An Efficient Hybrid Approach for Twitter Sentiment Analysis based on Bidirectional Recurrent Neural Networks

Khadija Shazly
Faculty of Computer and Information, Mansoura University, Egypt

Marwa Eid
Communications and Electronics Engineering Department
Delta Higher Institute for Engineering & Technology, Talkha 35111, Egypt

Hanaa Salem
Dept. of Comm. & Comp., Faculty of Engineering
Delta University for Science & Technology

ABSTRACT
Many optimization problems from various applications have been solved by many algorithms such as Grey Wolf Optimizer (GWO), Genetic Algorithm (GA), Whale Optimization Algorithm (WOA) and Particle Swarm Optimization (PSO). Sentiment Analysis (SA) is used to evaluate the polarity of reviews. In SA, feature selection phase is an important, as the best way to solve all optimization problems not happen yet so this paper proposes a hybrid approach that combines three modified hybrid algorithms [(GWO), (PSO) and (GA)], its name (HWPG) .To reduce the search space filter features selection, Information Gain (IG) has been used. (HWPG) used to select the best features for training Bidirectional Recurrent Neural Networks (BRNN) classifier. Arabic benchmark dataset which was collected from twitter on different topics used in our experimental. The proposed algorithm is compared with three well-known optimization algorithms the experiments and comparisons result to evaluate the quality and effectiveness of the (HWPG)

Keywords
Feature Selection, Genetic Algorithm, In Information Gain, Optimization problems, Particle Swarm Optimization, Sentiment analysis

1. INTRODUCTION
Due to the rapid growth of employing the internet of things (IoT) applications, huge amounts of the generated data are flooded our world. However, the biggest challenge is how to extract specific knowledge from these huge data sets [1].

Nowadays, big data has involved in several applications such as Internet of Things (IoT), medical imaging [2], bioinformatics, E-banking, and Online Social Networks (OSN). Hence, several researches had been concerned with data encryption [3], security systems [4,5,6], and Sentiment Analysis (SA). The real-world optimization problems are complex with high dimensional search space and therefore challenging to solve [7,8,9]. The typical machine learning techniques are suffered from several drawbacks like the class imbalance problem.

SA is a classification process, which is in light of deciding the extremity of a given audit into either positive, negative, or nonpartisan, as per the communicated sentiment [10].The principle point of SA is to discover the situation of survey author at archive level, sentence level, or viewpoint level [11,12,13]. Currently, one of the most common OSN is the Twitter application with more than 200 million users. These days, SA is broadly utilized in breaking down audits of various information spaces such item audits, film surveys, inn surveys, café audits, and some more such as engineering [14], machine learning [15], graphic rendering [16], bioinformatics [17], eMarketing [18], scheduling [19]. In optimization, the best optimal solution is a main goal, according to the problem description how to select the available solutions from a given problem [20]. The SA procedure can be mechanized utilizing various kinds of Artificial Iintelligence (AI) classifiers, which can likewise be improved with feature selection processes. In this paper, an efficient approach HWPG that combines three modified optimizer algorithms [(GWO), (GA) and (PSO)] is proposed to reduce the irrelevant features. Then, to enhance the classification accuracy, we used a Bidirectional Recurrent Neural Network (BRNN) for distinguishing the Arabic benchmark dataset which was collected from twitter OSN [21,22,24]. After comparisons, the proposed method can get the fit balance between the different aspects of exploration and exploitation in less elapsed time for processing.

