *J. Theor. Appl. Inf. Technol.*, vol. 96, no. 8, pp. 2208–2220, 2018.
[18] S. I. Syed Zulkarnain and H. Nor Azizah, "Social Media Use or Abuse : A Review," *J. Hum. Dev. Commun.*, no. August 2015, 2014.
[19] J. Kordonis, S. Symeonidis, and A. Arampatzis, "Stock price forecasting via sentiment analysis on Twitter," *ACM Int. Conf. Proceeding Ser.*, no. November, 2016.
[20] R. Chanklan, N. Kaoungku, K. Suksut, K. Kerdprasop, and N. Kerdprasop, "Runoff prediction with a combined artificial neural network and support vector regression," *Int. J. Mach. Learn. Comput.*, vol. 8, no. 1, pp. 39–43, 2018.
[21] C. N. Babu, "Prediction Of Time Series Data Using GA-BPNN based Hybrid ANN Model," 2017.
[22] O. M. E. Ebadati and M. T. Mortazavi, "An efficient hybrid machine learning method for time series stock market forecasting," *Neural Netw. World*, vol. 28, no. 1, pp. 41–55, 2018.
[23] B. J. Pillay and A. E. Ezugwu, "Metaheuristics optimized feedforward neural networks for efficient stock price prediction," 2019.
[24] X. S. Yang, *Nature-Inspired Optimization Algorithms*. 2014.

Authorized licensed use limited to: International Islamic University Malaysia. Downloaded on August 08,2021 at 12:29:21 UTC from IEEE Xplore. Restrictions apply.
View publication stats