Improved whale optimization algorithm for feature selection in Arabic sentiment analysis

Article *in* Applied Intelligence · May 2019

DOI: 10.1007/s10489-018-1334-8

4 authors:

Mohammad Tubishat
23 PUBLICATIONS 567 CITATIONS

Mohammad A. M. Abushariah
University of Jordan
50 PUBLICATIONS 959 CITATIONS

Norisma Idris
University of Malaya
80 PUBLICATIONS 1,279 CITATIONS

Ibrahim Aljarah
University of Jordan
155 PUBLICATIONS 8,533 CITATIONS

Applied Intelligence
https://doi.org/10.1007/s10489-018-1334-8

Improved whale optimization algorithm for feature selection in Arabic sentiment analysis

Mohammad Tubishat[1] · Mohammad A. M. Abushariah[2] · Norisma Idris[1] · Ibrahim Aljarah[3]

Abstract
To help individuals or companies make a systematic and more accurate decisions, sentiment analysis (SA) is used to evaluate the polarity of reviews. In SA, feature selection phase is an important phase for machine learning classifiers specifically when the datasets used in training is huge. Whale Optimization Algorithm (WOA) is one of the recent metaheuristic optimization algorithm that mimics the whale hunting mechanism. However, WOA suffers from the same problem faced by many other optimization algorithms and tend to fall in local optima. To overcome these problems, two improvements for WOA algorithm are proposed in this paper. The first improvement includes using Elite Opposition-Based Learning (EOBL) at initialization phase of WOA. The second improvement involves the incorporation of evolutionary operators from Differential Evolution algorithm at the end of each WOA iteration including mutation, crossover, and selection operators. In addition, we also used Information Gain (IG) as a filter features selection technique with WOA using Support Vector Machine (SVM) classifier to reduce the search space explored by WOA. To verify our proposed approach, four Arabic benchmark datasets for sentiment analysis are used since there are only a few studies in sentiment analysis conducted for Arabic language as compared to English. The proposed algorithm is compared with six well-known optimization algorithms and two deep learning algorithms. The comprehensive experiments results show that the proposed algorithm outperforms all other algorithms in terms of sentiment analysis classification accuracy through finding the best solutions, while its also minimizes the number of selected features.

Keywords Arabic sentiment analysis · Support vector machine · Information gain · Whale optimization algorithm

1 Introduction

Sentiment Analysis (SA) is a classification process, which is based on determining the polarity of a given review into either positive, negative, or neutral, according to the expressed opinion [1]. The main aim of SA is to find the position of review writer at document level, sentence level, or aspect level [2]. Nowadays, SA is widely used in analyzing reviews of different data domains such product reviews, movie reviews, hotel reviews, restaurant reviews, and many more. SA helps people and organization in decision making, However, the huge volume of available reviews, which is generated by the internet users is increasing rapidly. Therefore, this huge volume of data requires an automatic tool for analyzing the semantic orientation of the given reviews and cannot be conducted manually. To automate the process of review analysis and polarity determination, there are many methods have been used such as supervised and un-supervised methods. The SA process can be automated using different types of

✉ Norisma Idris
norisma@um.edu.my

Mohammad Tubishat
mtubishat@siswa.um.edu.my

Mohammad A. M. Abushariah
m.abushariah@ju.edu.jo

Ibrahim Aljarah
i.aljarah@ju.edu.jo

1 Department of Artificial Intelligence, Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia

2 Computer Information Systems Department, King Abdullah II School of Information Technology, The University of Jordan, Amman, Jordan

3 Business Information Technology Department, King Abdullah II School for Information Technology, The University of Jordan, Amman, Jordan

Published online: 26 November 2018

machine learning classifiers, which can also be improved with feature selection processes.

The huge volume of available reviews contains relevant and irrelevant features to the given used classifiers, which in turn decrease their performance. Therefore, feature selection algorithms for selecting the most important features and removing irrelevant data are required. Based on these feature selection algorithms, the most informative and relevant features will be selected, which in turn improve the performance of sentiment classification process.

Feature selection techniques can be classified into filter, wrapper, and hybrid based. The filter-based feature selection techniques are also known as traditional feature selection techniques such as Information Gain (IG). IG is used to reduce the number of features to be used by the classifiers and rank features' relevance in accordance with the classifier labels. However, the problem of traditional filter feature selection techniques is that they cannot interact with the used classifier directly. On the other hand, wrapper-based feature selection adopts optimization algorithms that have direct interaction with both the used classifier and features, but the computation time requirement is high. In addition, the hybrid-based feature selection techniques are used to take advantage of both filter and wrapper-based techniques and resolve their shortcomings [3, 4]. The process of optimization is existing in different application areas such as engineering, medical, agriculture, computer science, and many more. In optimization the main aim is to determine and select the optimal solution to a given problem from the pool of available solutions according to problem specifications. Furthermore, in theses optimization algorithms there is an objective function that should be minimized or maximized based on the problem to be solved. In this paper, WOA algorithm is used based on its capability and low number of parameters in comparison with other well recognized optimization algorithms [5]. WOA mimics the hunting behavior of humpback whale which use bubble-net hunting technique to surround and catches the preys. The WOA has been applied to different benchmark functions and it shows good results in comparison with other well recognized algorithms according to paper authors in [5].

Recently, WOA has been used to solve several problems, such as a research conducted by [6], WOA was used to find the optimal weights for training the neural network. On other hand, in work proposed by [7], they developed a multi objective version of WOA and applied it to the problem of forecasting the wind speed. Moreover, WOA was also used in [8] work to determine the optimal placement and size of capacitors which used in radial system. Furthermore, in work of [9], they utilized WOA in the problem of finding the optimal size used by distributed generator. In addition, in research by [10], they take the benefit of using WOA for MRI image segmentation. All these mentioned works used the standard WOA algorithm, however, the standard WOA algorithm is like other optimization algorithm which has the tendency of being trapped in local optima. To solve this problem, we used the Elite Opposition-Based Learning (EOBL) technique [11] to enhance the quality of solutions in the initialization phase of WOA. The basic idea of EOBL is to consider the current solution and its elite opposite solution and take the best one according the fitness values of the solutions.

Opposition Based Learning (OBL) technique has been used in many problems, for example in [12], the authors used OBL to solve the problem of slow convergence in Harmony Search (HS) algorithm. Further, in work conducted in [13], they utilized OBL for improving BAT algorithm slow convergence and increase the solutions diversity. In addition, in research by [14], they mentioned that many optimization algorithms such as Genetic Algorithm (GA), Particle Swarm Optimization (PSO), Gravitational Search Algorithm (GSA), and Simulated Annealing (SA) used OBL to improve its performance.Furthermore, OBL was applied in [15] and Generalized OBL was applied to optimizaton algorithms such as works in [16–22]. Another improvement on basic OBL technique include EOBL which developed in [11] and used in many studies. For example, in research conducted in [23], they used EOBL to improve the population diversity of Flower Pollination Algorithm (FPA). In [24], they utilized EOBL technique for improving Grey Wolf Optimizer (GWO) population diversity and convergence speed. Moreover, in research by [25], they applied EOBL strategy to enhance Cuckoo Search (CS) algorithm by making a balance between the exploitation and exploration. In our proposed Improved Whale Optimization Algorithm (IWOA), we used EOBL at initialization phase to enhance the quality of initial solutions generated in standard WOA. Moreover, Differential Evolution (DE) operators [26] which includes mutation, crossover, and selection was used at the end of each IWOA iteration to improve the local search capability of standard WOA and find more promising regions in the search space. DE algorithm used by many researches, for example in research by [27], they hybridized DE with Artificial Bee Colony (ABC) to improve the convergence speed and enhance the balance between the exploitation and exploration. In research of [28], they combined DE operators with Fireworks Algorithm (FA) at the end of each iteration in FA by taking the advantages good exploration capability of DE with good exploitation by FA. In [29] paper, they hybridized DE with Cultural Algorithms (CA) by taking the advantage of exploitation capability of DE together with the good exploration capability of CA. Lastly, to reduce the search space explored by IWOA algorithm, we used Information Gain (IG) as a feature ranking technique. The IG ranks the features in the used dataset and gives each feature score according to its importance.

Based on our literature investigation, research efforts have been devoted to SA for different languages such as

English and Chinese. However, despite the importance of Arabic language, its research is lacking in many aspects and research efforts on Arabic SA is still in its infancy and require more emphasis worldwide compared to other languages.

In this research, IWOA is an improvement of standard WOA algorithm using EOBL technique and DE evolutionary operators to solve the WOA problem of being stuck in local optima by increasing the diversity of the solutions and improving exploitation of WOA. Further, to reduce the search space explored by WOA and reduce the search time, IG is used to rank features and feed the best features to WOA as well as to discard irrelevant features. IWOA algorithm is used to optimize the feature selection process and improve the sentiment classification in Arabic language. In addition, IWOA algorithm is used to improve the performance of SVM classifier and discard the redundant and irrelevant features. SVM is selected based on its performance in the previous works, which has proven to outperform other machine learning classifiers for SA [30–33]. Furthermore, to provide reliable results and avoid overfitting problems, K-fold cross-validation method was used on the four used Arabic datasets.

In this work, the main contributions are summarized as follows:

- The IG filter feature reduction techniques and WOA optimization algorithm are combined to take the benefits of their advantages and solve their shortcomings.
- A variant version of standard WOA is presented and named IWOA by using EOBL strategy and DE evolutionary operators to improve both the exploration and exploitation of WOA algorithm as the following:

  a. EOBL strategy: used for enhancing WOA diversity of the initial solutions that generated by standard WOA
  b. DE evolutionary operators: used for solving the problem of local optima in standard WOA by using DE evolutionary operators at the end of each WOA iteration to improve the solutions found so far.
- The proposed IWOA tested on four Arabic sentiment analysis datasets and outperforms other algorithms. IWOA outperforms other optimization algorithms which include (WOA, DE, PSO, GA, Ant Lion Optimizer (ALO), and Grasshopper Optimization Algorithm (GOA). In addition, IWOA outperforms deep learning algorithms which include convolutional neural network (CNN) and long short-term memory (LSTM).
- The proposed IWOA proved its effectiveness based on the obtained results and provide quality of solutions.

The reminder of the paper is structured as follows: In Section 2, we discussed the previous related works. In Section 3, we discussed some preliminaries details. In Section 4, we presented the proposed improved algorithm. In Section 5, we discussed the conducted experiments to measure the performance of the algorithm with the results obtained in the experiments. Finally, we have the conclusion in Section 6.

