Flower Pollination Algorithm for Feature Selection in Tweets Sentiment Analysis

Muhammad Iqbal Abu Latiffi[1], Mohd Ridzwan Yaakub[2], Ibrahim Said Ahmad[3]
Center for Artificial Intelligence Technology, Faculty of Information Science and Technology
Universiti Kebangsaan Malaysia, Bangi Malaysia[1, 2]
Faculty of Computer Science and Information Technology, Bayero University Kano, Kano Nigeria[3]

***Abstract*—Text-based social media platforms have developed into important components for communication between customers and businesses. Users can easily state their thoughts and evaluations about products or services on social media. Machine learning algorithms have been hailed as one of the most efficient approaches for sentiment analysis in recent years. However, as the number of online reviews increases, the dimensionality of text data increases significantly. Due to the dimensionality issue, the performance of machine learning methods has been degraded. However, traditional feature selection methods select attributes based on their popularity, which typically does not improve classification performance. This work presents a population-based metaheuristic for feature selection algorithms named Flower Pollination Algorithms (FPA) because of their propensity to accept less optimum solutions and avoid getting caught in local optimum solutions. The study analyses tweets from Kaggle first with the usual Term Frequency-Inverse Document Frequency statistical weighting filter and then with the FPA. Four baseline classifiers are used to train the features: Naive Bayes (NB), Decision Tree (DT), Support Vector Machine (SVM), and k-Nearest Neighbor (kNN). The results demonstrate that the FPA outperforms alternative feature subset selection algorithms. For the FPA, an average improvement in accuracy of 2.7% is seen. The SVM achieves a better accuracy of 98.99%.**

***Keywords*—Sentiment analysis; metaheuristic algorithm; flower pollination algorithm; machine learning; feature selection**

I. Introduction

The introduction of the second generation of the web has resulted in the development of more interactive websites where users will play a significant role by contributing their thoughts, suggestions, and comments via the website's services. Large-scale communication, which is not limited to comments on news, services, or even goods or products, can also be used to facilitate conversation amongst website users within social networks. Social networks allow users to contact one another, engage, and collaborate to become content creators on a single platform known as social media. Through social media, they can provide views, comments, and experiences on an issue [1]–[6].

There are many web users, and there is also a large amount of information sharing in the form of reviews, which makes analyzing the reviews manually tricky and time-consuming. Sentiment Analysis (SA) has been introduced as a practical approach to determining the sentiments included in web materials to tackle this issue [7]–[10].

SA is a type of text classification problem that involves subjective statements. Another name for sentiment analysis is Opinion Mining (OP), where an opinion or comment is processed to determine the consumer's perception of a matter or issue. SA is one of the areas of data mining and text mining that falls under the category of web content mining techniques [8], [11]–[13]. It is defined as the computer study of public perceptions, beliefs, and moods around a specific topic [14]. It is a subset of natural language processing jobs that assesses a large amount of viewer content on social media, blog posts, e-commerce portals, and other user-editable online forums.

One of the primary challenges associated with sentiment analysis is the preprocessing phase of the text, in which they must handle user feedback on social media, webpages, and blog posts that contain irrelevant and noisy content. [15] studied the effects of text preprocessing in sentiment analysis. They found inconsistent text classification results, which is likely due to inefficient text preprocessing methods. Most researchers use only a fraction of the text preprocessing techniques [6], [16], [17]. In addition to improvements in the feature selection phase, the text preprocessing phase also plays a significant role in improving text or sentiment classification performance [2], [18]. According to [18], the habit will increase the results of text classification accuracy.

The magnitude of the feature dimension for text parsing is an additional key difficulty in sentiment analysis. Bag-of-Words is commonly used to represent document text in machine learning algorithms for sentiment categorization [19], [20]. Words in a text document contain feature vectors with large dimensions. As a result, the feature selection process focuses on selecting the optimal subset of features from a large-dimensional feature size. This is accomplished by removing cluttered and irrelevant features without altering the original data [14], [21].

According to [19], it is vital to select an optimal subset of features that represents the actual feature subsets to reduce feature size and increase classification accuracy. An optimization algorithm was utilized as the strategy for selecting features [14]. Previous research [22]–[25] has demonstrated that an optimization algorithm strategy for feature selection can handle feature selection and feature reduction issues in large amounts of data containing noise, redundant, and inaccurate information.

Flower Pollination Algorithm (FPA) is one of the algorithms inspired by nature. This algorithm has advantages in

terms of performance when it is applied in improving performance in various optimization issue problems as well as having few parameters [22], [26], [27]. According to [22], the FPA is a flexible and easily adaptable optimization method. However, research on the use of FPA in feature selection problems in sentiment analysis has not yet been conducted. So, this is necessary to study and further improve the performance of FPA for feature selection problems. Therefore, the development of feature selection methods that use FPA is expected to be able to select a more discriminatory feature set and improve the sentiment classification results.

On the other hand, the sentiment classification process is another issue in sentiment analysis technology due to the involvement of the textual data [5]. Furthermore, simple text classification and sentiment classification in text mining are two different things. Text classification in text mining identifies topics found in a data set. The classification of sentiment, on the other hand, is classified in terms of the type of sentiment in the text. When there are strategies to minimize the dimensions in large-sized data sets, sentiment classification performance can be enhanced [10], [28], [29]. This method generally serves to detect and eliminate irrelevant and overlapping data so that the sentiment classification results are meaningful.