1.1 Grey Wolf Optimizer (GWO)
Grey wolves are one of zenith predators, zenith predators are at the highest point of the natural pecking order. in the real live Grey wolves commonly bring closer in groups. The average of group numbers may be from 5–12. The most fascinating thing about grey wolves is that they have an extremely severe social driving pecking order. The leaders called alphas, are a male and a female. The alpha is responsible for settling on significant choices to the group such as hunting, resting place, etc. The whole group should obey the alpha's orders [12,23]. The alpha wolf is additionally called the leader wolf since the whole group should ought to comply with the alpha's requests. Beta is the second level in the hierarchy of grey wolves. Beta is the second level in the chain of command of dim wolves. The betas are second rate wolves that are liable for helping the alpha in settling on the correct choices or other gathering exercises. Omega is the least positioning of grey wolf. The omega assumes the job of substitute. Mega wolves consistently need to capitulate to all other predominant wolves. Also, they are the last wolves that are permitted to eat, as shown in figure (1)

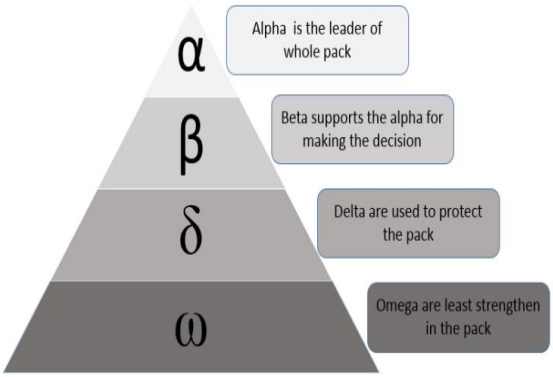

Figure [1] Hierarchy of Gray Wolves

The position of each wolf is updated using the following equations:

$$\vec{D} = \left|\vec{C} \,.\, \overrightarrow{X_p}\,(t) - \vec{X}\,(t)\right| \tag{1}$$

$$\vec{X}\,(t+1) = \overrightarrow{X_p}\,(t) - \vec{A} \,.\, \vec{D}$$

Where t refers to the current iteration, $\vec{A}$ and $\vec{C}$ are coefficient vectors, $\overrightarrow{X_p}$ is the preposition, and $\vec{X}$ is the position of the gray wolf. The vectors are calculated using the following equation:

$$\vec{A} = 2\vec{a} \,.\, \overrightarrow{r_1} - \vec{a} \tag{2}$$

$$\vec{C} = 2.\overrightarrow{r_2}$$

1.2 Genetic Algorithm

Evolutionary algorithm is the base idea of GA. Without any information about data, global search method can search in a large-scale, complex search space. Thus, it is known as a suitable method for solving optimization problems. This iterative algorithm starts with a population of chromosomes, each called an individual. By applying a set of operators to the individuals in each iteration, a new population is created. The selection, crossover, and mutation have been defined as three main operators for constructing new chromosomes. First, the selection operator is executed and two chromosomes (as the parents) are selected based on a fitness function. Second, the crossover operator is applied in order to construct two of spring that inherit some characteristics of their parents. Third, the mutation operator is performed that can change some characteristics of each ofspring. It is expected that the new population in the current iteration will suggest better solutions than the previous population in the former iteration. This procedure is performed until one of the stop criteria like, a fixed number of iterations reached, is met. In each iteration, the quality of the produced candidate solution via each individual is measured by a fitness function.

1.3 Particle Swarm Optimization

The PSO algorithm, it copies the insight of bird swarms in real live. velocity is the position Change of a particle. During time the of position particle are change. At the flight, particle's velocity is randomly accelerated toward its previous best position and toward a neighborhood best solution. As shown in the next equations

$$v_i^{k+1} = v_i^k + c_1 r_1 \left(Pbest_i^k + x_i^k\right) + c_2 r_2 \left(g_{best} - x_i^k\right) \tag{3}$$

$$x_i^{k+1} = x_i^k + v_i^{k+1} \tag{4}$$

2. THE PROPOSED FRAMEWORK

The proposed HWPG framework working through two phases:

First phase is preprocessing dataset the second phase is apply optimizer algorithms as shown in figure [2].