2 Related works

SA can be carried out at three levels which are document level, sentence level, and aspect level. Furthermore, SA based on document and sentence levels are concerned with extracting all opinion words in either the document or sentence, then classify these opinion words as positive, negative, or neutral for the whole document or sentence. On the other hand, SA based on aspect level is concerned with extracting the opinion words related to each aspect and determining the semantic orientation of each aspect individually. In addition, SA techniques can be either unsupervised, semi-supervised, or supervised. Unsupervised techniques do not require training data. However, supervised techniques require labelled data, and the semi-supervised techniques require little training data [1, 2]. Fig. 1 shows a classification of SA levels [2].

Several techniques have been used for Arabic SA by the previous works. For example, in [34] they employed five different machine learning classifiers namely, SVM, Stochastic Gradient Descent (SGD), Mutlinomial Naive Bayes (MNB), Decision Trees (DT), and Naive Bayes (NB) on the Large-Scale Arabic Book Reviews (LABR) dataset. In addition, in their work the MNB classifier reported the best performance compared to all other used classifiers. In addition, they experienced different feature extraction patterns with the used classifiers and achieved the best performance with MNB classifier using unigram. Finally, they utilized GA as a feature selection technique on the used dataset and provided the selected features to MNB classifier, which in turn improved its performance with 85% accuracy.

In research by [35], they investigated the use of three different machine learning classifiers namely, SVM, K-Nearest Neighbors (KNN), and NB on their own collected Arabic sentiment dataset. The used dataset was collected from Twitter, which included 2591 Arabic tweets. Moreover, they

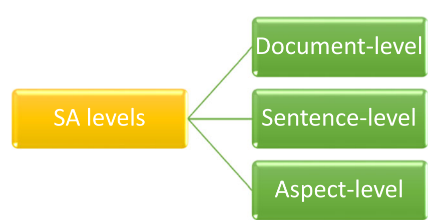

Fig. 1 Classification of SA levels

investigated different combinations between the three classifiers and three different types of weighting techniques. The used weighting techniques were, Term Frequency (TF), Binary Model (BM), and Term Frequency Inverse Document Frequency (TFIDF). They reported that the NB classifier with TF weighting techniques outperformed other combinations' accuracies with 69.97% accuracy, but the SVM classifier and TFIDF weighting technique combinations outperformed other classifiers in term of precision. In the work of [36], the authors investigated and compared the use of four reductive feature selection techniques for Arabic SA. They conducted their experiment on a Twitter Arabic dataset, which included 4800 Arabic tweets. Furthermore, they applied 10 cross validation and achieved the best accuracy of 74% by using combinations of Full Reduct by Attribute Weighting (FRAW) with LEM2 algorithms.

In work conducted by [37], they compared the combination of three different machine learning classifiers with three different features selection techniques for sentiment classification on a public sentiment Arabic dataset. The classifiers used in the experiments are N-gram model, Association Rule Mining (ARM), and Meta classifier. The used classifiers were combined with three different features selection methods, Gini Index (GI), IG, and Chi-square (CHI). In their experiment, they utilized the Opinion Corpus for Arabic (OCA) [38], which contains 500 Arabic comments. Moreover, the authors reported that the combination of CHI feature selection with Meta classifier yields the best accuracy with 90.8% accuracy, which outperforms other combinations. In [39], they improved the CHI feature selection technique, and investigated the different possible combinations and preprocessing. They utilized their own collected dataset, which was collected from Arabic newspapers that is available online with a total of 250 documents that are classified into five topics. Moreover, they investigated the use of either light stemming or hard stemming with the combinations of either the improved CHI or the original CHI features selection with decision tree classifier. As a result, they reported that the combinations of light stemming, improved CHI, and decision tree classifier achieved the best recall with 75.3% value, which in turn outperformed other combinations.

In research by [40], they created a new Arabic dataset solely for SA, which was collected from Yahoo-Maktoob. They also investigated two different machine learning classifiers SVM, and NB using their own dataset. In the preprocessing steps, the TFIDF weighting technique was used followed by stemming as feature selection to reduce the number of features. In a following step, they conducted their experiments by combining each classifier with selected and weighted features. As a result, the reported accuracy of SVM classifier was 68.2%, which outperformed the NB classifier accuracy. In [41], they compared two approaches for Arabic SA namely, corpus-based approach (supervised) and lexicon-based approach (unsupervised). In the first step, they created their own dataset by downloading 2000 Arabic tweets from Twitter about different Arabic topics. In addition, they constructed their own Arabic lexicon for SA through downloading 300 English words from SentiStrength website and translating the collected English words into corresponding Arabic words. Furthermore, they improved the lexicon by adding possible synonyms of the 300 translated words. They also studied the comparison between four machine learning classifiers namely SVM, KNN, NB, and decision tree for Arabic SA with the combination of three stemming situations such as: light stemming, root stemming, or no stemming. On the other hand, they also conducted the experiment using the constructed lexicon on the collected Arabic twitter dataset. They reported that the accuracy of corpus-based approach outperformed the lexicon-based approach. In addition, the best accuracy of corpus-based approach resulted using light stemming with SVM classifier with accuracy of 87.2%, while the best accuracy reached in lexicon-based approach was 59.6%.

In research by [42], they created a new Arabic SA dataset called Arabic Jordanian General Tweets (AJGT) Corpus. They compared the performance of SVM and NB classifiers using different preprocessing combinations, which was applied on the AJGT dataset. The new dataset was collected from Twitter about Jordanian different topics with a total of 1800 tweets. They also investigated and compared the performance of the two classifiers using different preprocessing techniques. In addition, they compared three N-grams feature extraction techniques namely Bigrams, Unigrams, and Trigrams. Furthermore, they combined the N-grams feature extraction techniques with either TF or TFIDF weighting technique. Besides, they investigated the use of three different stemming configuration techniques such as root stemming, light stemming, or no stemming. They reported that the SVM classifier combined with TFIDF weighting, Bigram feature extraction, and feature stemming outperformed other scenarios combination and reached an accuracy of 88.72% and F-measure 88.27%. In work of [43], they created a new Arabic dataset for SA and compared the performance between SVM and NB classifiers using the dataset. The created data was collected from Yahoo-Maktoob about several Arabic topics such as science, arts, social, politics, and technology. They also applied stemming as a feature selection with TFIDF weighting technique with each classifier. The accuracy reported from applying SVM classifier outperformed the NB classifier with accuracy of 64.1%. Furthermore, in [44], they conducted several comparative experiments for Arabic SA using three machine learning classifiers, KNN, SVM, and NB using the OCA dataset [38]. They combined each one of the used classifiers with one feature selection technique from seven different features selection techniques such as:

Table 1 Comparative summary of the related works in Arabic SA

Ref	Method Used	Classifier Used	Dataset Used	Best Performance Achieved
[34]	GA with classifiers.	SVM, SGD, MNB, DT, and NB.	LABR.	MNB classifier using unigram with 85% accuracy.
[35]	TF, BM, or TFIDF with classifiers.	SVM, KNN, and NB.	Arabic Twitter dataset.	NB classifier with TF with 69.97% accuracy.
[36]	Four reduct feature selection techniques.	N/A	Twitter Arabic dataset.	Best accuracy 74% by using of FRAW with LEM2 algorithm.
[37]	GI, IG, or CHI with classifiers.	N-gram model, ARM, and Meta classifier.	OCA.	CHI feature selection with Meta classifier with 90.8% accuracy.
[39]	Improved CHI or original CHI with classifier.	DT	Arabic newspapers reviews.	The combinations of light stemming, improved CHI, and DT classifier best recall with 75.3%.
[40]	TFIDF weighting technique followed by stemming with classifiers.	SVM, and NB	Arabic dataset collected from Yahoo-Maktoob.	SVM classifier with 68.2% accuracy.
[41]	Light stemming, root stemming, or no stemming with classifiers.	SVM, KNN, NB, and DT.	Arabic tweets from Twitter	Light stemming with SVM with accuracy 87.2%.
[42]	N-grams feature extraction with either TF or TFIDF with three different stemming configurations with the classifiers.	SVM and NB.	AJGT	SVM classifier combined with TFIDF weighting, Bigram feature extraction, and feature stemming with accuracy of 88.72%.
[43]	Stemming as feature selection with TFIDF weighting technique with each classifier.	SVM and NB.	Arabic reviews from Yahoo-Maktoob.	SVM with 64.1% accuracy.
[44]	CHI, PCA, GI, Relief, uncertainty, SVM, and IG. Also, they applied root stemming with the classifiers.	KNN, SVM, and NB.	OCA	SVM classifier with SVM as a feature selection with 92.4% accuracy.

CHI, Principal Components Analysis (PCA), GI, Relief, Uncertainty, SVM, and IG. They also applied root stemming on the selected features and removed the stop words. Based on their experiments, they concluded that the combination of SVM classifier with SVM as a feature selection was outperformed other combinations, which achieved 92.4% accuracy. Table 1 summarizes the comparison of these previous works, which highlighted the method, together with the performance of the classifier, and data used by each work.

Based on literature investigation of the mentioned works, little works utilized optimization algorithms to take benefits of their advantages for finding the optimal solution and reduce the space complexity. In literature, many research works have been proposed for minimizing the number of features used for training the SVM classifier. For example, in [45] work they used GSA algorithm with mutual information for feature selection. In [46] work, they used improved Firefly Algorithm (FA) with mutual information for feature selection. Further, in research by [47], they improved the Bacterial Algorithm (BA) by using new control and update techniques, then applied the improved BA algorithm on feature selection. In [48], they used binary quantum-inspired GSA for feature selection. Moreover, in work by [49] they developed multi objective version of ABC algorithm by hybridized it with sorting technique and applied it to feature selection problem. Although many researches works were conducted for feature selection problem but based on "No Free Lunch theorem in optimization" [50], which states that there is no single optimization algorithm capable of solving all optimization problems and no one single algorithm superior to all other optimization algorithms in solving all optimization problems. Therefore, one algorithm can outperform some algorithms in some problems but not all problems. These reasons motivated us to further do more contribution in this research area. The WOA algorithm [5], which is one of the newly optimization algorithms and shows good results according to previously mentioned. However, there is a possibility to improve the standard algorithm performance. This reason motivates us to improve the algorithm and apply it in feature selection for Arabic SA.