The remaining sections of this work are structured as follows: The second section examines previous works on feature selection and the necessary component of FPA. Next, Section 3 provides an overview of the technique applied in this study, while Section 4 details the FPA algorithm proposed for feature selection in sentiment analysis. The findings of the experiments are discussed in Section 5, and the study is concluded in Section 6.

II. Related Work

Different approaches for selecting features have been proposed and developed, and they are explored by researchers. Metaheuristic techniques improve solution performance over and over again [30]. These metaheuristic techniques are based on observations of natural phenomena such as ant colony optimization, bat algorithms, and flower pollination optimization. This metaheuristic is then used in the bandage feature selection method. This method has been gaining more attention recently [31] as it attempts to produce better solutions by applying knowledge gained from natural solutions. The wrapper approach in feature selection evaluates the quality of the selected features based on their classification performance. The wrapping method has two main steps: (1) finding a subset of features and (2) evaluating the selected features. Both of these steps will be repeated until the stop criteria are met. It starts with the production of a subset, and then the classification will evaluate the subset produced.

Previous studies proposed feature selection methods based on particle clustering optimization and the nearest k-neighbor classification [24]. In addition, based on an analysis from [32] also proposed a hybrid approach of feature selection based on genetic algorithms for the classification of Arabic texts using a wrapper model. In the first step, six feature evaluation methods are used simultaneously to select a subset of features. Then an enhanced genetic algorithm is used to optimize the selected subset.

The technique of selecting the optimal subset of features using firefly's algorithm for sentiment analysis problems has also proven the ability of this metaheuristic algorithm [16]. While in a study conducted by [17] where researchers improved the whale optimization algorithm for feature selection of two types of sentiment analysis data, namely in Arabic and English. Moreover, the hybridization between the ant colony optimization algorithm and k-Nearest Neighbors has successfully improved the sentiment classification results as it has been applied to the feature selection process [6]. These studies report that their method is more effective compared to the use of filter methods.

The FPA is an algorithm inspired by nature that imitates the basic pollination activity of flowers. In[33], four rules are idealized. Rule 1 of global pollination incorporates biotics and cross-pollination, with the pollinating agent transporting pollen according to the Lévy flight. Rule 2 requires abiotic and self-pollination for local pollination. Next, for Rule 3, the flower constant can be interpreted as the reproduction probability proportional to the degree of resemblance between two flowers. Rule 4 is the exchange probability p [0, 1], which can be regulated by external factors such as wind between local and global pollination. Local pollination accounted for a sizeable proportion of the total pollination activity.

FPA was successfully adapted for several domains of optimization problems [22]. For the field of electrical and power generation, [34] has introduced a Modified FPA (MFPA) that employs dynamic switching probabilities, the application of Real Coded Genetic Algorithm (RCGA) mutations for global and local searches, and differentiation between searches temporary localization and optimal solutions. The MFPA was subsequently assessed for ten power system benchmarks, and the experimental results revealed lesser fuel costs than the FPA. Another study by [35] presented MFPA to assess the fuel funding and time required to get to the globally optimal solution. The Institute of Electrical and Electronics Engineers 30 (IEEE 30) bus test system demonstrated that MFPA outperformed FPA and metaheuristic algorithms.

Next, research involving FPA was conducted on the signal and image processing domains. The Binary Flower Pollination Algorithm (BFPA) has been implemented to solve the challenge of lowering the number of sensors necessary to identify individuals' electroencephalogram signals [36]. BFPA is used to choose the ideal selection of channels that provide maximum accuracy. Based on the Optimum-Path Forest classifier, the findings of the BFPA experiment indicate an identification rate of up to 87 percent. In addition, [37] has done a thorough evaluation of BFPA's performance in solving the Antenna Positioning Problem (APP) area. BFPA was evaluated with real, artificial, and random data of various dimensions and compared with Population-Based Increment Learning (PBIL) and Differential Evolution (DE) algorithms, two efficient APP domain methods. In the sphere of APP, FPA obtains more competitive technical advances than PBIL and DE.

FPA is also not left behind to be applied to the clustering and classification domain. The performance of the modified FPA is tested through a clustering problem. This algorithm was evaluated with several different optimization algorithms, including bat, firefly, and conventional FPA on 10 cluster data sets. From the 10 data sets, 8 were generated from pattern recognition, and 2 were generated artificially. The clustering result is calculated in terms of the value of the objective function and the time taken by the CPU on each run. The distribution length graph illustrates the convergence behavior of the algorithm. The results show that the modified FPA exceeds the comparable algorithm to achieve the best fitness value and reduce CPU processing time [38].

