

Optimizing Extreme Learning Machine using GWO Algorithm for Sentiment Analysis

Mustafa Abdul Salam
Artificial Intelligence Dept,
Faculty of Computers and Artificial Intelligence
Benha University, Benha, Egypt

Mahmoud Ali
Scientific Computing Dept.,
Faculty of Computers and Artificial Intelligence
Benha University, Benha, Egypt

ABSTRACT

Sentiment analysis on social media is one of the most popular text mining application and many researchers have devoted more efforts in this interesting field. Sentiment analysis is a method for analyzing data and extracting the feeling it represents. Twitter is considered one of the most common social media forums used by people on various occasions to express their opinions and express feelings. Twitter's sentiment analysis has grown gradually over the past few decades. Due to the format of small tweet, a new dimension is created for problems such as slang usage, abbreviations, etc. This paper proposes a hybrid approach that optimizes extreme learning machine (ELM) classifier with one of the most recent swarm intelligence algorithms which is grey wolf optimization algorithms (GWO). GWO is used to self-adaptation of hidden neurons weights rather than manually selection. Also avoid the overfitting problem and make the model more generalized and robust. Results represented in this paper showed that the proposed hybrid model GWO-ELM overcame the problems found in classical ELM model and achieved best accuracy against all compared models.

Keywords

Sentiment Analysis, Machine Learning, Extreme Learning Machine, Grew Wolf Optimization, Text Mining.

1. INTRODUCTION

Sentiment analysis or opinion mining is the process of extracting and summarizing conclusions from opinions with monstrous quantity of data which the average human reader is not able to process [1]. Nowadays, Individuals from different locations within the worldwide be part of to each other through social media web sites such Facebook, Google+, Twitter, Instagram, etc.

As Twitter has huge range of users (more than 300 million users) and enormous data, it has usually been used as an informative resource by various organizations to research public opinions and gather critical feedback [2]. Due to this limited tweet length, people write in a very concise manner by using slangs. It is beneficial for the marketers as they can examine and analyze the opinion of the public towards their brand and existing/newly released products which would help them to evaluate their performance and improve it [3].

Users can communicate their views in the form of tweets on Twitter, using only 280 characters. This causes individuals to

Use slang, abbreviations ,emoticons, brief forms, etc. to compact their statements. Using sarcasm nd polysemy, individuals express their views along with this.

Machine learning (ML) methods such as artificial neural networks (ANNs), support vector machines (SVMs), extreme

learning machine (ELM), are considered the most commonly used ML models in classification, regression. Also were used in natural language processing (NLP), and text mining such as social media sentiment analysis. But these methods may suffer from local minima and overfitting problems due to using local optimization training algorithms such as gradient descent algorithm in ANN [4].

Swarm Intelligence algorithms such as particle swarm optimization (PSO), follower pollination algorithm (FPA), ant colony optimization (ACO), and artificial bee colony (ABC), can solve the problems or drawbacks of machine learning models such as ANN, SVM, and ELM methods [5,6].

Using swarm intelligence or meta-heuristic algorithms in optimizing and training classical machine learning models can enhance the accuracy and generalization ability of these methods [7-10].

In this paper, a proposed hybrid model called GWO-ELM, combining ELM with a recent swarm intelligence algorithm called GWO is proposed and applied for twitter's sentiment analysis. GWO selects the optimal weights and biases to enhance the performance of ELM model. The proposed model enhances the classification accuracy and generalization performance of the traditional ELM method.

The rest of this paper is organized as follows: Section 2 presents the related work. Section 3 presents the proposed model. The results and discussions are introduced in section 4. The main conclusion of this paper is presented in Section 5.

2. RELATED WORK

The information that people listen to and share in social media is largely influenced by the social circles and relationships they form online [11]. Researchers, politicians, people as well as individuals around the world would be interested in the feat of accurately monitoring the spread of false messages and particularly news content. This can be done by using "personal sensor devices" that are accurate and important.[12]. In countries that have accepted and welcomed technology as part of their electoral process and have therefore implemented e-voting, this need is more significant. [13].