3 Preliminaries

3.1 Whale optimization algorithm

WOA is one of the newly metaheuristic algorithms, which proved its capability in balancing between the exploration and exploitation compared to the state-of-the-art optimization algorithms. The idea of WOA is originated from mathematically formulating the hunting behavior used by humpback whale. Moreover, the humpback whales use the bubble-net hunting technique to encircle and catches their preys, which is also considered as the core intelligent mechanism used by the algorithm. In addition, whales can communicate with each other, learn, and normally live in groups [5]. The hunting mechanism used by humpback whales is based on hunting groups of small fishes that are close to the surface. The whales go down the surface and dive about 12 m below the preys, then they start creating bubbles in '9' or circles shapes to encircle the small fishes inside the created bubbles and make them as traps for these fishes. Consequently, the whales begin to go upward to the surface to hunt these small fishes [5]. WOA algorithm is mathematically composed of three phases including 1) encircling prey phase, 2) exploitation (bubble-net attacking) phase, and 3) exploration (search for a prey) phase, which formulate the spiral bubble hunting mechanism used by the whales [5]:

1) Encircling prey phase: In this phase, the whales (search-agents) identify the locations of prey, then they encircle them. After that, WOA at this stage finds the best candidate search-agents from a set of randomly generated set of search agents. Moreover, other whales (search-agents) try to update their positions with reference to this initially specified best candidate so far. The whole process is formulated using the following Eqs. (1) and (2) [5]:

$$\vec{D} = \left| \vec{C} \cdot \vec{X}^*(t) - \vec{X}(t) \right| \tag{1}$$

$$\vec{X}(t+1) = \vec{X}^*(t) - \vec{A} \cdot \vec{D} \tag{2}$$

where $\vec{A}$ and $\vec{C}$ are representing coefficient vectors, t stores the current iteration, $\vec{X}^*$ vector indicates the best solution position vector as obtained until this phase, and $\vec{X}$ represents the position vector. The values of $\vec{A}$ and $\vec{C}$ coefficient vectors will be calculated using Eqs. (3) and (4) [5]:

$$\vec{A} = 2\vec{a} \cdot \vec{r} - \vec{a} \tag{3}$$

$$\vec{C} = 2 \cdot \vec{r} \tag{4}$$

where the value of the variable $\vec{a}$ is decreased from 2 to 0 linearly over the algorithm iterations. Moreover, the vector $\vec{r}$ represents a random value over 0 to 1.

2) Exploitation (bubble-net attacking) phase: This exploitation phase is based on two techniques namely 1) shrinking encircling technique, and 2) spiral updating position.

- Shrinking encircling technique: they obtain and formulate the shrinking encircling behavior by decreasing $\vec{a}$ variable value from 2 to 0 linearly that is used in Equation (3) over the algorithm iterations, which in turn means the value of $\vec{A}$ vector is a random value over [−a, a].
- Spiral updating position: to formulate this behavior, they determined the distance between the current whale position and the prey. After finding the distance, to mimic the whale's spiral movement from its current position to the prey position they created a spiral equation as shown in Equation (5) [5]:

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

where $\vec{D'} = \left| \vec{X}^*(t) - \vec{X}(t) \right|$ represents the distance between the current whale and the prey, b represents a constant used to define the spiral movement shape by the whales, and/is a random value over the interval [−1,1]. In modeling the behavior of whale's movements around the prey, it is noticed that the humpback whales move concurrently using shrinking circling technique and move in spiral path toward the prey. Moreover, the probability of whales switching between the two behaviors is 50% and it is modeled using equation (6) [5]:

$$\vec{X}(t+1) = \begin{cases} \vec{X}^*(t) - \vec{A} \cdot \vec{D} & if\ p < 0.5 \\ \vec{D'} \cdot e^{bl} \cdot cos(2\pi l) + \overrightarrow{X^*}(t) & if\ p \geq 0.5 \end{cases} \tag{6}$$

where p is a random number over the interval [0,1].

3) Exploration (search for a prey) phase: In this phase the humpback whales conduct a global search (exploration). Like the exploitation phase, the exploration phase is based on $\vec{A}$ vector value, whereby if the $\vec{A}$ value is greater or equal to 1, it will be an exploration, else it will be an exploitation. In the exploration phase, the whales update their position with reference to the randomly selected whale instead of updating it to the best whale to make global search. Therefore, the whale new position will be calculated using Equations (7) and (8) [5]:

$$\vec{D} = \left| \vec{C} \cdot \overrightarrow{X_{rand}} - \vec{X} \right| \tag{7}$$

$$\vec{X}(t+1) = \overrightarrow{X_{rand}} - \vec{A} \cdot \vec{D} \tag{8}$$

where X_{rand} is a random value that represents the position of the randomly selected whale from the available whales, Fig. 2 shows the WOA algorithm.

Initialize the whales population X_i (i = 1, 2, ..., n)
Calculate the fitness of each search agent
X=the best search agent*
while *(t < maximum number of iterations)*
for *each search agent*
Update a, A, C, l, and p
if1 *(p<0.5)*
if2 *(|A| < 1)*
Update the position of the current search agent by the Eq. (1)
else if2 *(|A|≥1)*
Select a random search agent (X_{rand})
Update the position of the current search agent by the Eq. (8)
end if2
else if1 *(p≥0.5)*
Update the position of the current search by the Eq. (5)
end if1
end for
Check if any search agent goes beyond the search space and amend it
Calculate the fitness of each search agent
Update X if there is a better solution*
t=t+1
end while
*return X**

Fig. 2 WOA algorithm [5]

3.2 Elite opposition-based learning

- Opposition-Based Learning (OBL): The OBL is an optimization strategy used to improve the diversity of optimization algorithm and enhance its generated solutions. In common, optimization algorithms start its steps toward optimal solution by initially generating a set of random solutions. However, these generated solutions normally not based on previous knowledge and just generated randomly along the problem search space. Moreover, most of optimization algorithms when it updates the position of the search agent it based on distance toward the current best solution, but this is not guaranteeing to reach global optimal solution. To solve the optimization algorithm problem the OBL can be used. The OBL strategy gives a useful concurrent search in both direction, which includes the current solution and its opposite solution, then take the best one based on its fitness for further processing [51].
- Opposite number: according to strategy proposed by [51], which states that if x is real number over interval $x \in [lb, ub]$ where lb is the lower bound value of variables in the current j dimension, while ub is the upper bound value of variables in the current j dimension, the opposite number of x is $\tilde{x}$ and its value can be determined according to equation (9) [51]:

$$\tilde{x} = lb + ub - x \tag{9}$$

For example, in feature selection problem $lb = 0$ and $ub = 1$ where the value of 1 means the given feature is selected; otherwise value of 0 means the feature not selected. The same equation can be generalized and applied in multidimensional search space, in this case the search agent solution can be represented as the following equations (10) and (11):

$$x = [x_1, x_2, x_3, \ldots, x_n] \tag{10}$$

$$\tilde{x} = \left[\tilde{x}_1, \tilde{x}_2, \tilde{x}_3, \ldots, \tilde{x}_n\right] \tag{11}$$

Where equation (10) represents the dimensions of the current solution and equation (11) represents opposite solution dimensions of the current solution. Now for each element in $\tilde{x}$ its value will be determined according equation (12):

$$\tilde{x}_j = lb_j + ub_j - x_j \quad \text{where } j = 1, 2, 3, \ldots, \text{n} \tag{12}$$

- Optimization Based on Opposition: In this technique, the current candidate solution x is replaced with its corresponding opposite solution $\tilde{x}$ according to its fitness value. Let the fitness function is $f(.)$, then in every iteration the fitness value of each solution in the search space and its corresponding opposite will be calculated, then the fittest solutions will be selected from the original and opposite solutions set. For example, if fitness value $f(\tilde{x})$ is better than fitness value $f(x)$ of its corresponding solution x then $x = \tilde{x}$, otherwise $x = x$ and the new value will go for next iterations in the optimization algorithm [51].
- Elite Opposition-based Learning (EOBL): In work of [11], they improved the basic OBL technique and make EOBL as a new opposition-based learning strategy that based on finding the elite solution from the current population. In EOBL for each solution its opposite solution will be determined based on predetermined elite solution. Furthermore, as in the initialization phase in all optimization algorithms the solutions are randomly generated. The optimization algorithm evolves over the iterations toward the global solution and there is a possibility that many of initial solutions are very far from the optimal solution. In addition, worst case happens when there are many initial solutions are in opposite locations with the optimal solution. To solve this problem EOBL can be used, which able to search in both directions of the original initial solution and the new generated opposite solutions and take the fittest solutions for next iterations. Further, EOBL was used in many studies such as [23–25], and according to their experimental results it shown that EOBL achieved more interesting and efficient results than the basic OBL. Motivated by these finding, the idea of EOPL is incorporated in WOA initialization to improve its diversity and search capability. The basic idea of EOBL was formulated according to the following equations [11], assume the elite solution is $x_e = [x_{e1}, x_{e2}, x_{e3}, \ldots, x_{e\ dim}]$ which is determined from the initial generated n solutions as the fittest solution among these n solutions, now for every solution x_i its elite opposition solution $\tilde{x}_i$ can be determined using

the following equation (13)

$$\tilde{x}_{ij} = k(lb_j + ub_j) - x_{ej} \quad \text{where } j = 1, 2, 3, \ldots, \text{dim and } i = 1, 2, 3, \ldots, \text{n} \tag{13}$$

Where k in this case is a random value over the interval $[0, 1]$, the upper ub_j and low value lb_j used in elite opposite solutions calculation and the formula for finding its values are (14–15):

$$lb_j = \min(x_{i,j}) \tag{14}$$

$$ub_j = \max(x_{i,j}) \tag{15}$$

Now to ensure that the new opposite values are feasible and inside the boundary of the search space the following equation (16) will be used after finding $\tilde{x}_{i,j}$

$$\tilde{x}_{i,j} = rand(lb_j, ub_j) \left[if\ \tilde{x}_{i,j} < x_{min} OR\ \tilde{x}_{i,j} > x_{max} \right] \quad i = 1, 2, 3, \ldots, \text{n}, \quad j = 1, 2, 3, \ldots, \text{dim} \tag{16}$$

Where $x_{i,\ j}$ represents the j^{th} value in the vector of current i^{th} solution of problem population, $\tilde{x}_{i,j}$ is the elite opposite solution of $x_{i,\ j}$, lb_j represents the minimum value of the j^{th} dimension in the search space, ub_j represents the maximum value of the j^{th} dimension in the search space, $rand(lb_j, ub_j)$ is a random value over the interval $[lb_j, ub_j]$, the maximum and minimum bounds of $\tilde{x}_{i,j}$ is $[x_{min}, x_{max}]$ which is the constraints if the new value of $\tilde{x}_{i,j}$ is jump out of the boundary, dim is the problem dimension, and n is the size of population. Thus, EOBL can be embedded in the initialization phase to get fitter solutions than the initially generated solutions.