Aside from that, the BFPA was applied to feature selection problems and tested on six data sets, where it outperformed Particle Swarm Optimization (PSO), Harmony Search (HS), and Firefly Algorithm (FA) [39]. BCFA is a hybrid algorithm that combines the Clonal Selection Algorithm (CSA) with FPA to tackle feature selection issues. It was introduced in [40]. Using the Optimum-Path Forest classifier as an objective function with the proposed hybrid algorithm (BCFA) has led to superior performance compared to existing metaheuristics. A new methodology for multi-objective feature selection is based on FPA and rough set theory to identify the optimal classification feature set [41]. This model selects features using the filter and wrapper method. The filter approach is a data-driven methodology, whereas the wrapper method is a classification-based technique. Comparing this method against FPA, PSO, and genetic algorithms, the performance of the suggested method was validated using eight UCI data sets, which revealed that this method is extremely competitive. The addition of FPA to the Ada-Boost algorithm enhances the classification accuracy of text documents during the initial phase of feature selection. In contrast, it is utilized to categorise text materials. Three standard data sets were used to evaluate the performance of the proposed algorithm: CADE 12, WEBKB, and Reuters-21578. The experimental findings demonstrated that the suggested algorithm outperformed Ada-Boost and other algorithms [42].

This paper proposes a metaheuristic approach called the flower pollination algorithm for text feature selection based on the Twitter dataset to increase the accuracy of sentiment classification.

III. Methodology

In general, the SA methodology is intended to yield the optimal subset of features and the most accurate sentiment classification. As illustrated in Fig. 1, the approach for this study consists of four phases: text preprocessing, feature selection, sentiment classification, testing, assessment, and analysis. The tasks associated with these phases are described in the sections that follow.

A. Phase 1: Text Preprocessing

As tweets are composed by regular people who are not language specialists, the Twitter dataset underwent a cleansing procedure. As a result, misspellings, grammatical errors such as faulty punctuation and capitalization, slang words that do not exist in dictionaries, and abbreviations or acronyms for common terminology are likely to be present in the datasets. The dataset is therefore subjected to two forms of text preprocessing: linguistic processing and natural language processing. This study's linguistic processing comprises five text processing approaches, including lowercase conversion, removal of '#', '@', and other symbols, and removal of punctuation. The document then undergoes spelling correction, an NLP approach. The preprocessing techniques used in this work are shown in Table I.

B. Phase 2: FPA for Feature Selection

Wind or pollinating agents, such as insects, butterflies, bees, birds, and bats, carry pollen from one flower to another during pollination. Flowering plants have evolved to generate nectar or honey in order to entice pollinators and ensure pollination [43]. In addition, a number of pollination agents and plant species, such as squirrels and ornithophilous flowering plants that are pollinated by birds, constitute a number of floral evolution constants [44]. Based on the main characteristics of pollination, a flower pollination algorithm has been developed by [33] developed an algorithm for flower pollination.

There are two fundamental types of flower pollination: biotic and abiotic processes. Biotic pollination, also known as cross-pollination, is the most common type of pollination and is carried out by pollinating agents like insects, birds, and others. This method of pollination is utilized by over 90 percent of blooming plants. When the pollinating agent travels and even flies at varying speeds, the movement of the pollen is quite remote; such pollination can also be considered global pollination when Lévy Flight rules [33], [45], [46] are used. If pollen is encoded as a solution vector, this operation corresponds to a global search. Abiotic pollination, also known as self-pollination, does not require an external pollinator. Approximately 10 percent of flowering plants utilize this kind of pollination, according to estimates. Since self-pollination and localised pollination are more likely to occur in this way, wind dispersal is an option [43], [44]. Typically, the distance traversed by a local movement is shorter, therefore the search might be deemed local. The aforesaid characteristics were utilised to plan the creation of the flower pollination algorithm (FPA) [33], an optimization algorithm.

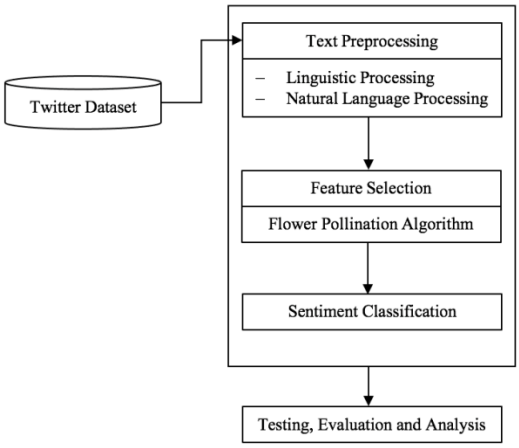

Fig. 1. Methodology for the Study.

TABLE I. TEXT PREPROCESSING AND EXAMPLES

Techniques	Raw	Processed
Conversion to lowercase	@Piwi_47 I hated the da Vinci code, the movie with a passion, it was boring and made me sad!! that a movie blaspheming the name of Jesus is being played world wide $. #davincicode	@piwi_47 i hated the da vinci code, the movie with a passion, it was boring and made me sad!! that a movie blasphying the name of jesus is being played world wide $. #davincicode
Removal of @	@piwi_47 i hated the da vinci code, the movie with a passion, it was boring and made me sad!! that a movie blasphying the name of jesus is being played world wide $. #davincicode	i hated the da vinci code, the movie witha passion, it was boring and made me sad!! that a movie blasphying the name of jesus is being played world wide $. #davincicode
Removal of punctuation	i hated the da vinci code, the movie witha passion, it was boring and made me sad!! that a movie blasphying the name of jesus is being played world wide $. #davincicode	i hated the da vinci code the movie with a passion it was boring and made me sad that a movie blasphying the name of jesus is being played world wide $ #davincicode
Removal of #	i hated the da vinci code the movie witha passion it was boring and made me sad that a movieblasphying the name of jesus is being played world wide $ #davincicode	i hated the da vinci code the movie with a passion it was boring and made me sad that a movie blasphying the name of jesus is being played world wide $
Removal of symbol	i hated the da vinci code the movie with a passion it was boring and made me sad that a movie blasphying the name of jesus is being played world wide $	i hated the da vinci code, the movie with a passion it was boring and made me sad that a movie blasphying the name of jesus is being played world wide
Spelling correction	it was really ironic that he spent the first part of class talking about his own **professot** at Harvard who was a pompous arrogant ass	it was ironic that he spent the first part of class talking about his **professor** at Harvard who was a pompous arrogant ass