Swarm Intelligence (SI) is another strong CI type used to solve the problems of optimization. SI algorithms model and replicate natural swarms or communities or structures like fish schools, swarms of birds, bacterial growth, colonies of insects, and herds of animals.[14]. In addition to the interactions, most SI algorithms concentrate on the behavior of swarm members and their lifestyle, and relationships between swarm members to find food sources. SI algorithms include many ant colony optimization (ACO) algorithms [15], GWO is one of the recently proposed swarm intelligence-based algorithms, which is developed by Mirjalili et al. [16]. In order to solve problems

of continuous optimization as well as combinatorial optimization (such as TSP, knapsack problem, vehicle routing problem), a number of such nature-inspired SI algorithms are developed. Some of these are the optimization of particle swarm (PSO) [17]. A hybrid dragonfly algorithm with an extreme learning machine for forecasting is introduced, in this proposed technique, DA optimizer improved the performance of ELM in forecasting or prediction [18]. Meta-heuristics for prime factorization is introduced [19]. Some swarm intelligence methods have been applied to solve this problem using specific objective functions [20-22].

Ankita Gupta et al. [23] introduced, KNN, and SVM hybrid model. This model enhances the performance of the classical classifiers. The introduced model has three classes positive, negative in addition to neutral class. The simulation results show that the hybrid model improves the performance of class prediction.

Barbosa et al. [24] implemented a model for classifying tweets. Tweets classified as objective (positive) or subjective (negative).

Bharat Naiknaware et al. [25] The SVM classifier gave the best accuracy performance for seven datasets.

Bifet and Frank et al. [26] implemented multinomial NB, SGD, and the Hoeffding tree. SGD was the best one.

Christianini and Taylor. [27] implemented SVM which was giving better accuracy than compared methods.

Daoud. et al. [28] proposed a method consists of four parts, AdaBoost, BNN, SVM and Signal-to-Noise.

Go and Huang et al. [29] introduced twitter's sentiment analysis. NB, MaxEnt and SVM were used. SVM gave better accuracy than other models.

Akhtar, M.S, et al.[30] Worked on a feature selection technique that relies on an aspect-based study of the feelings. The main focus of this research work was on the creation of a cascade collection of features and classifiers. PSO (Particle swarm optimization) was the premeditated technique. In two steps the aspect-based SA (Sentiment Analysis) was operated. The two fundamental steps on which SA was carried out were extraction and classification of aspects. Additionally, the features were used to identify the properties of different classifiers and their domain. Essentially, the maximum entropy, conditional random field, and help vector machine were used as three types of classifiers. The results that were produced recently were more successful.

Kang et al. [31] proposed a senti-lexicon of restaurant reviews for the sentiment analysis. The author proposed improved NB, INB to improved the accuracy when applied with bigram.

Kopel and Schler et al. [32] states that negative and positive messages alone no meaning. Corpus has been used for testing both polarity of a message.

Li and wu. et al. [33] presented polarity identifying as sentiment analysis of text mining. The authors made a model which used unsupervised text mining method, with SVM and K-means clustering. K-means and SVM gave the similar results.

Anjaria et al. [34] implemented supervised ML models and NN to classify tweets presidential elections, SVM was the best classifier. The introduced model used to forecast the outcome of election results by studying the influence value of the person. To reduce the dimensions PCA combined with SVM.

Martineau et al. [35] introduced a model called D-TFIDF. SVM gave better performance with D-TFIDF to movies reviews datasets.

Mohammad et al. [36] presented two SVM models, level task to define sentiment of a word. Also, message level specifies the phrase sentiment like, tweets and short messages. The introduced model has 88.93 F-score in term-level stage and 69.02 F-score in message-level stage.

3. PROPOSED MODEL

The phases of proposed methodology are described in figure (1). Starting with data gathering, preprocessing, feature extraction, hybrid model training and optimizing, then classifying and computing the accuracy.

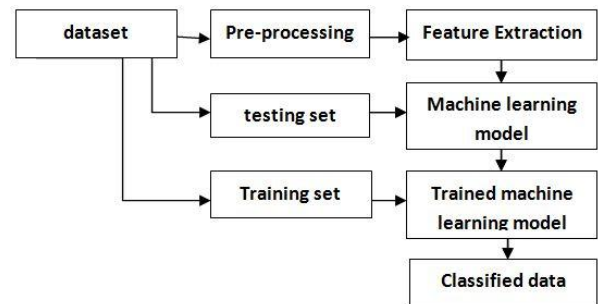


Fig. 1 – Phases of sentiment analysis methodology

3.1 Tweet Collection

Gathering relevant tweets about the specified subject or topic. The tweets are gathered by Twitter's API [37], [38], for the needed time period. Then convert the text to suitable form like JSON [38], [39].