3.3 Information gain

In Information Gain (IG), the relevance of each feature to the class labels is estimated according to following equation (17), whereby IG ranks each feature according to its entropy and selects the most important features according to the prespecified threshold [52].

$$\begin{aligned} IG(f) = &- \sum_{n=1}^{m} P(C_n) \log P(C_n) \\ &+ P(f) \sum_{n=1}^{m} P(C_n|f) \log P(C_n|f) \\ &+ P(\overline{f}) \sum_{n=1}^{m} (C_n|\overline{f}) log\ P(C_n|\overline{f}) \end{aligned} \tag{17}$$

where C_n stand for the n^{th} class category, f is the feature, $P(C_n)$ is the percentage portion of reviews with C_n class category, $P(f)$ is the percentage portion of reviews in which the feature f exists, $P(\overline{f})$ is the percentage portion of reviews in which the feature f does not exist, $P(C_n|f)$ is the percentage portion of reviews from class category C_n that have the feature f, and $P(C_n|\overline{f})$ is the percentage portion of reviews from class category C_n that does not have the feature f [52].

3.4 Differential evolution(DE)

DE algorithm which is proposed by [26], is a type of evolutionary search algorithms that mimic biological operations by applying three inspired biological operators. The main DE evolutionary operators include mutation, crossover, and selection.

3.4.1 Mutation

Mutation operator involves creating three indices randomly in range over [1, n] where n is the population size, then from the current solutions in the search space three solution vectors with the given indices will be selected. Based on the three randomly selected solutions a new solution will be generated using equation (18):

$$V_i = X_{r1} + F(X_{r2} - X_{r3}) \tag{18}$$

Where X_{r1}, X_{r2}, and X_{r3} are the solution vectors that are randomly selected and F is the mutation scaling factor and its value over [0,1]. V_i is the newly generated mutant solution vector.

3.4.2 Crossover

Crossover operator in DE involves making crossover between the new mutant solution vector V_i and the original solution vector X_i according to the following equation (19):

$$U_{ij} = \begin{cases} V_{ij} & if\ rand(0,1) \leq CR \quad or \quad j = j_{rand} \\ X_{ij} & otherwise \end{cases} \tag{19}$$

Where U_{ij} is new solution vector resulted from crossover operator and this trial vector represented by $U_{ij} = \{U_{i1}, U_{i2}, U_{i3}, \ldots U_{i\,dim}\}$, CR is the crossover rate, and j_{rand} is random value over $[1, dim]$ which represent the randomly chosen index in the solution vector. Dim is the dimensionality of the problem.

3.4.3 Selection

Selection operator in DE involves makes selection between the original solution X_i or the new trial solution U_i based on their fitness value by selecting the fittest solution among the two solutions for next iteration according to following equation (20):

$$X_i^{t+1} = \begin{cases} U_i & \text{if } f(U_i) > f(X_i) \\ X_i & \text{otherwise} \end{cases} \tag{20}$$

In this selection strategy, the current candidate solution X_i is replaced with U_i or remain as it based on their fitness value. Let the fitness function is $f(.)$, If fitness value $f(U_i)$ is better than fitness value $f(X_i)$ then $X_i^{t+1} = U_i$ otherwise $X_i^{t+1} = X_i$.

4 Proposed algorithm

This section presents our proposed improved algorithm IWOA and all required details for Arabic language SA. Figure 3 shows an architecture of the proposed improved algorithm, which consists of two major phases: Filter Phase and Wrapper Phase.

4.1 Input dataset phase

To evaluate our proposed improved IWOA algorithm, four publicly available Arabic datasets for sentiment analysis were used. The first used dataset is called OCA [38] which was collected from different number of Arabic movie blogs and web pages about movie reviews written in Arabic language. OCA corpus includes a total of 500 Arabic reviews, which are organized into two groups, whereby the first group contains 250 positive reviews, and the second group contains 250 negative reviews. In addition, the collected reviews were manually preprocessed by removing the HTML tags and special symbol characters, correcting the misspelled words, and annotating them [38]. The statistics of the OCA dataset is shown in Table 2.

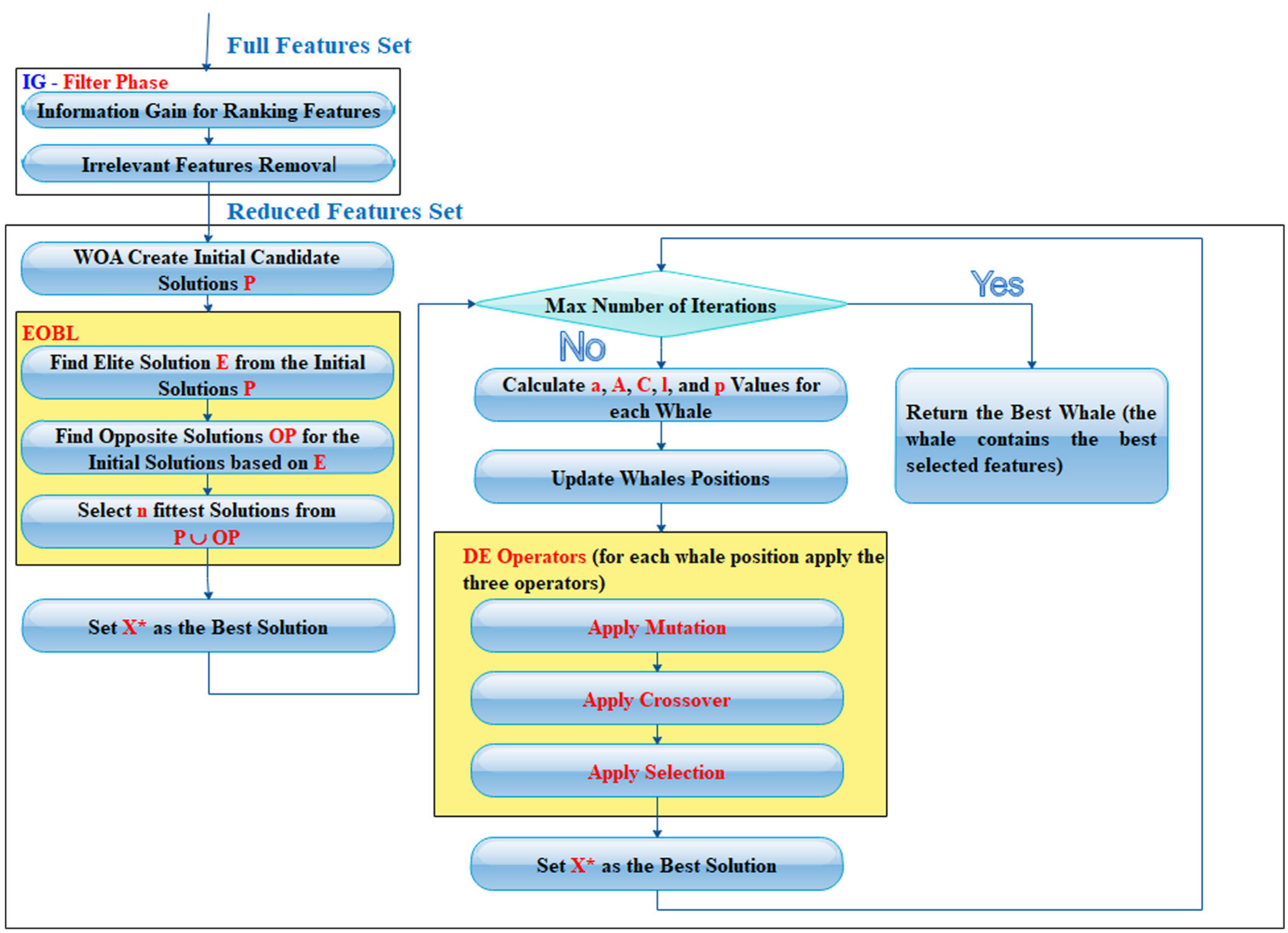

Fig. 3 Proposed architecture of IWOA

Table 2 Statistics of OCA dataset

Review Polarity Criteria	Positive Reviews	Negative Reviews
Total number of review documents	250	250
Total number of words	121,392	94,556
Average number of words in each file	485	378
Total number of sentences	3137	4881
Average number of sentences in each file	13	20

The second used dataset from [53], which was collected from twitter on different topics such as arts and politics. This Arabic twitter corpus contains 2000 tweets reviews with 1000 positive tweets and 1000 negative tweets. These collected tweets were written using both Modern Standard Arabic (MSA) and the Jordanian dialect [53]. The corpus collected tweets were annotated manually by two domain experts. If both annotators agree in annotation label of the given tweet, then this annotation is approved. Otherwise, they asked a third domain expert to take the decision of annotation. The statistics of Arabic twitter dataset [53] is shown in Table 3. In addition, the collected tweets were preprocessed by removing the repeated letters, correct misspellings words, and normalization of Arabic letters [53].

The third and fourth used datasets are software and political datasets from [54], the reviews were collected manually by authors from number of websites such as Qaym website, Jeeran website, Google play, Facebook, and Twitter. Furthermore, the annotation of datasets was conducted manually by two Arabic native speakers into either negative or positive reviews. In addition, as a preprocessing of the collected reviews they removed non-Arabic words. The statistics of political and software datasets are shown in Table 4 [54].

4.2 Preprocessing phase

The preprocessing phase includes very crucial processes that help in producing a ready to use dataset for training and testing purposes. In our proposed methodology, this phase includes: 1) tokenization, 2) stop words removal, 3) stemming, and 4) term weighting.

During the tokenization process, all individual words (tokens) in each review are identified by splitting each review into individual tokens. In the stop words removal process, words that do not carry important meaning for the text and classification process are removed. The stop words such as "في,In", "على,On", and "من,From" normally occur with high frequencies within the review text. Furthermore, it is an essential step to remove these words as they affect the performance of the classifier. In the stemming process, the root or stem of Arabic words are extracted by removing all suffixes and prefixes from each token. Furthermore, the tokens that share the same root or stem are placed together as they all have related meanings. For example, the following three Arabic words (Office, "مكتب"), (Library, "مكتبة"), (Writing, "كتابة") are created from the same Arabic root word (Write, "كتب"). Once the Arabic words are preprocessed according to the previous steps, each review will be transformed into a vector, whereby each vector includes the left terms after preprocessing. The vector is then used to calculate the weight of the included terms using TF-IDF. The following equations (1–3) are used to calculate the TF-IDF weight:

$$TF(i,j) = \frac{\text{Frequency of term } i \text{ in review } j}{\text{Total number of terms in review } j} \tag{21}$$

$$IDF(i,j) = log\frac{Total\ number\ of\ reviews\ in\ the\ dataset}{Number\ of\ reviews\ which\ include\ i\ term} \tag{22}$$

$$W_{i,j} = TF_{i,j} \times IDF_{i,\mathrm{j}} \tag{23}$$

where TF is calculated first using Equation (21), which refers to the frequency of term i in review j, and the IDF is then calculated using Equation (22). After calculating TF and IDF for each term in the review, the term weight is calculated using Equation (23).