The development of the FPA is based on the pseudo-code shown in Fig. 2. In general, the development of the FPA consists of three main parts - parameter declaration, initiation, and searching.

```
    Parameter Declaration
1   Initialize population size randomly
2   Initialize stopping criteria (iteration number)
3   Set the switch probability
4   Set the best solution

    Initiation phase
5   Check if stopping criteria meet
6   {

    Searching Phase
        Global pollination
7       Choose the best solution
8       Random solution < switch probability
9       Global pollination occurs
10      New solution produced
11      If new solution better
12          Subs the current solution with the new
13      Else
14          New solution rejected
        Local pollination
15      Choose two random solution
16      Local pollination occurs
17      New solution produced
18      If new solution better
19          Subs the current solution with the new
20      Else
21          New solution rejected
22      Repeat step # 6
23  }
24  Select the optimal solution
```

Fig. 2. Pseudocode for FPA.

1) Parameter declaration phase: There are two types of parameters that must be set during the general parameter declaration phase: population size and the number of iterations. The population parameter choice for this experiment is based on a study by [47], which determined that the optimal population size for optimal results is 25. The number of generations is determined based on a study conducted by [33], i.e., the most appropriate number of generations is 100. This number of generations will be the termination criterion for the algorithm.

2) Initiation phase: In this initial phase, a subset of solutions will be randomly generated and stored in the form of a one-dimensional array, as in the example shown in Fig. 3. The illustration depicts a solution subset with ten feature attributes labeled F1 through F10. A cell with a value of 1 indicates the attribute of the selected feature, while a cell with a value of 0 indicates the attribute of the unselected feature. Next, all subsets of these generated initial solutions will be evaluated based on the performance of the sentiment classification accuracy assessment and sorted based on the classification accuracy score values obtained. A subset of this initial solution will be used to generate the following generation subset.

3) Searching phase: For the searching phase, the algorithm is divided into two parts, namely, global pollination and local pollination. Before proceeding to the search phrase, the termination criteria will be reviewed first. If the termination criteria are not met, the search phase will continue. Conversely, suppose the termination criteria have been met. In that case, the subset of features with the best sentiment classification score will be returned and used as a feature list for sentiment classification.

Global pollination begins when a solution is randomly generated. If its value is smaller than the exchange probability, the Levy flight rule will generate a new solution. Existing solutions will be selected along with these new solutions to get better scores. The crossover method in the global pollination method is illustrated in Fig. 4. The global pollination session will result in two new solutions, and both will go through a sentiment classification performance appraisal process. Only new solutions that are capable of producing higher scores than the existing solution score list is accepted into the population, and their position within the population is decided by the scores produced. On the other hand, this new solution will not be accepted if it gets a score that is less than the score of the existing population.

F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
1	0	1	0	1	1	1	1	0	0

Fig. 3. Solution Subset Array.

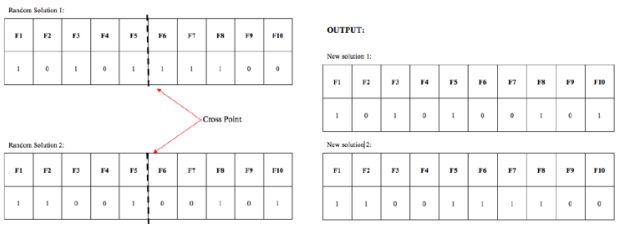

Fig. 4. Global Pollination Process.

On the other hand, for local pollination sessions, it occurs when a solution is randomly generated and if its value is greater than the exchange probability. Therefore, two randomly generated solutions were selected to undergo local pollination. As depicted in Fig. 5, the process of learning between the two candidates in the FPA algorithm occurs via the crossover approach. The outcomes of the local pollination will generate two new solutions, which are then evaluated based on their performance in sentiment classification. If based on the obtained sentiment evaluation score, this new solution is deemed superior to the previous solution, it will be accepted and its population position will be changed. Alternatively, this solution will not be approved if it has a low score among the current solutions.

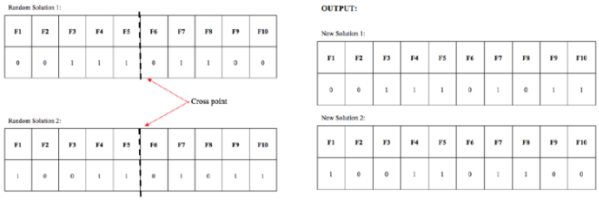

Fig. 5. Local Pollination Process.