The dataset must be gathered carefully to test the model with robust datasets. Dataset is splatted into training part and testing part.

3.2 Tweets Preprocessing

Data preprocessing is a very important phase as it specifies the effectiveness of next steps. It includes, as required, syntactic tweet correction. The steps concerned should strive to make the information readable by more machine to decrease ambiguity in the extraction of features. Below are a few steps for tweet preprocessing.

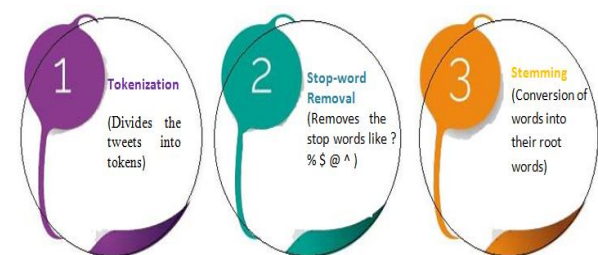


Fig. 2- Pre-processing Steps

3.3 Feature Extraction

The gathered dataset is used to select and extract features that will be used in training of the classifiers.

There are currently various methodologies for extracting characteristics. Term frequency-Frequency Inverse Document is an effective method. TF-IDF is a numerical statistic reflecting a word's significance for the entire document (here, tweet). Selection of effective words from tweets is considered the feature selection or extraction.

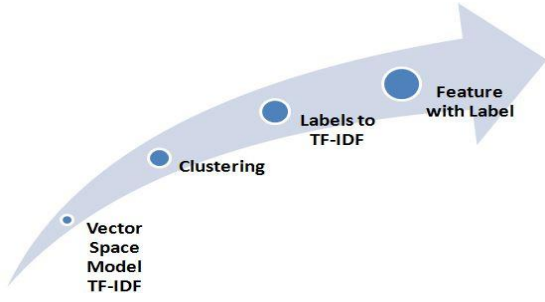


Fig. 3 - Process for extraction of features

3.4 Extreme Learning Machine (ELM)

ELM model has been introduced for single layer NN. In ELM model, the links between the input and hidden layers are selected randomly and still fixed with no change [40]. The output weights are then adapted through minimizing the error rate across a linear system [41].

When we are training ANN with N hidden neurons and transfer function $f(x)$ to learn M distinct samples (x_i, t_i) , where: $x_i = [x_{i1}, x_{i2}, \dots, x_{ik}]^T \in R^k$, $t_i = [t_{i1}, t_{i2}, \dots, t_{id}]^T \in R^d$. this convert nonlinear to linear system as follows:

$$H\beta = T \quad (1)$$

where H is the matrix of hidden layer denoted by:

$$H = \begin{bmatrix} f(w_1 \cdot x_1 + b_1) & \dots & f(w_n \cdot x_1 + b_n) \\ \vdots & \ddots & \vdots \\ f(w_1 \cdot x_m + b_1) & \dots & f(w_n \cdot x_m + b_n) \end{bmatrix} \quad (2)$$

where $w_j = [w_{j1}, w_{j2}, \dots, w_{jk}]^T, (j = 1, 2, \dots, N)$ is the vector of weight linking jth hidden node and input node, and b_j the hidden bias of jth node; $w_j \cdot x_i (i = 1, \dots, M)$ denotes the inner product of w_j and x_i ; $\beta = [\beta_1, \beta_2, \dots, \beta_N]^T$, the weights output matrix $\beta_j = [\beta_{j1}, \beta_{j2}, \dots, \beta_{jd}]^T$,

$$t_i = \sum_{j=1}^n b_j f(w_j \cdot x_i + b_j), \quad (3)$$

The Woutput are computed using the LS method introduced in equation (4) as follows:

$$\hat{\beta} = H^*T \quad (4)$$

where H^* is inverse matrix of H. ELM with MP inverse approach leads to obtain better performance with enhanced training rate [42].