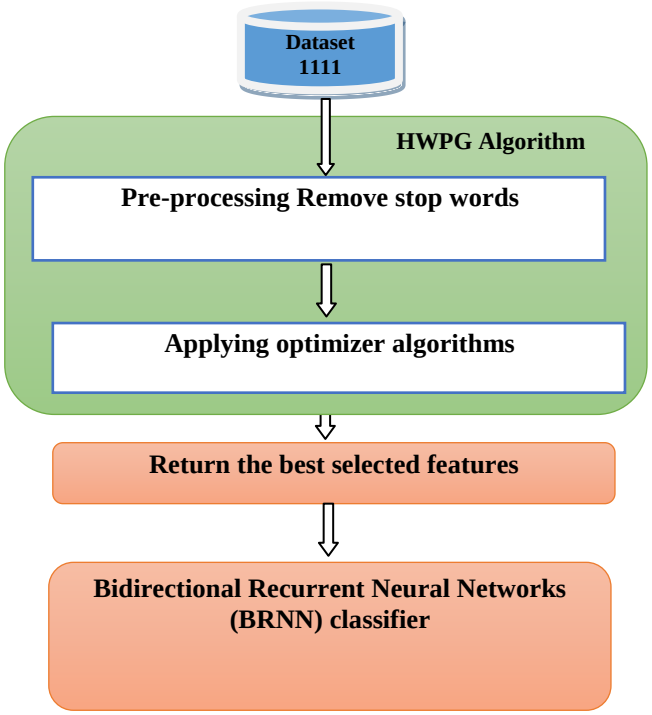

Figure [2] The proposed Framework Phases.

2.1 Preprocessing Dataset

The preprocessing phase incorporates exceptionally essential procedures that help in delivering a prepared to utilize dataset for preparing and testing purposes. In the stop words removal process, words that don't convey significant importance for the content and characterization procedure will be expelled. Such as:

On	علي
In	في
From	من

the tokens that offer a similar root or stem are put together as they all have related implications. For instance, the accompanying three Arabic words are made

[(Office, " مكتب "), (Library, " مكتبة "), (Writing, " كتابة ")] from a similar Arabic root word[(Write, " كتب ")] . Once the Arabic words are preprocessed as indicated by the past steps, each audit will be changed into a vector.

2.2 Optimizer Algorithms

The proposed approach is equipped with three powerful algorithms. The first algorithm is PSO in which individuals are moving influenced by their local best positions and by the global best position. The second optimizer in our proposed hybrid approach is Grey Wolf Optimizer. GWO is a swarm-

based meta-heuristic optimizer that mimics the social hierarchy and the foraging behavior of the grey wolves. Individuals in GWO move influenced by the place of the three leader's alpha, beta, and delta. the third is genetic algorithm which accrue convergence by enhance the position of a specific solution around randomly selected leaders this is call Mutation operator random changes one or more components of the offspring. The (HWPG) algorithm shown in figure (2)

```
Initialize the grey wolf population Xi (i = 1, 2, ..., n)
Calculate α, β, and δ
Calculate a as per Eq (3.2)
for each search agent
Update  A, and C
Calculate Xα, Xβ, and Xδ
crossover ( Xα , Xβ , Xδ)
        solution (t+1) = Mutate
end for
While (t < Max_iter)
if  t % 2 == 0:
           Calculate fitness(3)
Else
Calculate fitness using Eq (4)
Update A, a, C
Evaluate all particles using the objective function
Update the positions of the three best agents α;β;δ
 t=t+1
        return α
```

Fig. 3 Pseudocode of the Proposed (HWPG)

Information Gain

In Information Gain (IG), the pertinence of each component to the class marks is evaluated by following condition, whereby IG positions each feature as indicated by its entropy and chooses the most significant highlights as indicated by the prespecified limit.

$$IG(f) = -\sum_{n=1}^{m} P(C_n)\, log\, P(C_n) + P(f)\ \sum_{n=1}^{m} P(C_n|f)\ log\, P(C_n|f) + P(\bar{f})\ \sum_{n=1}^{m} P\left(C_n|\bar{f}\right) log\, P\left(C_n|\bar{f}\right) \quad (5)$$

f is the feature, $P(C_n)$ is the percentage portion of reviews, $P(f)$ percentage portion, $P(f)$ is the percentage portion of reviews in which the feature *f* does not exist, $P(C_n|\, f)$ is the percentage portion of reviews, C_n class category