These global and local pollination processes will continue until the termination criteria are reached. If the criteria for stopping are not reached, this procedure will be repeated. As illustrated in Fig. 6, the result of this algorithm is a subset of quality and modest characteristics that will be employed in the sentiment classification step.

F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
1	0	1	0	1	1	1	0	0	0

Fig. 6. New Feature Subset that Selected.

C. Phase 3: Sentiment Classification

Classification is one of the key processes in machine learning that refers to determining input data assigned to one of predetermined categories or classes [48]. Classifiers will learn with a learning algorithm or classification algorithm, also known as a classification initiator, a supervised machine learning algorithm. Machine learning algorithms use a set of examples to learn to classify the class label of something that has not been seen or has not been learned. The classifier that has been learned will take the feature value or attribute of an object as input and the class label that has been defined as the output. A set of class labels is defined as part of a problem by the user.

Therefore, the third phase involves four machine learning classifier algorithms to carry out the sentiment classification process by matching a subset of the features generated with the feature information found in the text. The algorithms used are SVM, NB, DT, and kNN.

D. Phase 4: Testing, Evaluation and Analysis

1) Dataset: For this study, a benchmark Twitter data set in which has been used by the previous researcher [49] can be obtained from the Kaggle repository. This data set consists of 7086 positive and negative Twitter comments written in English.

2) Baseline model: The performance of the FPA feature selection technique was evaluated through a comparison with two baseline algorithms, namely, TF-IDF and Binary Cuckoo Search, by [49]. The evaluation is based on the sentiment classification accuracy from the tweets by using a subset of features selected from phase 3. A subset of these features is obtained upon the text preprocessing process, reduction of feature dimension size, and subsequently feature selection process. This study will not cover the time required for the text preprocessing phase, feature selection, and even classification. Thus, measurements of the time rate will not be performed.

3) Evaluation: The accuracy measure was used to evaluate the efficacy of the FPA feature selection technique in getting the optimal subset of features. The evaluation is based on the accuracy of the sentiment classification process outcomes. The classification algorithm generates a confusion matrix, which is used to guide the evaluation process. The confusion matrix displays information about the actual number of classes as well as the number of predictions made by the classification algorithm. True positive (TP) is a condition in which a positive case is successfully classified as positive. True negative (TN) are conditions in which a negative case is successfully classified as negative. A false positive (FP) is a negative case but is misclassified as a positive. False-negative (FN) are positive cases but misclassified as negative ones, as shown in Table II.

The accuracy of the proposed feature selection algorithm was used to evaluate its effectiveness. Accuracy is a simple performance metric derived as the ratio of successfully predicted values to total values. The equation is shown in Equation 1.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

TABLE II. THE CONFUSION MATRIX

		Actual Class	
		Yes	**No**
Classification Result	**Yes**	TP	FN
	No	FP	TN

IV. Flower Pollination for Feature Selection

For this experiment, the evaluation performance was evaluated based on accuracy. These values are compared with the results obtained from baseline algorithms, as presented in Table III.

Table III shows that the FPA feature selection algorithm achieves the highest accuracy value compared to the TF-IDF and Binary Cuckoo Search (BCS). Fig. 7 displays the overall accuracy values for the APB algorithm and TF-IDF and BCS based on the classification algorithm used. Based on Table III, it is found that the accuracy values for TF-IDF for NB, SVM, DT, and k-NN are 87.71%, 89.21%, 89.90%, and 88.42%, respectively. While for Binary Cuckoo Search, the resulting accuracy values are 96.26%, 96.54%, 96.26%, and 95.56% for the NB, SVM, DT, and k-NN, respectively. Based on Fig. 7, the FPA feature selection algorithm has produced the highest accuracy compared to the other two feature selection algorithms. This is where the accuracy values obtained for the NB, SVM, DT, and k-NN are 98.79%, 98.99%, 98.76%, and 98.91%, respectively.

This experiment went through 100 iterations. As shown in Fig. 8, at the 45th iteration, the algorithm is approaching the optimum value at a rapid pace of convergence. This indicates that the algorithm can discover a solution quickly. The increment on accuracy rate then slows until the 75th iteration. This is because this algorithm has consolidated nearly every high feature set. Achieving a subset of features capable of obtaining higher accuracy values got progressively challenging because the current feature set already had a relatively high quality of accuracy values—the accuracy rate value peaks as it reaches the 75th iteration. As a result, until the 100th iteration, this accuracy value remains constant.

TABLE III. Accuracy Performance for FPA, TF-IDF and BSC using Naïve Bayes, Support Vector Machine, Decision Tree and k-Nearest Neighbor

Classifiers	FPA	BCS	TF-IDF
Naïve Bayes	**98.79**	96.26	87.71
Support Vector Machine	**98.99**	96.54	89.12
Decision Tree	**98.76**	96.26	89.90

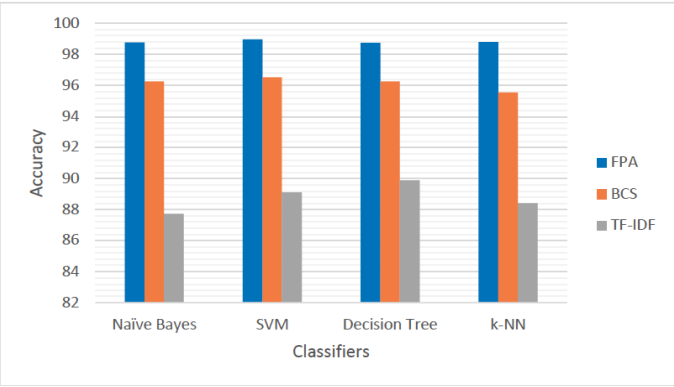

Fig. 7. Comparison of Accuracy for FPA, TF-IDF, and BCS.