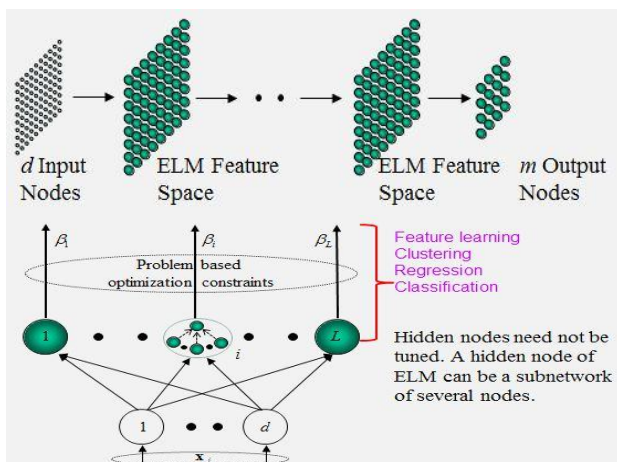


Fig. 4- Hidden neurons in hierarchical ELM

3.5 Grey Wolf Optimization (GWO)

GWO is a meta-heuristic algorithm based primarily on grey wolves. GWO was introduced by Mirjalili et al. in 2014, which is considered promising global optimization algorithm [16]. This algorithm mimics the leadership hierarchy of gray wolves and their hunting methods. Four distinct gray wolves are employed to simulate their hierarchical management. It's alpha, beta, delta, omega.

In addition, there are three main steps for execution involving hunting and prey searching, followed by encircling and attacking prey.

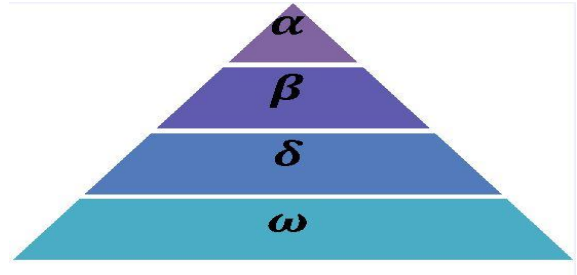


Fig.5- Gray wolf hierarchy (dominance falls from top to down).

Encircling prey: during the chase, grey wolves surround the prey. The following equations were proposed for mathematically modeling encircling behavior [16].

$$D = |C * x_p(t)| \quad (5)$$

$$x(t + 1) = x_p(t) - A * D \quad (6)$$

Where t represents the current iteration, A and C are coefficient vectors, XP is the prey's position vector, and X indicates the gray wolf's position vector. The vectors A & C are computed as follows:

$$A = 2\alpha * r_1 - a \quad (7)$$

$$C = 2 * r_2 \quad (8)$$

Where components of a are reduced linearly from 2 to 0. Random vectors in [0, 1] are random vectors during iterations and r_1, r_2 .

Hunting: to imitate the hunting activity of gray wolves mathematically, it is believed that the alpha and beta (best candidate solution) have a better knowledge of the possible location of prey. the first three best solutions are recorded, and obliging the other search agents to change their locations coping with the location of the best agents in search space. In this respect, the following formulas are proposed.

$$D_\alpha = |C_1 * X_\alpha - X|, D_\beta = |C_2 * X_\beta - X|, D_\delta = |C_3 * X_\delta - X| \quad (9)$$

$$X_1 = X_\alpha - A_1 * (D_\alpha), X_2 = X_\beta - A_2 * (D_\beta), X_3 = X_\delta - A_3 * (D_\delta) \quad (10)$$

$$X(t + 1) = \frac{X_1 + X_2 + X_3}{3} \quad (11)$$

Fscore accuracy will be calculated using equation (12) as follows:

$$Fscore = 2 * [(Pre * Rec) / (Pre + Rec)] \quad (12)$$

Where Pre is the precision value and Rec is the recall value which

computed as follows in equation (13) and equation (14):

$$Pre = TP / (TP + FP) \quad (13)$$

$$Rec = TP / (FN + TP) \quad (14)$$



Fig. 6- Grey wolf hunting behavior

3.6 Hybrid GWO-ELM Sentiment classifier

The proposed method is designed with gray wolf optimization (GWO) and SVM called as the GWO- SVM. The proposed GWO-ELM model consists of two main phases.

In the first step, GWO algorithm is employed to optimize ELM to predict the tweet polarity. GWO algorithm chooses the hidden weights and biases automatically to avoid over-fitting and local minima problems and improve classification accuracy and then ELM classification.

The steps of the GWO-ELM algorithm are introduced in the following Algorithm1.

The fitness function used in GWO to evaluate the best parameters of SVM is shown in equation (12).