BRNN

HWPG is utilized to improve the presentation of BRNN classifier and dispose of the excess and immaterial highlights. BRNN is chosen dependent on its presentation in the past works, which has demonstrated to outflank other AI classifiers for SA. Moreover, to give solid outcomes and maintain a strategic distance from overfitting issues

(BRNN) can be considered as an extension of the unidirectional recurrent neural networks by adding a second hidden layer, where the connections between the hidden to hidden layers are in the opposite direction. Consequently, this model can exploit data from both directions, the past and the future. The output ct can be calculated by

Configurations parameters of HWPG

$$Q(c_t|\{a_i\}_{i\neq t}) = \sigma(W_c^f h_t^f + W_c^b h_t^b + b_c) \quad (6)$$

Where

$$h_t^f = \tanh(W_h^f h_{t-1}^f + W_x^f x_t + b_h^f) \quad (7)$$

$$h_t^b = \tanh(W_h^b h_{t+1}^b + W_a^b a_t + b_h^b) \quad (8)$$

BRNN parameter optimization

Weights in a BRNN have two main roles; the first one is, deciding how much the output is affected by the input and the second one is, controlling the learning rate of the hidden layers. Exactly as slope in linear regression, where the output is calculated by multiplying the weights to the inputs then added up. Weights are numerical values that control how much neurons are affecting each other. For any neuron, if the inputs are a1, a2, and a3, and weights applied to them are w1, w2, and w3. The output is:

$$c = f(a) = \sum_{j=1}^{n} a_j w_j \quad (9)$$

Where n is the number of inputs. Generally, the weighted sum can be calculated by performing this array multiplication. Bias is an additional variable that can be used to adjusting the output along with the weighted sum of the inputs to the neuron. The final output of a neuron is:

$$c = f(a) = \sum_{j=1}^{n} a_j w_j + b \quad (10)$$

Where b is the bias

Each dataset is divided into three randomly equal-size parts: training, validation, and test. Training is used to train BRNN classifier during the learning phase. As shown in table 1

Table 1: Configurations parameters

Parameter	value
No of search agents	10
No of iterations	80
Problem dimension	Number of features in the data
Search domain	[0 1]
No. repetitions of runs	20
Inertia factor of PSO	0.1
α Parameter in the fitness function	0.99
β Parameter in the fitness function	0.01

3. RESULTS AND DISCUSSION

Dataset

which was gathered from twitter on various themes, for example, expressions and legislative issues. This Arabic twitter corpus contains 2000 tweets audits with 1000 positive tweets and 1000 negative tweets. These gathered tweets were composed utilizing both Modern Standard Arabic (MSA)The measurements of Arabic twitter dataset are appeared in Table 2. What's more, the gathered tweets were preprocessed by expelling the rehashed letters, right incorrect spellings words, and standardization of Arabic letters

Table 2: Statistics of Database

Review Polarity Criteria	Positive Reviews	Negative Reviews
Total number of tweets	1000	1000
Total number of words	7189	9769
Average number of words in each tweet	7.19	9.97
Average number of characters in each tweet	40.04	59.02

To assess and research the exhibition and viability of the proposed HWPG calculation. The analyses actualized utilizing scikit-learn Machine Learning in Python. This procedure is rehashed multiple times. Ultimately, the average accuracy, average fitness, and average number of selected features over 10 runs are accounted for. The HWPG calculation is contrasted and other surely understood calculations counting PSO, GA, and WOA. Besides, in this work BRNN classifier is applied to the various investigations as exhibited in the accompanying Tables from these different investigations were directed utilizing Arabic twitter sentiment analysis datasets, by contrasting our proposed HWPG calculation with well-known algorithms. The directed analyses exhibited that the proposed HWPG beat different algorithms as far as SA classification accuracy and fitness value, reduces the selected features. Thus, the proposed HWPG done Arabic SA feature selection task adequately