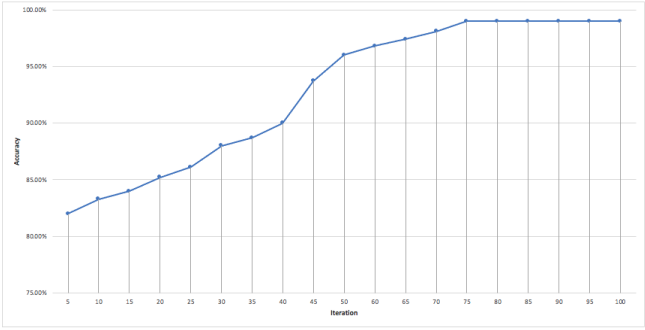

Fig. 8. Classification.

The proposed algorithm for feature selection was able to attain greater accuracy than the two baseline techniques. This result demonstrates that FPA is more effective at extracting features than the baseline techniques.

V. Conclusion

This study compares and contrasts the NLP method, spelling correction in text preprocessing techniques, with other conventional text preprocessing techniques. In addition, the use of FPA algorithms for feature selection strategies to enhance the performance of sentiment classification has been proposed. Based on the accuracy results, our approach achieved promising results, confirming that implementing NLP (spelling correction) approach on text preprocessing technique and FPA algorithm on feature selection technique improved sentiment classification performance by 2.68 % compared to the baseline model.

In the future, we would like to employ the proposed technique for solving sentiment analysis tasks on a larger data set containing product review evaluations.

Acknowledgment

The authors gratefully acknowledge Universiti Kebangsaan Malaysia and the Industrial Grant Scheme with Perunding Tamadun Teras Pte Ltd for supporting this research project through grant no. GP-2020-K011466 and TT-2020-008.

References

[1] Dritsas, G. Vonitsanos, and I. E. L. B, “Pre-processing Framework for Twitter,” IFIP Int. Fed. Inf. Process. 2019, vol. 2, pp. 138–149, 2019, doi: 10.1007/978-3-030-19909-8.

[2] S. R. Priya and M. Devapriya, “The role of pre-processing on unstructured and informal text in diabetic drug related twitter data,” Int. J. Sci. Technol. Res., vol. 8, no. 10, pp. 607–611, 2019.

[3] L. Zhou et al., “Text preprocessing for improving hypoglycemia detection from clinical notes – A case study of patients with diabetes,” Int. J. Med. Inform., vol. 129, no. January, pp. 374–380, 2019, doi: 10.1016/j.ijmedinf.2019.06.020.

[4] I. S. Ahmad, A. Abu Bakar, M. R. Yaakub, and M. Darwich, “Sequel movie revenue prediction model based on sentiment analysis,” Data Technol. Appl., vol. 54, no. 5, pp. 665–683, 2020, doi: 10.1108/DTA-10-2019-0180.

[5] M. R. Yaakub, M. I. A. Latiffi, and L. S. Zaabar, “A Review on Sentiment Analysis Techniques and Applications,” IOP Conf. Ser. Mater. Sci. Eng., vol. 551, no. 1, 2019, doi: 10.1088/1757-899X/551/1/012070.

[6] S. R. Ahmad, A. A. Bakar, and M. R. Yaakub, “Ant colony optimization for text feature selection in sentiment analysis,” Intell. Data Anal., vol. 23, no. 1, pp. 133–158, 2019, doi: 10.3233/IDA-173740.

[7] B. Liu, "Sentiment Analysis and Opinion Mining," Morgan Claypool Publ., no. May, pp. 1–108, 2012, doi: 10.2200/S00416ED1V01Y201204HLT016.

[8] M. T. Khan, M. Durrani, A. Ali, I. Inayat, S. Khalid, and K. H. Khan, "Sentiment analysis and the complex natural language," Complex Adapt. Syst. Model., vol. 4, no. 1, 2016, doi: 10.1186/s40294-016-0016-9.

[9] M. R. Yaakub, Y. Li, and J. Zhang, "Integration of Sentiment Analysis into Customer Relational Model: The Importance of Feature Ontology and Synonym," Procedia Technol., vol. 11, pp. 495–501, Jan. 2013, doi: 10.1016/J.PROTCY.2013.12.220.

[10] I. S. Ahmad, A. A. Bakar, M. R. Yaakub, and M. Darwich, "Beyond Sentiment Classification : A Novel Approach for Utilizing Social Media Data for Business Intelligence," Int. J. Adv. Comput. Sci. Appl., vol. 11, no. 3, pp. 437–441, 2020.