ELM has low learning time while giving regression and also classification performance with a huge number of hidden neurons are used. There is no confirmation of finding global solution using set of weights and bases on the hidden layer. In proposed GWO-ELM a small number of neurons in the hidden layer is chosen to accelerate the learning speed of the ELM while guarantee the generalization and optimality. GWO algorithm is used to choose the optimal weights and biases of the hidden layer which maximizes the whole ELM accuracy.

Algorithm1: The Hybrid GWO_ELM Model

Input : n : SearchAgents_no
 max_iter : Max_iteration Number
 lb : Lower bound of data
 ub : Upper bound of data
 dim : Number of dimensions
 trn The training dataset.
 Output : fbest optimal weights and biases
 bestpos final best X_α
 fscore classification accuracy

- 1) Initialization of :
 - Population size (n) and the grey wolf population X_i ($i = 1, 2, \dots, n$).
 - Initial weights and biases.
 - Maximum number of iterations.
 - Generation limits lb & ub and number of dimensions (dim).
 - The best search agent (X_α)
 - The second best search agent (X_β)
 - The third best search agent (X_δ)
 - Function (f) that apply to each search agent.
 - 2) Generate the initial population randomly.
 - 3) while (t < Max_iteration)
 - for each search agent do
 - Construct a neural model given hidden layer weights and biases of grey wolf agent X_i
 - Calculate the fitness of each search agent by equation (12)
 - Calculate the output layer weights using the MP generalized inverse matrix given the training data
 - Assign ELM output parameters
 - Run ELM
 - Compute classification accuracy
 - Update fitness of ($X_\alpha, X_\beta, X_\delta$).
 - end for
 - Update w
 - for each search agent do
 - for each dimension do
 - Update w and b.
 - Update $X_\alpha, X_\beta,$ and X_δ
 - Update the position of the current search agent
 - end for
 - end for
 - t=t+1
 - end while
 - 4) Pass X_α to f
 - 5) Assign final w , b to ELM.
 - 6) Perform ELM.
 - 7) Return classification accuracy.
-

4. RESULTS AND DISCUSSION

4.1 Dataset description:

Twitter dataset used in this paper. Twitter API used to collect the dataset. In collected data files, we use.csv format as data consists of many fields. Dataset splatted into 70 % training and 30 % testing. The data set includes 6000 tweets distributed over positive, negative sentiment groups, and the initial test set consists of 4082 negative and 1918 positive tweets.

4.2 Simulation Results:

Comparison of proposed GWO-ELM model with classical ELM, and SVM is introduced in table 4. Also Proposed model compared with hybrid SVM_ACO, and SVM_PSO models. From table we notice that proposed model achieved the best accuracy, precision, recall, and Fscore. ELM is little less than SVM with advantage of fast training compared to SVM. Also, GWO-ELM is better than both SVM_ACO, and SVM_PSO with fast training also and less training time which is considered one of the advantages of ELM model over SVM models.

Table 1. Performance evaluation of proposed model

Function	GWO_ ELM	ELM	SVM	SVM_ ACO	SVM_ PSO
Accuracy	82.93	75.93	73.23	76.71	80.54
Precision	76.82	68.59	70.32	70.45	73.98
Recall	69.44	55.2	47.07	47.78	50.17
F Score	72.94	61.17	56.39	56.94	59.79

Figures from Fig.7 to Fig.10 represent the performance evaluation of proposed GWO-ELM model. From figures one can notice that proposed model achieves the best accuracy, precision, recall, and F-score and outperforms all compared models.