Table 3. The sentiment classification accuracy

No. of Features	Features Proportion %	BRNN	SVM	KNN
2257	100	**89.97**	81.73	69.89

Table 4. HWPG Classification Accuracy

Number of input features based on IG ratio	IG ratio (%)	Average accuracy (%)
135	6	**88.75**
271	12	85.64
406	18	84.79

Table 5. Average Classification Accuracy in 10 Runs

Number of input features based on IG ratio	IG ratio	HWPG	PSO	GA	WOA
135	6	**88.89**	83.54	84.35	87.84
271	12	**87.79**	85.61	85.81	86.78
406	18	**86.64**	86.12	86.38	85.41

Table 5. Average Number of Selected Features in 10 Runs

Number of input features based on IG ratio	IG ratio	HWPG	PSO	GA	WOA
135	6	97	**70**	94	88
271	12	**129**	142	136	139
406	18	**181**	209	201	197

Table 6. Average Fitness in 10 Runs

Number of input features based on IG ratio	IG ratio	HWPG	PSO	GA	WOA
135	6	**13.71**	16.23	17.58	15.98
271	12	**10.98**	12.56	13.78	13.56
406	18	**9.97**	11.87	12.41	11.74

4. CONCLUSIONS

This paper concentrated on Arabic SA to add to its cutting edge. This is endeavored to improve the standard HWPG algorithm by improve introduction GWO and upgrade nearby pursuit ability including transformation, utilizing (IG) and training (BRNN) classifier to choose the best features than assess and rank it. We utilized and assessed the HWPG in correlation with WOA, PSO, and GA the predominance of HWPG precision brings about

examination with different algorithms as demonstrated by strong text style.

5. REFERENCES

[1] El-Kenawy, E. S. M. T., El-Desoky, A. I., & Sarhan, A. M. (2014). A bidder strategy system for online auctions trust measurement. International Journal of Strategic Information Technology and Applications (IJSITA), 5(3), 37-47.

[2] El-kenawy, E. S. M., El-Desoky, A. I., & Al-rahamawy, M. F. (2012). Distributing Graphic Rendering using Grid Computing with Load Balancing. International Journal of Computer Applications, 975, 888.

[3] El-kenawy, E. S. M. T. (2019). A Machine Learning Model for Hemoglobin Estimation and Anemia Classification. International Journal of Computer Science and Information Security (IJCSIS), 17(2).

[4] Hassib, E. M., El-Desouky, A. I., El-kenawy, E. S. M., & Elghamrawy, S. (2019). An Imbalanced Big Data Mining Framework for Improving Optimization Algorithms Performance. IEEE Access.

[5] El-Knawy, E. S. M. T., & El-Desoky, A. I. (2016). TRUST MEASUREMENT FOR ONLINE AUCTIONS: PROPOSAL OF NEW MODEL. INTERNATIONAL JOURNAL OF INNOVATIVE COMPUTING INFORMATION AND CONTROL, 12(2), 385-394.

[6] El-sayed, M., El-Desoky, A. I., & Sarhan, A. M. (2014). A bidder behavior learning intelligent system for trust measurement. International Journal of Computer Applications, 89(8).

[7] Reham Arnous, El-Sayed Towfek M El-kenawy and M Saber. A Proposed Routing Protocol for Mobile Ad Hoc Networks. International Journal of Computer Applications 178(41):26-30, August 2019.

[8] El-Sayed Towfek M El-kenawy. Trust Model for Dependable File Exchange in Cloud Computing. International Journal of Computer Applications 180(49):22-27, June 2018.

[9] El-Sayed Towfek M El-kenawy, M Saber and Reham Arnous. An Integrated Framework to Ensure Information Security Over the Internet. International Journal of Computer Applications 178(29):13-15, July 2019

[10] H. Hassan, A. I. El-Desouky, A. Ibrahim, E. M. El-kenawy and R. Arnous, (2020) "Enhanced QoS-based Model for Trust Assessment in Cloud Computing Environment," in IEEE Access. doi: 10.1109/ACCESS.2020.2978452

[11] Hassib, E. M., El-Desouky, A. I., Labib, L. M., & El-kenawy, E. S. M. WOA+ BRNN: An imbalanced big data classification framework using Whale optimization and deep neural network. Soft Computing, 1-20.