[11] M. K. Sohrabi and F. Hemmatian, "An efficient preprocessing method for supervised sentiment analysis by converting sentences to numerical vectors: a twitter case study," Multimed. Tools Appl., 2019, doi: 10.1007/s11042-019-7586-4.

[12] J. Awwalu, A. A. Bakar, and M. R. Yaakub, "Hybrid N-gram model using Naïve Bayes for classification of political sentiments on Twitter," Neural Comput. Appl., 2019, doi: 10.1007/s00521-019-04248-z.

[13] S. R. Ahmad, A. A. Bakar, and M. R. Yaakub, "A review of feature selection techniques in sentiment analysis," Intell. Data Anal., vol. 23, no. 1, pp. 159–189, 2019, doi: 10.3233/IDA-173763.

[14] A. Unler and A. Murat, "A discrete particle swarm optimization method for feature selection in binary classification problems," Eur. J. Oper. Res., vol. 206, no. 3, pp. 528–539, Nov. 2010, doi: 10.1016/j.ejor.2010.02.032.

[15] K. Ganesan, "Text Preprocessing for Machine Learning & NLP," Kavita Ganesan - Build Beautiful NLP & Data Applications., 2019. [Online]. Available: https://kavita-ganesan.com/text-preprocessing-tutorial/#.XrcyMxMzab8. [Accessed: 12-May-2020].

[16] K. Akshi and K. Renu, "Firefly Algorithm for Feature Selection in Sentiment Analysis," Comput. Intell. Data Mining, Adv. Intell. Syst. Comput., vol. 556, pp. 693–703, 2017, doi: 10.1007/978-981-10-3874-7.

[17] M. Tubishat, M. A. M. Abushariah, N. Idris, and I. Aljarah, "Improved whale optimization algorithm for feature selection in Arabic sentiment analysis," Appl. Intell., vol. 49, no. 5, pp. 1688–1707, 2019, doi: 10.1007/s10489-018-1334-8.

[18] A. Krouska, C. Troussas, and M. Virvou, "The effect of preprocessing techniques on Twitter sentiment analysis," IISA 2016 - 7th Int. Conf. Information, Intell. Syst. Appl., 2016, doi: 10.1109/IISA.2016.7785373.

[19] W. Shang, H. Huang, H. Zhu, Y. Lin, Y. Qu, and Z. Wang, "A novel feature selection algorithm for text categorization," doi: 10.1016/j.eswa.2006.04.001.

[20] Z. Zhai, B. Liu, H. Xu, and P. Jia, "Constrained LDA for grouping product features in opinion mining," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2011, vol. 6634 LNAI, no. PART 1, pp. 448–459, doi: 10.1007/978-3-642-20841-6_37.

[21] A. Bagheri, M. Saraee, and F. De Jong, "Care more about customers: Unsupervised domain-independent aspect detection for sentiment analysis of customer reviews," Knowledge-Based Syst., vol. 52, pp. 201–213, Nov. 2013, doi: 10.1016/j.knosys.2013.08.011.

[22] Z. A. A. Alyasseri, A. T. Khader, M. A. Al-Betar, M. A. Awadallah, and X. S. Yang, "Variants of the flower pollination algorithm: A review," Stud. Comput. Intell., vol. 744, pp. 91–118, 2018, doi: 10.1007/978-3-319-67669-2_5.

[23] M. H. Aghdam, N. Ghasem-Aghaee, and M. E. Basiri, "Text feature selection using ant colony optimization," Expert Syst. Appl., vol. 36, no. 3 PART 2, pp. 6843–6853, Apr. 2009, doi: 10.1016/j.eswa.2008.08.022.

[24] M. H. Aghdam and S. Heidari, "Feature Selection Using Particle Swarm Optimization in Text Categorization," J. Artif. Intell. Soft Comput. Res., vol. 5, no. 4, pp. 231–238, Oct. 2015, doi: 10.1515/jaiscr-2015-0031.

[25] S. Nemati, M. E. Basiri, N. Ghasem-Aghaee, and M. H. Aghdam, "A novel ACO-GA hybrid algorithm for feature selection in protein function prediction," Expert Syst. Appl., vol. 36, no. 10, pp. 12086–12094, Dec. 2009, doi: 10.1016/j.eswa.2009.04.023.

[26] Z. Abdi, A. Alyasseri, A. T. Khader, M. A. Al-betar, M. A. Awadallah, and X. Yang, "Nature-Inspired Algorithms and Applied Optimization," vol. 744, no. October 2017, 2018, doi: 10.1007/978-3-319-67669-2.

[27] J. Too and S. Mirjalili, "A Hyper Learning Binary Dragonfly Algorithm for Feature Selection: A COVID-19 Case Study," Knowledge-Based Syst., vol. 212, p. 106553, Jan. 2021, doi: 10.1016/j.knosys.2020.106553.

[28] B. Seerat and F. Azam, "Opinion Mining: Issues and Challenges (A survey)," Int. J. Comput. Appl., vol. 49, no. 9, pp. 975–8887, 2012, doi: 10.5120/7658-0762.

[29] U. Pervaiz, S. Khawaldeh, T. A. Aleef, V. H. Minh, and Y. B. Hagos, "Activity monitoring and meal tracking for cardiac rehabilitation patients," Int. J. Med. Eng. Inform., vol. 10, no. 3, pp. 252–264, 2018, doi: 10.1504/IJMEI.2018.093365.