Table 3 Statistics of Arabic twitter dataset

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

Table 4 Statistics of political and software datasets

Review Polarity Criteria	Positive Reviews	Negative Reviews
Total number of tweets in Political	600	600
Total number of tweets in Software	600	600

4.3 Filter phase

In this phase, IG feature ranking technique used to remove irrelevant features. Further, IG used to reduce the features size and select the top ranked features for training the SVM classifier. Thus, select best features will give superior classification performance than using the whole set of original features.

4.4 Wrapper phase (improved WOA based on DE operators and EOBL)

This section introduces the improvement to the standard WOA, in which the EOBL is embedded inside the standard WOA initialization to improve its solutions. In addition, DE mutation, crossover, and selection operators incorporated at the end of each WOA iteration.

As shown in Fig. 3, the improvement to standard WOA contains two phases. The first improvement as displayed in Fig. 4, after the WOA algorithm generates its initial whales' solutions, then EOBL will be applied to find elite opposite solution of each initial solution using equation (13) in EOBL. Next, to ensure that the new elite opposite solutions not fall out of the boundary, it applied equation (16) from EOBL. After that, from the whole set of initial solutions and set of elites opposite solutions EOBL select the best solutions based on their fitness. For example, if the number of initial solutions is 30, then EOBL find 30 elite opposite solutions for the set of initial solutions. Finally, the whole set now contains 60 solutions, from these 60 solutions EOBL select best 30 solutions based on their fitness values.

Another improvement at the end of each iteration after WOA updates the whales' positions, DE evolutionary operators will be applied on each whale position to find better positions and improve WOA local search capability by using equations (18–20). During the wrapper phase, the improved IWOA algorithm is used for feature selection. The improved IWOA algorithm takes the reduced features, which are resulted from applying IG to select the optimal feature combinations from the reduced features. Furthermore, the optimal selected features are used as input to the classifier for Arabic sentiment classification. Therefore, the improved IWOA algorithm is used right after IG to select the optimal features in order to improve the sentiment classification performance and to reduce the number of used features simultaneously.

Fig. 4 Proposed Improved WOA algorithm based on DE evolutionary operators and EOBL

Initialize the whales population X_i (i = 1, 2, ..., n)
// **Apply GOBL In Initialization**

1. *Find elite solution* $\mathbf{X_e}$ *from initial population*
2. *Apply EOBL to calculate the elite opposition positions of all initial whales by Eq. (13).*
3. *Check if any new opposite search agent goes beyond the search space reposition it by Eq (16)*
4. *Calculate the fitness of each search agent in X_i and its corresponding elite opposite whale $\tilde{X}_i$*
5. *Select the n fittest search agents from the set of initial solutions and elite opposite solutions set.*

X=the best search agent*
while *(t < maximum number of iterations)*
  for *each search agent*
  Update a, A, C, l, and p
    if1 *(p<0.5)*
      if2 *(|A| < 1)*
        Update the position of the current search agent by the Eq. (1)
      else if2 *(|A|≥1)*
        Select a random search agent (X_{rand})
        Update the position of the current search agent by the Eq. (8)
      end if2
    else if1 *(p≥0.5)*
      Update the position of the current search by the Eq. (5)
    end if1
  end for
  Check if any search agent goes beyond the search space and amend it
  for *each search agent*
    apply DE mutation
    apply DE crossover
    apply DE selection
  end for
  Update X if there is a better solution*
  t=t+1
end while

Table 5 Parameters setting of the used algorithms

Algorithm	Parameter
DE	Crossover ratio = 0.1 beta_min = 0.3 beta_max = 0.7
PSO	Acceleration constants C1 = 1.5 C2 = 2 Inertia Weight W1 = 1 W2 = 0.9
GA	Crossover ratio = 0.9 Mutation ratio = 0.1
GOA	cMax = 1 cMin = 0.00004
WOA	a = [2,0] b = 1
IWOA	a = [2,0] b = 1 For mutation, crossover, and selection: Crossover ratio = 0.1 beta_min = 0.3 beta_max = 0.7

IWOA for feature selection The proposed improved algorithm for feature selection technique in our study is accomplished through two phases as shown in Fig.3. In the first phase, we used filter feature selection approach by calculating the IG weight of each feature, therefore the relevant features will be identified. In the second phase, we applied IWOA during the wrapper mode in the previously selected relevant features by IG. IWOA is applied to find the optimal features combinations from these features resulted from applying IG filter approach. Finally, this proposed hybrid model takes the advantages of both modes by taking the advantage of efficiency resulted by IG filter mode and the accuracy advantage resulted from applying IWOA in wrapper mode. In IWOA, the selection of features is represented as a binary value, such that the value of "1" means the corresponding feature is selected; otherwise the value of "0" means the corresponding feature is not selected. The following steps list and describe our proposed IWOA algorithm:

1. Feature ranking and pruning: In the first step, IG filter technique is used to rank features and select the best subset discriminative features from the full set of original features vector. Furthermore, this step is mainly needed to minimize the number of features, to avoid IWOA of searching irrelevant space areas by minimizing the search space, to minimize the searching time used by IWOA algorithm, and to improve the SA classification performance.
2. WOA initialization: In the second step, only the reduced features are selected by IG filter technique and then used as input to IWOA. In addition, the IWOA randomly generates a number of whales (search-agents) according to the prespecified number of whales. Furthermore, each generated search-agent represents a possible solution, which contains randomly selected subset of features from the full set of features generated in step 1 by IG.
3. Apply EOBL: In this step, EOBL used to find the elite solution from step 2, then based on this elite solution it finds the elite opposite position of all solutions generated in step 2 by using equation (13) and if any new opposite solution goes out the boundary, then reposition it by using equation (16). Finally, from the set of initial solution in step 2 and from a set of elites opposite solutions generated in step 3 EOBL take the best n solutions based on their fitness. In Whales fitness evaluation, IWOA evaluates the fitness value for each whale (possible solution) according to SVM classification accuracy. In addition, IWOA feature selection works by selecting the features in the wrapper phase. Moreover, the training dataset is used to train the SVM classifier, while testing the dataset is used by IWOA to calculate the accuracy of SVM sentiment classification. Further from these selected solutions, IWOA will select the best solution X*
4. Whales positions updates: In this step, IWOA updates the positions of search-agents either with reference to the best search-agents obtained so far or with reference to the randomly selected search-agents positions.
5. Apply DE evolutionary operators: for each whale position the DE evolutionary operators including mutation, crossover, and selection will be applied to find better positions.
6. IWOA termination: The IWOA algorithm will repeat steps 4 and 5 *t* times (*t* represents the number of iterations that are predefined for IWOA to iterate over solutions) and update the better solution at the end of each iteration if it finds a solution better than the current best solution.

Table 6 The sentiment classification accuracy from using each classifier

Dataset	No. of Features	Proportion of Features Used (%)	SVM (%)	KNN (%)	NB (%)
OCA [38]	8057	100	70.16	66.33	53.95
Arabic twitter [53]	2257	100	80.23	61.46	72.17
Political [54]	2237	100	74.74	51.25	66.29
Software [54]	2378	100	82.22	74.49	65.10

Table 7 classification accuracy applying IG feature selection with SVM classifier

Dataset	Number of input features based on IG ratio	IG ratio (%)	Average accuracy (%)
OCA [38]	483	6	84.60
OCA [38]	966	12	87.43
OCA [38]	1450	18	**88.86**
Arabic twitter [53]	135	6	**83.98**
Arabic twitter [53]	271	12	81.33
Arabic twitter [53]	406	18	83.43
Political [54]	134	6	**81.14**
Political [54]	268	12	79.49
Political [54]	403	18	78.07
Software [54]	143	6	**84.88**
Software [54]	285	12	80.31
Software [54]	428	18	79.07

7. Solution: The best whale (search-agent) with best sentiment classification accuracy using the SVM classifier will be returned by IWOA. The contained features in the best whale represent the optimal features for training the SVM classifier.
8. Testing: the best selected features will be used to test the testing data part.

5 Experimental results and analysis

To evaluate and investigate the performance and effectiveness of the proposed IWOA algorithm, all experiments were performed on four Arabic sentiment analysis datasets including OCA [38], Arabic twitter [53], Political [54], and Software [54]. The experiments implemented using RapidMiner and MATLAB software tools. During the classification process, we employed 10-fold cross-validation for all experiments by dividing the dataset in each run into 10-folds, whereby one-fold is used for testing purposes and the remaining 9 folds are used for training purposes. This process is repeated 10 times. Lastly, the average accuracy, average number of selected features, and average fitness across 10 runs are reported. The IWOA algorithm is compared with other well-known algorithms including PSO, GA, GOA [55], DE, ALO [56], and standard WOA. In addition, IWOA is compared with CNN and LSTM deep learning algorithms. The parameter setting of the used algorithms is shown in Table 5. Where the number of search agents in all algorithms is 10 and the number of iterations is 40. In the conducted experiments, classification accuracy is the main measure used for testing all algorithms performance, where classification accuracy is defined as the percentage of correct classified instances according to the actual correct classes.

Furthermore, in this work SVM classifier is applied to four different experiments as presented in the (Tables 6, 7, 8, 9, 10, 11, 12 and 13) in this section, whereby each experiment is represented by different combination of SVM as the main classifier used with IG as filter feature selection techniques and optimization algorithms (WOA, DE, IWOA, PSO, GA, ALO, and GOA). In first experiments, SVM only is applied on the whole features set without filter feature selection technique and optimization algorithm and is compared to KNN and NB classifiers. The second experiment involves applying IG feature selection with SVM classifier. The third experiment involves applying optimization algorithms including (WOA, DE, IWOA, PSO, GA, ALO, and GOA) with SVM on the whole dataset without using IG feature selection. In the fourth experiment its involves applying optimization algorithms including (WOA, DE, IWOA, PSO, GA, ALO, and GOA) and IG feature selection with SVM classifier on the used datasets. In the fifth experiment, the IOWA and IWOA+IG are compared to deep learning algorithms include CNN and LSTM.