[12] E.-S. El-Kenawy and M. Eid, "Hybrid gray wolf and particle swarm optimization for feature selection," INTERNATIONAL JOURNAL OF INNOVATIVE COMPUTING INFORMATION AND CONTROL, vol. 16, no. 3, pp. 831–844, 2020.

[13] El-Kenawy, E. S. M., Eid, M., & Ismail, A. H. A New Model for Measuring Customer Utility Trust in Online Auctions. *International Journal of Computer Applications*, *975*, 8887.

[14] E. M. El-Kenawy and M. Saber , "Design and implementation of accurate frequency estimator depend on deep learning" International Journa l of Engineering &Technology (IJET), vol. 9 , Issue 2, PP. 367-377 , 2020 , DOI: 10.14419/ijet.v9i2.30473

[15] E. M. El-Kenawy, M. M. Eid, M. Saber and A. Ibrahim, "MbGWO-SFS: Modified Binary Grey Wolf Optimizer Based on Stochastic Fractal Search for Feature Selection," in IEEE Access, vol. 8, pp. 107635-107649, 2020, doi: 10.1109/ACCESS.2020.3001151.

[16] Tharwat, A., Ibrahim, A., Hassanien, A. E., & Schaefer, G. (2015, June). Ear recognition using block-based principal component analysis and decision fusion. In International Conference on Pattern Recognition and Machine Intelligence (pp. 246-254). Springer, Cham.

[17] Ibrahim, A., Tominaga, S., & Horiuchi, T. (2009, March). Material classification for printed circuit boards by spectral imaging system. In International Workshop on Computational Color Imaging (pp. 216-225). Springer, Berlin, Heidelberg.

[18] Ibrahim, A., Mohammed, S., & Ali, H. A. (2018, February). Breast cancer detection and classification using thermography: a review. In International Conference on Advanced Machine Learning Technologies and Applications (pp. 496-505). Springer, Cham.

[19] M. M. Fouad, A. I. El-Desouky, R. Al-Hajj and E. M. El-Kenawy, "Dynamic Group-based Cooperative Optimization Algorithm," in IEEE Access, doi: 10.1109/ACCESS.2020.3015892.

[20] Ibrahim, A., Tharwat, A., Gaber, T., & Hassanien, A. E. (2018). Optimized superpixel and AdaBoost classifier for human thermal face recognition. Signal, Image and Video Processing, 12(4), 711-719.

[21] Ibrahim, A., Tominaga, S., & Horiuchi, T. (2009, May). Unsupervised Material Classification of Printed Circuit Boards Using Dimension-Reduced Spectral Information. In MVA (pp. 435-438).

[22] Gaber, T., Tharwat, A., Ibrahim, A., Snáel, V., & Hassanien, A. E. (2015, September). Human thermal face recognition based on random linear oracle (rlo) ensembles. In 2015 International Conference on Intelligent Networking and Collaborative Systems (pp. 91-98). IEEE.

[23] El-kenawy, E. S. M. T. (2018). Solar Radiation Machine Learning Production Depend on Training Neural Networks with Ant Colony Optimization Algorithms. IJARCCE, 7(5). DOI10.17148/IJARCCE.2018.751.

[24] El-kenawy, E. S. T., El-Desoky, A. I., & Al-rahamawy, M. F. (2012). Extended max-min scheduling using petri net and load balancing. *Int. J. Soft Comput. Eng.(IJSCE)*, *2*(4), 198-203.

[25] Ibrahim, A., & Tharwat, A. (2014). Biometric authentication methods based on ear and finger knuckle images. International Journal of Computer Science Issues (IJCSI), 11(3), 134.