[30] S. Asghari and N. J. Navimipour, "Review and Comparison of Meta-Heuristic Algorithms for Service Composition in Cloud Computing," 2015.

[31] B. O. Alijla, C. P. Lim, L. P. Wong, A. T. Khader, and M. A. Al-Betar, "An ensemble of intelligent water drop algorithm for feature selection optimization problem," Appl. Soft Comput. J., vol. 65, pp. 531–541, Apr. 2018, doi: 10.1016/j.asoc.2018.02.003.

[32] A. S. Ghareb, A. A. Bakar, and A. R. Hamdan, "Hybrid feature selection based on enhanced genetic algorithm for text categorization," Expert Syst. Appl., vol. 49, pp. 31–47, May 2016, doi: 10.1016/j.eswa.2015.12.004.

[33] X. Yang, "Flower Pollination Algorithm for Global Optimization," pp. 240–249, 2012.

[34] P. H. P. Sarijiya and T. A. Saputra, "Modified Flower Pollination Algorithm for Non smooth and Multiple Fuel Options Economic Dispatch," in 2016 8th International Conference on Information Technology and Electrical Engineering (ICITEE), 2016, pp. 346–350.

[35] F. P. Sakti, S. Sarjiya, and S. P. Hadi, "Optimal Power Flow Using Flower Pollination Algorithm: A Case Study of 500 kV Java-Bali Power System," IJITEE (International J. Inf. Technol. Electr. Eng., vol. 1, no. 2, pp. 45–50, Sep. 2017, doi: 10.22146/ijitee.28363.

[36] D. Rodrigues, G. F. A. Silva, J. P. Papa, A. N. Marana, and X.-S. Yang, "EEG-based person identification through Binary Flower Pollination Algorithm," Expert Syst. Appl., vol. 62, pp. 81–90, Nov. 2016, doi: 10.1016/j.eswa.2016.06.006.

[37] Z. A. E. M. Dahi, C. Mezioud, and A. Draa, "On the efficiency of the binary flower pollination algorithm: Application on the antenna positioning problem," Appl. Soft Comput. J., vol. 47, pp. 395–414, Oct. 2016, doi: 10.1016/j.asoc.2016.05.051.

[38] P. Agarwal and S. Mehta, "Enhanced flower pollination algorithm on data clustering," Int. J. Comput. Appl., vol. 38, no. 2–3, pp. 144–155, 2016, doi: 10.1080/1206212X.2016.1224401.

[39] R. Douglas, X.-S. Yang, A. N. de Souza, and J. P. Papa, "Binary Flower Pollination Algorithm and Its Application," in Recent Advances in Swarm Intelligence and Evolutionary Computation, no. April 2016, 2015, p. 303.

[40] S. A.-F. Sayed, E. Nabil, and A. Badr, "A binary clonal flower pollination algorithm for feature selection," Pattern Recognit. Lett., vol. 77, pp. 21–27, 2016, doi: 10.1016/j.patrec.2016.03.014.

[41] H. M. Zawbaa, A. E. Hassanien, E. Emary, W. Yamany, and B. Parv, "Hybrid flower pollination algorithm with rough sets for feature selection," 2015 11th Int. Comput. Eng. Conf. Today Inf. Soc. What's Next?, ICENCO 2015, pp. 278–283, 2016, doi: 10.1109/ICENCO.2015.7416362.

[42] H. Majidpour and F. G. Soleimanian, "An Improved Flower Pollination Algorithm …," Sari Branch, Islamic Azad University, Feb. 2018.

[43] B. Glover, Understanding Flowers and Flowering : An integrated approach. Oxford University Press, 2008.

[44] M. L. Kawasaki and A. D. Bell, "Plant Form. An Illustrated Guide to Flowering Plant Morphology.," Brittonia, vol. 43, no. 3, p. 145, Jul. 1991, doi: 10.2307/2807042.

[45] F. B. Ozsoydan and A. Baykasoglu, "Analysing the effects of various switching probability characteristics in flower pollination algorithm for solving unconstrained function minimization problems," Neural

Comput. Appl., vol. 31, no. 11, pp. 7805–7819, Nov. 2019, doi: 10.1007/s00521-018-3602-2.

[46] I. Pavlyukevich, "Lévy flights, non-local search and simulated annealing," J. Comput. Phys., vol. 226, no. 2, pp. 1830–1844, Oct. 2007, doi: 10.1016/j.jcp.2007.06.008.

[47] X. S. Yang, M. Karamanoglu, and X. He, "Flower pollination algorithm: A novel approach for multiobjective optimization," Eng. Optim., vol. 46, no. 9, pp. 1222–1237, 2014, doi: 10.1080/0305215X.2013.832237.

[48] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, Classification and regression trees. CRC Press, 2017.

[49] A. Kumar, A. Jaiswal, S. Garg, S. Verma, and S. Kumar, "Sentiment Analysis Using Cuckoo Search for Optimized Feature Selection on Kaggle Tweets," Int. J. Inf. Retr. Res., vol. 9, no. 1, pp. 1–15, 2018, doi: 10.4018/ijirr.2019010101.