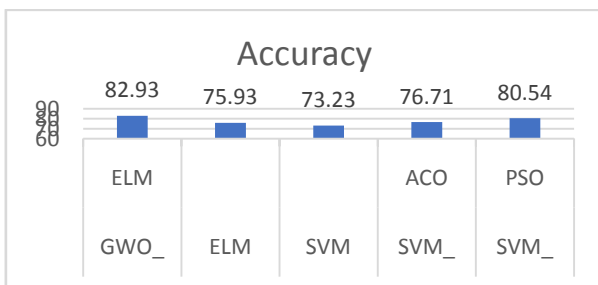


Fig. 7- Accuracy of the proposed and compared models

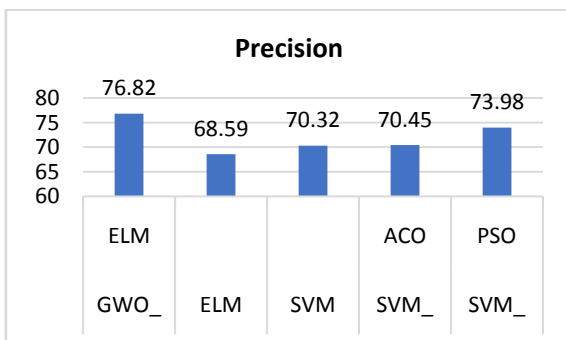


Fig. 8- Precision of the proposed and compared models

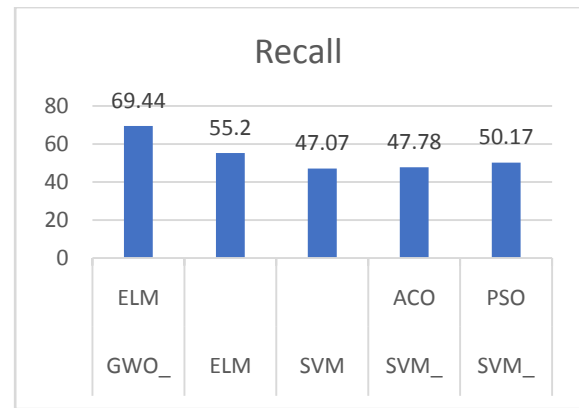


Fig. 9- Recall of the proposed and compared models

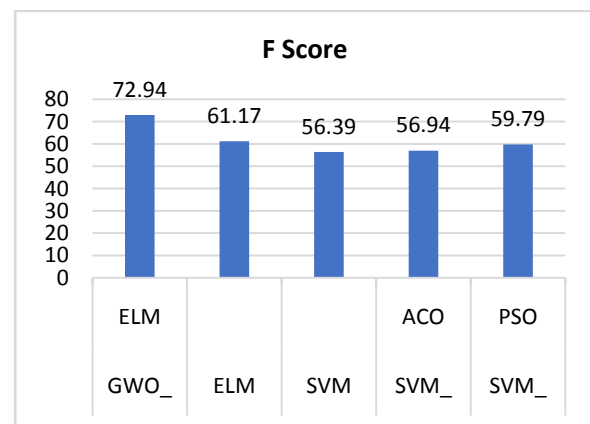


Fig. 10- F-Score of the proposed and compared models

5. CONCLUSION

Analysis of Twitter sentiment falls under the category of mining of text and opinion. It focuses on evaluating the tweet feelings and feeding the information into a machine learning model to train it and then checking its precision so that we can use this model according to the outcomes for future use. It includes measures such as data collection, pre-processing of text, classification of sentiment, instruction and model testing. A hybrid GWO-ELM model is proposed to overcome the drawbacks of traditional ELM model. Results presented in this paper indicate that the performance of GWO-ELM model outperforms ELM and SVM with generalization ability. Also, GWO-ELM is better than both SVM_ACO, and SVM_PSO with fast training also and less training time which is considered one of the advantages of ELM model over SVM models. The accuracy of the proposed model for topics other than the one in consideration is not yet tested. Analysis of sentiment, therefore, has a very bright future development scope.

6. REFERENCES

- [1] A. Cobo, D. Parra, and J. Navon. Identifying relevant messages in a twitter-based citizen channel for natural disaster situations. In Proceedings of the 24th International Conference on World Wide Web Companion, pages 1189–1194, 2015.
- [2] E. Boiy, P. Hens, K. Deschacht and M. Moens, 'Automatic sentiment analysis in on-line text', 11th International Conference on Electronic Publishing, vol. 349360, 2007.
- [3] D. Zimbra, M. Ghiassi and S. Lee, "Brand-Related Twitter Sentiment Analysis using Feature Engineering and the Dynamic Architecture for Artificial Neural Networks",