Table 8 Comparison between IWOA and other algorithms in terms of average classification accuracy in 10 runs without using IG feature selection

Dataset	Number of input features.	Algorithms average accuracy (%)						
		WOA	DE	IWOA	GOA	PSO	ALO	GA
OCA [38]	8057	93.30	95.33	**95.93**	91.28	94.92	89.64	95.54
Arabic twitter [53]	2257	83.05	85.80	**87.61**	76.21	81.73	76.21	83.09
Political [54]	2237	86.04	86.96	**88.45**	80.64	87.70	78.90	86.86
Software [54]	2378	89.45	91.27	**92.36**	86.46	91.53	84.55	92.02

Table 9 Comparison between IWOA and other algorithms in terms of average number of selected features in 10 runs without using IG feature selection

Dataset	Number of input features.	Average number of selected features						
		WOA	DE	IWOA	GOA	PSO	ALO	GA
OCA [38]	8057	3910	4004	3157	4014	3996	4040	3886
Arabic twitter [53]	2257	1227	1082	1102	1127	1107	1125	**1033**
Political [54]	2237	1650	1109	1489	1122	1114	1119	**1068**
Software [54]	2378	1648	1177	1317	1192	1188	1188	**1158**

The accuracy measure is considered as the main measure in our experiments to compare between the performance of different methods and combinations, whereby the accuracy required in SA as our target is to minimize the classification error rates as we have either positive or negative reviews only. The experimental results of all experiments are presented in Tables 6 through 14.

5.1 Experiment 1: Results using different types of classifiers

The first experiment is represented by evaluating the performance of the proposed method using either SVM, KNN, or NB classifier, on which each classifier is applied only on the whole features set without filter feature selection technique and optimization algorithm. Table 6 shows the sentiment classification accuracy results from the first experiment. Based on Table 6, the best achieved accuracy obtained from using SVM classifier which was 70.16% on OCA dataset, 80.23% on Arabic twitter dataset, 74.74% on political dataset, and 82.22% on software dataset. Based on the obtained results it is confirmed that the SVM classifier outperforms other machine learning classifiers including KNN and NB. In addition, based on results of the previous works [40–44, 57–64] its demonstrated the outperformance of results obtained using SVM in comparison with other classifiers over Arabic classification. Thus, based on these findings and promising results obtained from SVM, the SVM is selected in this research.

The obtained results on Table 6 show a necessity for employing filter feature selection techniques and optimization algorithms to obtain better accuracy results.

5.2 Experiment 2: Results from applying SVM classifier with IG feature selection

In this experiment, IG filter feature selection technique is applied to rank all extracted features. It is important to note that in the second experiment, the top features are selected from the whole features set according to their resulted IG weights based on 6%, 12%, and 18% ratios. Tables 7 show the results after applying IG feature selection according to different ratios using SVM classifier with best results in bold.

Based on results from Table 7, the highest obtained accuracy on OCA dataset was 88.86%, which is achieved when the top 18% of the ranked features using IG are selected and used. In Arabic Twitter dataset the highest accuracy achieved 83.98% when the top 6% of ranked features by IG are selected. Furthermore, the highest accuracy achieved using political dataset was 81.14% when top 6% features ranked by IG were selected. Finally, the highest accuracy achieved on software dataset was 84.88% when the top 6% IG ranked features were selected. In summary, the results obtained from combination of SVM with IG feature selection in second experiment according to Table 7 are outperforming the results from experiment 1 in Table 6 resulting from applying SVM only on the whole set of features. In addition, the application IG on the datasets minimizes the number of selected features. This worth noting that feature reduction techniques improve sentiment classification accuracy results based on Table 7 results.

5.3 Experiment 3: IWOA comparisons with other metaheuristics algorithms

To examine the effectiveness and performance of the proposed IWOA algorithm, we applied optimization algorithms

Table 10 Comparison between IWOA and other algorithms in terms of average fitness in 10 runs without using IG feature selection

Dataset	Number of input features.	Algorithm average fitness (%)						
		WOA	DE	IWOA	GOA	PSO	ALO	GA
OCA [38]	8057	7.13	5.11	**4.39**	9.12	5.50	10.74	4.89
Arabic twitter [53]	2257	17.39	14.52	**12.76**	24.04	18.54	24.04	17.19
Political [54]	2237	14.60	13.40	**12.16**	19.66	12.65	21.38	13.47
Software [54]	2378	11.21	9.12	**8.17**	13.89	8.86	15.79	8.38

Table 11 Comparison between IWOA and other algorithms in terms of average classification accuracy in 10 runs using IG feature selection

Dataset	Number of input features based on IG ratio	IG ratio	Algorithms average accuracy (%)						
			WOA	DE	IWOA	GOA	PSO	ALO	GA
OCA [38]	483	6%	97.56	97.97	98.78	92.49	97.77	92.28	97.96
OCA [38]	966	12%	97.15	98.78	99.18	96.34	98.78	93.91	98.58
OCA [38]	1450	18%	97.56	98.98	99.39	96.14	99.18	94.11	99.19
Arabic twitter [53]	135	6%	85.08	84.63	85.93	79.17	83.88	78.67	84.28
Arabic twitter [53]	271	12%	87.53	87.18	88.83	82.13	87.73	81.48	87.58
Arabic twitter [53]	406	18%	88.89	88.63	89.68	83.18	88.53	8233	88.49
Political [54]	134	6%	82.14	82.96	84.63	77.40	82.39	76.32	83.13
Political [54]	268	12%	85.63	85.05	87.37	79.16	85.39	78.74	84.97
Political [54]	403	18%	87.79	87.54	88.54	81.56	87.38	79.24	87.46
Software [54]	143	6%	89.53	89.78	91.94	85.05	90.36	84.71	90.19
Software [54]	285	12%	90.95	91.28	92.44	86.54	92.02	85.13	91.44
Software [54]	428	18%	91.52	92.36	93.27	86.54	91.85	92.85	86.29

including (WOA, DE, IWOA, PSO, GA, ALO, and GOA) in the third experiment without employing IG filter features selection technique. In addition, SVM is used as the main classifier. Based on the third experiment results from Table 8 though Table 10, it is found that the IWOA algorithm outperform other algorithms over all datasets in terms of accuracy as indicated in bold font while it also reducing the number of selected features. In addition, IWOA outperforms all other algorithms in terms of their fitness values as IWOA resulted in the lowest values of the objective function in comparison with other algorithms on all datasets. Furthermore, the accuracy results of using IWOA on the whole set of features is superior to using only SVM or IG feature selection only in comparison with results in experiments 1 and 2. By referring to Table 9 its clearly observed that IWOA algorithm outperform the standard WOA in feature reduction in all datasets and got the best results among all algorithm using OCA dataset.

5.4 Experiment 4: IWOA with IG feature selection comparisons with other algorithms used with IG

To examine the effectiveness and performance of the proposed approach, in this experiment, IG filter features selection technique is applied with optimization algorithms including (WOA, DE, IWOA, PSO, GA, ALO, and GOA). Like all previous experiments, SVM is used as the main classifier. Based on the results from Table 11 - Table 13, it is clearly noticed that IWOA outperforms other optimization algorithms in terms of accuracy and fitness over all datasets while it also

Table 12 Comparison between IWOA and other algorithms in terms of average number of selected features in 10 runs using IG feature selection

Dataset	Number of input features based on IG ratio	IG ratio	Average number of selected features						
			WOA	DE	IWOA	GOA	PSO	ALO	GA
OCA [38]	483	6%	285	232	**179**	239	237	243	226
OCA [38]	966	12%	500	471	**385**	484	472	481	435
OCA [38]	1450	18%	767	707	**601**	720	715	721	664
Arabic twitter [53]	135	6%	104	72	94	70	**69**	71	73
Arabic twitter [53]	271	12%	206	136	179	136	138	138	**132**
Arabic twitter [53]	406	18%	335	200	257	202	204	208	**199**
Political [54]	134	6%	110	**66**	79	**66**	68	69	**66**
Political [54]	268	12%	197	**133**	168	136	134	135	**133**
Political [54]	403	18%	320	198	259	200	201	200	**187**
Software [54]	143	6%	108	**69**	92	70	72	71	70
Software [54]	285	12%	193	140	182	139	141	143	**133**
Software [54]	428	18%	329	213	267	214	208	214	**201**

Table 13 Comparison between IWOA and other algorithms in terms of average fitness in 10 runs using IG feature selection

Dataset	Number of input features based on IG ratio	IG ratio	Algorithm average fitness (%)						
			WOA	DE	IWOA	GOA	PSO	ALO	GA
OCA [38]	483	6%	3.05	2.48	1.54	7.91	2.65	8.15	2.48
OCA [38]	966	12%	3.33	1.68	1.22	4.10	1.65	6.52	1.85
OCA [38]	1450	18%	2.96	1.48	**1.02**	4.30	1.26	6.32	1.25
Arabic twitter [53]	135	6%	15.67	15.77	**14.72**	21.15	16.47	21.64	16.10
Arabic twitter [53]	271	12%	13.21	13.19	**11.78**	18.19	12.63	18.84	12.77
Arabic twitter [53]	406	18%	11.89	11.74	**10.94**	17.14	11.83	18.01	11.88
Political [54]	134	6%	18.55	17.36	**15.89**	22.87	17.91	23.96	17.19
Political [54]	268	12%	15.05	15.29	**13.18**	21.14	14.94	21.55	15.37
Political [54]	403	18%	12.94	12.81	**12.09**	18.74	12.95	21.04	12.87
Software [54]	143	6%	11.20	10.60	**8.68**	15.27	10.00	15.64	10.19
Software [54]	285	12%	9.73	9.11	**8.17**	13.79	8.35	15.22	8.93
Software [54]	428	18%	9.24	8.05	**7.34**	13.82	8.51	14.07	7.54

minimizing the number of selected features. It is obviously noticed from Table 11 the superiority of IWOA algorithm accuracy results in comparison with other algorithms as indicated by bold font. Furthermore, when comparing the results of hybridization of IWOA with IG feature selection on Table 11 it outperforms the results in Table 8 when the optimization algorithms used only on the whole features set without using IG feature selection. In addition, the selected features as shown in Table 12 are fewer than the number of features selected in Table 9.