- IEEE 1530-1605, 2016.
- [4] O. Hegazy, O.S. Soliman, and M. Abdul Salam, "A Machine Learning Model for Stock Market Prediction", *International Journal of Computer Science and Telecommunications*, Vol. (4), Issue (12), pp. 17-23, December 2013.
- [5] O. Hegazy, O.S. Soliman, and M. Abdul Salam, "LSSVM-ABC Algorithm for Stock Price Prediction", *International Journal of Computer Trends and Technology (IJCTT)*, Vol. (7), Issue (2), pp. 81-92, Jan 2014.
- [6] O. Hegazy, O.S. Soliman, and M. Abdul Salam, "Optimizing LS-SVM using Modified Cuckoo Search algorithm (MCS) for Stock Price Prediction", *International Journal of Advanced Research in Computer Science and Management Studies*, Vol. (3), Issue (2), pp. 204-224, February 2015.
- [7] O. Hegazy, O.S. Soliman, and M. Abdul Salam, "Comparative Study between FPA, BA, MCS, ABC, and PSO Algorithms in Training and Optimizing of LS-SVM for Stock Market Prediction", *International Journal of Advanced Computer Research* Vol.(5), Issue (18), pp.35-45, March-2015.
- [8] O. Hegazy, O.S. Soliman, and M. Abdul Salam, "FPA-ELM Model for Stock Market Prediction", *International Journal of Advanced Research in Computer Science and Software Engineering*, Vol.(5), Issue (2), pp.1050-1063, February 2015.
- [9] R. Salem, M. Abdul Salam, H. Abdelkader and A. Awad Mohamed, "An Artificial Bee Colony Algorithm for Data Replication Optimization in Cloud Environments," in *IEEE Access*, vol. 8, pp. 51841-51852, 2020, doi: 10.1109/ACCESS.2019.2957436.
- [10] EM Badr, M. Abdul Salam, M Ali, H Ahmed, Social Media Sentiment Analysis using Machine Learning and Optimization Techniques, *International Journal of Computer Applications (0975 – 8887) Volume 178 – No. 41, August 2019.*
- [11] J. Leskovec and J.J. McAuley. 2012. Learning to discover social circles in ego networks. In *Advances in neural information processing systems*. 539–547.
- [12] S. Schifferes, N. Newman, N. Thurman, verifying news through social media: Developing a user-centred tool for professional journalists. *Digital Journalism* 2, 3 (2014), 406–418.
- [13] A. Ceron, L. Curini, S.M Iacus, and G. Porro. 2014. Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' political preferences with an application to Italy and France. *New Media & Society* 16, 2 (2014), 340–358.
- [14] AP. Engelbrecht (2006), *Fundamentals of computational swarm intelligence*. Wiley, Hoboken.
- [15] M. Dorigo, LM. Gambardella (1997), Ant colony system: a cooperative learning approach to the traveling salesman problem. *IEEE Trans Evol Comput* 1(1):53–66.
- [16] S. Mirjalili, S.M. Mirjalili, and A. Lewis, Grey wolf optimizer. *Advances in Engineering Software*, 69:46–61, 2014.
- [17] R. Eberhart and J. Kennedy, A new optimizer using particle swarm theory. In *Micro Machine and Human Science, 1995. MHS'95., Proceedings of the Sixth International Symposium on*, pages 39–43. IEEE, 1995.
- [18] E. Emary K.K.A. Ghany B. Parv M.A. Salam, H.M. Zawbaa, A hybrid dragonfly algorithm with extreme learning machine for prediction. *International Symposium on INnovations in Intelligent SysTems and Applications (INISTA)*, pages 1–6, 2016.
- [19] P. Dass, H. Sharma, J.C. Bansal, and K.E. Nygard, Meta heuristics for prime factorization problem. In *Nature and biologically inspired computing (NaBIC), 2013 World Congress on*, pages 126–131. IEEE, 2013.
- [20] M. Mishra, U. Chaturvedi, and S.K. Pal, A multithreaded bound varying chaotic firefly algorithm for prime factorization. In *Advance Computing Conference (IACC), 2014 IEEE International*, pages 1322–1325. IEEE, 2014.
- [21] EC Laskari, GC Meletioui, and MN Vrahatis. Problems of cryptography as discrete optimization tasks. *Nonlinear Analysis: Theory, Methods & Applications*, 63(5):e831–e837, 2005.
- [22] M. Abdul Salam, H.M. Abdel Kader, " Evaluation of Differential Evolution and Particle Swarm Optimization Algorithms at Training of Neural Network for Stock Prediction", *Int. Arab. J. e Technol.* 2 (3), 145-151, 2012.
- [23] A Gupta, J.Pruthi, N.Sahu" Sentiment Analysis of Tweets using Machine Learning Approach " *International Journal of Computer Science and Mobile Computing*, Vol.6 Issue.4, April- 2017, pg. 444-458.
- [24] L. B, J. Feng. "Robust Sentiment Detection on Twitter from Biased and Noisy Data". *COLING 2010: Poster Volume*, pp. 36-44.
- [25] B. Naiknaware, B. Kushwaha, S. Kawathekar , "Social Media Sentiment Analysis using Machine Learning Classifiers" *International Journal of Computer Science and Mobile Computing*, Vol.6 Issue.6, June- 2017, pg. 465-472.
- [26] A. Bifet and E. Frank, "Sentiment Knowledge Discovery in Twitter Streaming Data", In *Proceedings of the 13th International Conference on Discovery Science, Berlin, Germany: Springer, 2010*, pp. 1-15.
- [27] A. Andrew, "An Introduction to Support Vector Machines and Other Kernel- based Learning Methods" Nello Christianini and John Shawe- Taylor. *An Introduction to Support Vector Machines and Other Kernel- based Learning Methods*. Cambridge University Press, 2000. xiii + 189 pp., ISBN: ISBN 0- 521- 78019- 5 Hardback:£27.50", *Kybernetes*, vol. 30, no. 1, pp. 103-115, 2001.
- [28] E. AL-Daoud, "Integration of Support Vector Machine and Bayesian Neural Network for Data Mining and Classification", *World Academy of Science, Engineering and Technology*, vol. 64, 2010.
- [29] A. Go, R. Bhayani, L.Huang. "Twitter Sentiment Classification Using Distant Supervision". *Stanford University, Technical Paper, 2009.*
- [30] M. Shad Akhtar ,D. Gupta ,A. Ekbal & Bhattacharyya, P. (2017). Feature selection and ensemble construction: A twostep method for aspect based sentiment analysis. *KnowledgeBased Systems*, 125, 116-135.