According to results shown on Table 12, the IWOA algorithm outperform the standard WOA algorithm on all used datasets as it selected less number of features. In addition, IWOA algorithm outperforms other algorithm by selecting less number of features when OCA dataset was applied. However, in other datasets such as Arabic Twitter dataset GA algorithm outperform other algorithms when 12% or 18% IG ranked features are used, while PSO outperforms other algorithms using Arabic Twitter dataset when 6% of IG ranked features are used. Moreover, GA algorithm outperform other algorithms in features reduction when software dataset used with 12% or 18% IG ranked features are used, while DE algorithms outperforms other algorithms in features reduction when software dataset used and 6% of IG ranked features are used. Lastly, DE, GOA, and GA outperform other algorithms when 6% of IG ranked features are used from political dataset. In addition, DE and GA algorithms outperforms other algorithms in features reduction when political dataset with 12% of IG ranked features are used, while GA algorithm outperform other algorithms in features reduction when political dataset used with 6% of IG ranked features are used from the dataset. In summary, IWOA attained comparable results with other algorithms and outperform the standard WOA.

Table 14 Comparison in terms of average classification accuracy in 10 runs between IOWA and deep learning algorithms

Dataset	Algorithms average accuracy (%)			
	IWOA	IWOA + IG	CNN	LSTM
OCA [38]	**95.93**	**99.39**	86.28	78.84
Arabic twitter [53]	**87.61**	**89.68**	75.86	77.97
Political [54]	**88.45**	**88.54**	68.93	70.42
Software [54]	**92.36**	**93.27**	76.90	77.15

As shown on Table 13, IWOA algorithm fitness results indicated in bold font outperforms other algorithms through all used datasets. These results occurred because of capability of IOWA algorithm to select best combinations of features for best solutions with the lowest values of objective function in comparison with other algorithms through all used datasets. Based on our experimental results, it is found that removing irrelevant features by making reduction of features according to IG technique could improve the sentiment classification performance and reduce the search space to be explored by IWOA optimization algorithm. Furthermore, adopting IWOA algorithm as an optimization algorithm after filter features reduction techniques improves the sentiment classification performance and reduce the number of features by selecting the optimal combinations of features. Furthermore, the proposed hybrid model not only increases the accuracy of sentiment classification, but also reduces the number of features by selecting only the most relevant features and remove any irrelevant features.

5.5 Experiment 5: IWOA comparisons with other deep learning algorithms

To examine the effectiveness and performance of the proposed approach in comparison with other deep learning algorithms, experiment 5 was conducted. The accuracy of sentiment classification using IOWA only, and IWOA with IG algorithms are compared with CNN and LSTM deep learning algorithms. The comparison results are shown in Table 14.

It is obviously noticed from Table 14 the superiority of IWOA algorithm accuracy results in comparison with other deep learning algorithms as indicated by bold font. The results of applying IWOA algorithm without IG is outperforming CNN and LSTM over all datasets. Furthermore, using IOWA with IG further improve the accuracy and outperform CNN and LSTM deep learning algorithms. These obtained results confirmed the ability of IOWA to improve classification performance.

6 Conclusion

Due to the increase in review data volumes, automatic and computerized solutions are very crucial. SA has proven to be a great contribution for making decisions for individuals and organizations through identifying the polarity of a review at document, sentence, or aspect (feature) levels. It is widely used in various domains including products, movies, hotels, restaurants, and many others. In the last decade, many research efforts were conducted on SA for languages such as English and Chinese. However, less research efforts were devoted towards Arabic language despite its importance and wide usage. This research focused on Arabic SA to contribute to its state-of-the-art. This research attempted to improve the standard WOA algorithm by improve initialization phase in WOA by using EOBL and enhance local search capability of WOA by using DE evolutionary operators including mutation, crossover, and selections at end of each iteration. The improved algorithm IWOA applied for Arabic SA for reducing the irrelevant features fed to the classifier. Therefore, we used IG filter features selection technique to evaluate and rank the features. Consequently, the best ranked features were fed as input to the IWOA optimization algorithm that works on wrapper mode to further reduce the selected best ranked features into more tightly selected, informative, and relevant features. For optimization purposes, we used and evaluated the IWOA in comparison with WOA, DE, PSO, GA, ALO, and GOA algorithms. In addition, we compared the IWOA algorithm with CNN and LSTM deep learning algorithms.

The main and novel contributions of this study is the Arabic SA hybrid model that takes the advantages of filter feature reduction techniques and optimization algorithms and solves their shortcomings. In addition, the improvement of standard WOA by improve population diversity and quality through using of EOBL strategy. Also, the improvement of WOA by using DE evolutionary operators to avoid it of being stuck at local optima. Furthermore, the proposed two-phase feature selection was accomplished by the hybridization of SVM classifier with IG filter feature reduction technique with IWOA optimization algorithm in wrapper mode. This hybrid model was used for finding and selecting the significant features subsets, and its application to SA in Arabic language. The proposed hybrid model comprises of two phases, whereby the first phase used IG feature selection as a pre-feature reduction technique, which was used to reduce the search space complexity by ranking all features according to IG and removing irrelevant features. In the second phase, the best selected features according to IG weights were used by IWOA algorithm in wrapper mode, which searches for the optimal features combination from these received and selected by IG. In addition, IWOA was used to reduce the number of features and improve the SA classification performance.

In this work, various experiments were conducted using four benchmark publicly available Arabic sentiment analysis datasets, by comparing our proposed IWOA algorithm to different well-known algorithms. The conducted experiments demonstrated that the proposed IWOA algorithm outperformed other algorithms in terms of SA classification accuracy and its fitness value, while it also reduces the selected features. Hence, the proposed IWOA algorithm can perform the Arabic SA feature selection task sufficiently and is comparable to the related works on Arabic SA.

References

1. Schouten K, Frasincar F (2016) Survey on aspect-level sentiment analysis. IEEE Trans Knowl Data Eng 28(3):813–830
2. Rana TA, Cheah YN (2016) Aspect extraction in sentiment analysis: comparative analysis and survey. Artif Intell Rev 46(4):459–483
3. Ahmad SR, Bakar AA, Yaakub MR (2015) Metaheuristic algorithms for feature selection in sentiment analysis. In Science and Information Conference (SAI), 2015 (pp. 222–226). IEEE
4. Liu H, Yu L (2005) Toward integrating feature selection algorithms for classification and clustering. IEEE Trans Knowl Data Eng 17(4):491–502
5. Mirjalili S, Lewis A (2016) The whale optimization algorithm. Adv Eng Softw 95:51–67
6. Aljarah I, Faris H, Mirjalili S (2018) Optimizing connection weights in neural networks using the whale optimization algorithm. Soft Comput 22(1):1–15
7. Wang J, Du P, Niu T, Yang W (2017) A novel hybrid system based on a new proposed algorithm—multi-objective whale optimization algorithm for wind speed forecasting. Appl Energy 208:344–360
8. Prakash DB, Lakshminarayana C (2017) Optimal siting of capacitors in radial distribution network using whale optimization algorithm. Alexandria Eng J 56(4):499–509
9. Reddy PDP, Reddy VV, Manohar TG (2017) Whale optimization algorithm for optimal sizing of renewable resources for loss reduction in distribution systems. Renewables: Wind Water Solar 4(1):3
10. Mostafa A, Hassanien AE, Houseni M, Hefny H (2017) Liver segmentation in MRI images based on whale optimization algorithm. Multimed Tools Appl 76(23):24931–24954