- [31] H. Kang, S.J. Yoo and D. Han, "Senti-lexicon and improved Naïve Bayes algorithms for sentiment analysis of restaurant reviews", *Expert Systems with Applications*, Elsevier, vol. 39, no. 5, pp. 6000–6010, 2012.
- [32] M. Koppel and j. Schler, "the importance of neutral examples for learning sentiment", *computational intell.*, vol. 22, no. 2, pp. 100-109, 2006.
- [33] D. Romero, B. Meeder and J. Kleinberg, "Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on Twitter", *Proceedings of the 20th international conference on World wide web*, pp. 695--704, 2011.
- [34] M. Anjaria and R. Guddeti, "Influence factor based opinion mining of Twitter data using supervised learning", *Communication Systems and Networks (COMSNETS), 2014 Sixth International Conference on*, pp. 1--8, 2014.
- [35] J. Martineau and T. Finin, "Delta TFIDF: An Improved Feature Space for Sentiment Analysis", *Proceedings of the Third International ICWSM Conference*, vol. 9, 2009.
- [36] SM. Mohammad, S. Kiritchenko and X. Zhu, "NRC-Canada: Building the State-of- the-Art in Sentiment Analysis of Tweets", *arXiv preprint arXiv:1308.6242*, 2013.
- [37] P. Barnaghi, J.G. Breslin and P. Ghaffari, "Opinion Mining and Sentiment Polarity on Twitter and Correlation between Events and Sentiment", *2016 IEEE Second International Conference on Big Data Computing Service and Applications*.
- [38] B.S. Halima and S Chitrakala, "Trending Topic Analysis Using Novel Sub Topic Detection Model", (IEEE) ISBN-978-1-4673-9745-2, 2016.
- [39] S. Yuan, J. Wu, L. Wang and Q. Wang, "A Hybrid Method for Multi-class Sentiment Analysis of Micro-blogs", ISBN- 978-1-5090-2842-9, 2016.
- [40] P.N. Suganthan G.B. Huang Q.Y. Zhu, A.K. Qin. Evolutionary extreme learning machine. *Pattern Recognition*, 38(10):1759–1763, 2005.
- [41] L. Zhiping C. Jiuwen. Extreme learning machines on high dimensional and large data applications: A survey. *Mathematical Problems in Engineering*, 2015(1):1–13, 2015.
- [42] C.K. Siew G.B. Huang, Q.Y. Zhu. Extreme learning machine: theory and applications. *Neurocomputing*, 70(1):489–501, 2006.