11. Zhou X, Wu Z, Wang H (2012) Elite opposition-based differential evolution for solving large-scale optimization problems and its implementation on GPU. In Parallel and Distrib Comput, Applications and Technologies (PDCAT), 2012 13th International Conference on (pp. 727–732). IEEE
12. Sarkhel R, Das N, Saha AK, Nasipuri M (2018) An improved harmony search algorithm embedded with a novel piecewise opposition based learning algorithm. Eng Appl Artif Intell 67:317–330
13. Shan X, Liu K, Sun PL (2016) Modified Bat Algorithm Based on Lévy Flight and Opposition Based Learning Scientific Programming, 2016
14. Mahdavi S, Rahnamayan S, Deb K (2017) Opposition based learning: a literature review. Swarm and Evolutionary Computation
15. Wang H, Wu Z, Liu Y, Wang J, Jiang D, Chen L (2009) Space transformation search: a new evolutionary technique. Proceedings of the first ACM/SIGEVO Summit on Genetic and Evol Comput (pp. 537–544). ACM
16. Wang H, Wu Z, Rahnamayan S, Liu Y, Ventresca M (2011) Enhancing particle swarm optimization using generalized opposition-based learning. Inf Sci 181(20):4699–4714
17. Guo Z, Wang S, Yue X, Yang H (2017) Global harmony search with generalized opposition-based learning. Soft Comput 21(8):2129–2137
18. Wang H, Wang W, Zhu H, Sun H (2013) Generalised opposition-based differential evolution for frequency modulation parameter optimisation. Int J Model Identif Control 18(4):372–379
19. Si T, De A, Bhattacharjee AK (2014) Particle swarm optimization with generalized opposition based learning in particle's pbest position. In Circuit, Power and Computing Technologies (ICCPCT), 2014 International Conference on (pp. 1662–1667). IEEE
20. Wang H, Wang W, Sun H (2015) Firefly algorithm with generalised opposition-based learning. Int J Wirel Mob Comput 9(4):370–376
21. Zhou X, Wu Z, Deng C, Peng H (2015) Enhancing artificial bee colony algorithm with generalised opposition-based learning. Int J Comput Sci Math 6(3):297–309
22. Chen X, Yu K, Du W, Zhao W, Liu G (2016) Parameters identification of solar cell models using generalized oppositional teaching learning-based optimization. Energy 99:170–180
23. Zhou Y, Wang R, Luo Q (2016) Elite opposition-based flower pollination algorithm. Neurocomputing 188:294–310
24. Zhang S, Luo Q, Zhou Y (2017) Hybrid grey wolf optimizer using elite opposition-based learning strategy and simplex method. Int J Comput Intell Appl 16(02):1750012
25. Huang K, Zhou Y, Wu X, Luo Q (2016) A cuckoo search algorithm with elite opposition-based strategy. J Intell Syst 25(4):567–593
26. Storn R, Price K (1997) Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. J Glob Optim 11(4):341–359
27. Jadon SS, Tiwari R, Sharma H, Bansal JC (2017) Hybrid artificial bee colony algorithm with differential evolution. Appl Soft Comput 58:11–24
28. Zheng YJ, Xu XL, Ling HF, Chen SY (2015) A hybrid fireworks optimization method with differential evolution operators. Neurocomputing 148:75–82
29. Awad NH, Ali MZ, Suganthan PN, Reynolds RG (2017) CADE: a hybridization of cultural algorithm and differential evolution for numerical optimization. Inf Sci 378:215–241
30. Wang G, Sun J, Ma J, Xu K, Gu J (2014) Sentiment classification: the contribution of ensemble learning. Decis Support Syst 57:77–93
31. Moraes R, Valiati JF, Neto WPG (2013) Document-level sentiment classification: an empirical comparison between SVM and ANN. Expert Syst Appl 40(2):621–633
32. Ye Q, Zhang Z, Law R (2009) Sentiment classification of online reviews to travel destinations by supervised machine learning approaches. Expert Syst Appl 36(3):6527–6535
33. Saleh MR, Martín-Valdivia MT, Montejo-Ráez A, Ureña-López LA (2011) Experiments with SVM to classify opinions in different domains. Expert Syst Appl 38(12):14799–14804
34. Aliane AA, Aliane H, Ziane M, Bensaou N (2016) A genetic algorithm feature selection based approach for Arabic sentiment classification, IEEE
35. Duwairi RM, Qarqaz I (2016) A framework for Arabic sentiment analysis using supervised classification. Int J Data Mining Model Manag 8(4):369–381
36. Al-Radaideh QA, Al-Qudah GY (2017) Application of rough set-based feature selection for Arabic sentiment analysis. Cogn Comput, 1–10
37. Sabri B, Saad S (2016) Arabic sentiment analysis with optimal combination of features selection and machine learning approaches. Res J Appl Sci Eng Technol 13(5):386–393
38. Rushdi-Saleh M, Teresa M-VM, Urea-Lopez AL, Perea-Ortega JM (2011) OCA: opinion corpus for Arabic. J Am Soc Inf Sci 62:2045–2054
39. Bahassine S, Madani A, Kissi M (2016) An improved Chi-sqaure feature selection for Arabic text classification using decision Trees In Intelligent Systems: Theories and Applications (SITA), 2016 11th International Conference on (pp. 1–5). IEEE
40. Al-Kabi MN, Abdulla NA, Al-Ayyoub M (2013) An analytical study of arabic sentiments: Maktoob case study. In Internet Technology and Secured Transactions (ICITST), 2013 8th International Conference for (pp. 89–94). IEEE
41. Abdulla NA, Ahmed NA, Shehab MA, Al-Ayyoub M (2013) Arabic sentiment analysis: Lexicon-based and corpus-based. In Applied Electr Eng and Computing Technologies (AEECT), 2013 IEEE Jordan Conference on (pp. 1–6). IEEE
42. Alomari KM, ElSherif HM, Shaalan K (2017) Arabic Tweets Sentimental Analysis Using Mach Learn In International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems (pp. 602–610). Springer, Cham
43. Abdulla NA, Al-Ayyoub M, Al-Kabi MN (2014) An extended analytical study of Arabic sentiments. Int J Big Data Intell 1 1(1–2):103–113
44. Omar N, Albared M, Al-Moslmi T, Al-Shabi A (2014) A comparative study of feature selection and Mach Learn algorithms for Arabic sentiment classification. In Asia Inf Retr symposium (pp. 429–443). Springer, Cham
45. Bostani H, Sheikhan M (2017) Hybrid of binary gravitational search algorithm and mutual information for feature selection in intrusion detection systems. Soft Comput 21(9):2307–2324
46. Zhang L, Shan L, Wang J (2017) Optimal feature selection using distance-based discrete firefly algorithm with mutual information criterion. Neural Comput & Applic 28(9):2795–2808
47. Wang H, Niu B (2017) A novel bacterial algorithm with randomness control for feature selection in classification. Neurocomputing 228:176–186
48. Barani F, Mirhosseini M, Nezamabadi-Pour H (2017) Application of binary quantum-inspired gravitational search algorithm in feature subset selection. Appl Intell 47(2):304–318
49. Hancer E, Xue B, Zhang M, Karaboga D, Akay B (2018) Pareto front feature selection based on artificial bee colony optimization. Inf Sci 422:462–479
50. Wolpert DH, Macready WG (1997) No free lunch theorems for optimization. IEEE Trans Evol Comput 1(1):67–82
51. Tizhoosh HR (2005) Opposition-based learning: a new scheme for machine intelligence. In Computational intelligence for modelling, control and automation, 2005 and international conference on intelligent agents, web technologies and internet commerce, international conference on (Vol. 1, pp. 695–701). IEEE

52. Song Q, Ni J, Wang G (2013) A fast clustering-based feature subset selection algorithm for high-dimensional data. IEEE Trans Knowl Data Eng 25(1):1–14
53. Abdulla N, Mahyoub N, Shehab M, Al-Ayyoub M (2013) Arabic sentiment analysis: Corpus-based and lexicon-based. In Proceedings of The IEEE conference on Applied Electrical Engineering and Computing Technologies (AEECT)
54. Al-Moslmi T, Albared M, Al-Shabi A, Omar N, Abdullah S (2017) Arabic senti-lexicon: constructing publicly available language resources for arabic sentiment analysis. J Inf Sci, 0165551516683908
55. Saremi S, Mirjalili S, Lewis A (2017) Grasshopper optimisation algorithm: theory and application. Adv Eng Softw 105:30–47
56. Mirjalili S (2015) The ant lion optimizer. Adv Eng Softw 83:80–98
57. Bahassine S, Madani A, Al-Sarem M, Kissi M (2018) Feature selection using an improved chi-square for Arabic text classification. Journal of King Saud University-Computer and Information Sciences
58. Al-Smadi M, Al-Ayyoub M, Jararweh Y, Qawasmeh O (2018) Enhancing aspect-based sentiment analysis of Arabic hotels' reviews using morphological, syntactic and semantic features. Inf Process Manag
59. Abdul-Mageed M, Diab M, Kübler S (2014) SAMAR: subjectivity and sentiment analysis for Arabic social media. Comput Speech Lang 28(1):20–37
60. Soliman TH, Elmasry MA, Hedar A, Doss MM (2014) Sentiment analysis of Arabic slang comments on facebook. Int J Comput Technol 12(5):3470–3478
61. Nasser A, Dinçer K, Sever H (2016) Investigation of the feature selection problem for sentiment analysis in Arabic language. Res Comput Sci 110:41–54
62. Al-Rubaiee H, Qiu R, Li D (2016). Identifying Mubasher software products through sentiment analysis of Arabic tweets. In Industrial Informatics and Computer Systems (CIICS), 2016 International Conference on (pp. 1–6). IEEE
63. Cherif W, Madani A, Kissi M (2015) A new modeling approach for Arabic opinion mining recognition. In 2015 Intelligent Systems and Computer Vision (ISCV) (pp. 1–6). IEEE
64. Ibrahim HS, Abdou SM, Gheith M (2015) Sentiment analysis for modern standard Arabic and colloquial. arXiv preprint arXiv: 1505.03105

Mohammad A. M. Abushariah is working as an Associate Professor at the Department of Computer Information Systems, King Abdullah II School of Information Technology, The University of Jordan. He obtained his bachelor degree in information technology from the International Islamic University Malaysia, Malaysia in 2005. He then obtained his Master degree in software engineering and Ph.D. in computer science and information technology specialized in Natural Language Processing (NLP) and speech processing from University of Malaya, Malaysia in 2007 and 2012, respectively. He is a supervisor of various Ph.D. and Master research students specialized in Arabic NLP, speech processing, and software engineering. He has over 50 publications in ISI journals, IEEE international conferences, and technical reports. He was appointed as the Guest Editor of a special issue (Volume 19, No. 2) on Arabic Natural Language Processing and Speech Recognition: A study of algorithms, resources, tools, techniques, and commercial applications, which was published in 2016 by the International Journal of Speech Technology (IJST), Springer. In January 2017, he was appointed as an External Associate Editor for the Malaysian Journal of Computer Science (MJCS), which is indexed by ISI/WoS (JCR). In April 2015, he was appointed as an Editorial Team Member of the International Journal of Open Information Technologies (INJOIT). He also serves as a reviewer for various specialized international journals indexed in ISI/WoS (JCR) and conferences. In April 2014, he was appointed as a Consultant of NLP and speech recognition for Samsung R&D Institute in Jordan working with a distributed team with Samsung Headquarter in South Korea. In September 2014, he was appointed as the assistant dean for student affairs at School of Graduate Studies, The University of Jordan. In September 2015, he was appointed as the assistant dean for development and quality affairs at School of Graduate Studies, The University of Jordan. Since September 2016 until September 2018, he was appointed as the deputy dean at King Abdullah II School of Information Technology, The University of Jordan. His research interests include: Arabic NLP, Arabic speech processing, text and speech corpora, language resources production, speaker recognition using biometrics, speech emotion recognition, Arabic sign language, sentiment analysis, and software engineering.

Mohammad Tubishat received a BSc degree in Computer Science from Yarmouk University - Jordan, and master's degree in Computer and Information Sciences from Yarmouk University - Jordan. He is currently a PhD student in Artificial Intelligence department at University of Malaya. His research interest includes natural language processing, data mining, artificial intelligence, machine learning, optimization algorithms, and sentiment analysis

Dr. Norisma Idris joined Faculty of Computer Science and IT, University of Malaya in 2001. She is currently the senior lecturer in Artificial Intelligence Department. She received her Ph.D in Computer Science from University of Malaya in 2011. Her research interest is on Natural Language Processing where the main focus is on developing efficient algorithms to process texts and to make their information accessible to computer applications, mainly for text normalization and sentiment analysis

Dr. Ibrahim Aljarah is an associate professor of BIG Data Mining and Computational Intelligence at the University of Jordan - Department of Business Information Technology, Jordan. He obtained his bachelor degree in Computer Science from Yarmouk University - Jordan, 2003. Dr. Aljarah also obtained his master degree in computer science and information systems from the Jordan University of Science and Technology - Jordan in 2006. He also obtained his Ph.D. In computer Science from the North Dakota State University (NDSU), USA, in May 2014. He organized and participated in many conferences in the field of data mining, machine learning, and Big data such as NTIT, CSIT, IEEE NABIC, CASON, and BIGDATA Congress. Furthermore, he contributed in many projects in USA such as Vehicle Class Detection System (VCDS), Pavement Analysis Via Vehicle Electronic Telemetry (PAVVET), and Farm Cloud Storage System (CSS) projects. He has published more than 35 papers in refereed international conferences and journals. His research focuses on data mining, Machine Learning, Big Data, MapReduce, Hadoop, Swarm intelligence, Evolutionary Computation, Social Network Analysis (SNA), and large scale distributed algorithms.