Feature Selection using Particle Swarm Optimization for Sentiment Analysis of Drug Reviews

Afifah Mohd Asri[1], Siti Rohaidah Ahmad[2], Nurhafizah Moziyana Mohd Yusop[3]
Department of Computer Science-Faculty of Defence, Science and Technology,
Universiti Pertahanan Nasional Malaysia, Sungai Besi, Kuala Lumpur, Malaysia

***Abstract*—Feature selection (FS) is an essential classification pre-processing task that eliminates irrelevant, redundant, and noisy features. The primary benefits of performing this task include enhanced model performance, reduced computational expense, and modified "curse of dimensionality". The goal of performing FS is to find the best feature group that can be used to build an effective pattern recognition model. Drug reviews play a significant role in delivering valuable medical care information, such as the efficacy, side effects, and symptoms of drug use, facilities, drug pricing, and personal drug usage experience to healthcare providers and patients. FS can be used to obtain relevant and valuable information that can produce an optimal subset of features to help obtain accurate results in the classification of drug reviews. The FS approach reduces the number of input variables by eliminating redundant or irrelevant features and narrowing the collection of features to those most significant to the machine learning model. However, the high dimensionality of the feature vector is a major issue that reduces the accuracy of sentiment classification, making it challenging to find the best feature subset. Thus, this article presents a perceptive method to perform FS by gathering information from the potential solutions generated by a particle swarm optimization (PSO) algorithm. This research aimed to apply this algorithm to identify the optimal feature subset of drug reviews to improve the classification accuracy of sentiment analysis. The experimental results showed that PSO provided a better classification performance than a genetic algorithm (GA) and ant colony optimization (ACO) in most datasets. The results showed that PSO demonstrated the highest levels of performance, with an average of 49.3% for precision, 73.6% for recall, 59% for F-score, and 57.2% for accuracy.**

Keywords—Sentiment analysis; feature selection; particle swarm optimization; drug reviews

I. Introduction

As the world becomes increasingly digitised, there is a growing need for the automation of various processes. One area that has experienced significant expansion is sentiment analysis, which is a crucial component of natural language processing. Sentiment analysis involves identifying the emotional responses elicited by a given text. One area where sentiment analysis has proven useful is in the analysis of drug reviews [1]. Sentiment analysis can identify the overall sentiment of a drug review, which can help healthcare providers understand the effectiveness and side effects of a given drug. However, a significant challenge in sentiment analysis is figuring out which features, or aspects of a text are essential in determining the sentiment. Feature selection (FS) involves identifying a relevant subset of features from a larger set of potential features [2], [3]. By decreasing the number of features, FS can enhance the accuracy and efficiency of a sentiment analysis model [4], [5].

Particle Swarm Optimization (PSO) is a well-known optimization algorithm commonly used in FS due to its ability to simulate the behaviour of a swarm of particles. Each particle represents a possible solution to an optimization problem, and they interact with each other as they move through a solution space in search of the optimal solution [6], [7]. PSO is a suitable method for feature selection in sentiment analysis, as it can determine the optimal feature subset to maximise the model's accuracy [8]. The features to be selected must be encoded as binary values, with 1 indicating their inclusion in the subset and 0 indicating their exclusion.

The PSO algorithm commences the optimization process by randomly assigning a set of particles to represent feasible feature subsets to form a swarm. The fitness value of each particle is then calculated by evaluating the accuracy of the sentiment analysis model with its corresponding feature set. Subsequently, the particles navigate through the solution space, influenced by their individual best position (pbest) and the best position of the swarm (gbest). This iterative process persists until a pre-defined termination criterion is reached, such as reaching a maximum number of iterations, or a specific level of accuracy [6], [9]. Once the PSO algorithm has identified the optimal subset of features, the sentiment analysis model can be trained using only these features, which would result in a more efficient and accurate model. Thus, utilising PSO for FS in a sentiment analysis of drug reviews can enhance the accuracy and efficiency of the model, ultimately leading to a more comprehensive understanding of the effectiveness and side effects of reviewed drugs.

This study aimed to utilise PSO as part of a feature selection method to identify an optimal set of features in drug review datasets that can enhance the classification accuracy. The use of PSO for FS has been demonstrated to be effective in identifying relevant and non-redundant subsets of features that can enhance the performance of machine learning algorithms [10], [11]. The results of these studies showed that PSO outperformed two algorithms by demonstrating the highest level of performance. The application of PSO in drug review datasets can uncover feature subsets that can enhance the accuracy of sentiment analysis models. This achievement can ultimately lead to a deeper comprehension of the efficacy and adverse reactions of various drugs. The experimental results indicated that PSO possessed the capability to produce high-quality feature subsets, which increased the classification accuracy of the sentiment analysis model.

This article begins with an introduction to the background of this study, leading to a clear overview of the conducted experiments. The main body of this article is divided into several sections, with each section focusing on a specific topic, such as literature reviews, related works, brief explanation on PSO, research methodology, experimental setup, and analysis of results. Within each section, subtopics are introduced to offer explanations, support evidence, examples, and analysis. The conclusion section summarises the main points, restates the experiments, and ends with a concluding statement or call to action. The reference section is included to list the sources used in this article.

II. Literature Work

A. Sentiment Analysis

Sentiments can be defined as an individual's attitude or belief that are often influenced by emotions rather than logical reasoning. Therefore, the process of analysing opinions is also commonly known as sentiment analysis, which aims to extract subjective information from various sources, such as speech, text, tweets, and databases [12]. Sentiment analysis involves using Natural Language Processing (NLP) techniques to automatically detect emotions, perspectives, opinions, and attitudes that are expressed in texts and to categorise them either as neutral, positive, or negative [13]. According to [14], sentiment analysis can be applied to social media and medical records to acquire information on the effectiveness of medical treatments or medications. By analysing public archives and social media posts, specific adverse effects of drugs can be identified more efficiently, leading to potential benefits for the pharmaceutical industry in terms of pharmacovigilance [15]. Thus, sentiment analysis can be used to extract information that can assist in making accurate judgments regarding public health and substance safety. Machine Learning and Lexicon-based approaches are commonly used in sentiment analysis.

Lexicon-based methods primarily depend on a sentiment vocabulary, or a set of well-known and precompiled emotional words, sentences, and idioms, which is created for conventional communication categories [16]. The Machine Learning (ML) approach relies on the quantity of labelled data marked by specialists to train a classification [2]. ML can be divided into two categories, namely, supervised and unsupervised methods. An appropriate collection of features must be selected and retrieved to identify sentiments in order to improve the performance of these approaches [17]. The analysis of sentiments using machine learning was the focus of this study.

B. Drug Reviews

In recent years, there has been a significant surge in the number of social networks dedicated to discussing health-related matters. This has led to a substantial increase in the availability of healthcare information in the form of drug reviews on the Internet [18]. According to [15], Drugs.com is the biggest and most viewed medicinal information website that serves customers among the public and healthcare experts. According to [18], a substantial number of drug reviews has been created and published through various healthcare and drug-related online platforms, including AskaPatient and Drugratingz, where users express their first-hand experiences and reactions to various medications.

Drug reviews are written evaluations of numerous drugs by users based on their personal experiences and preferences [19]. Drug reviews can also be described as an individual's impressions regarding several drug-related fields, such as efficacy, adverse effects, convenience, and value [14]. These reviews offer abundant data that can be utilised to make informed decisions concerning public health and medication safety [20].

III. Related Works

This section presents related studies on sentiment analysis of drug reviews. Table I shows an extended summary of related studies based on the techniques used in the sentiment analysis of drug reviews [20].

TABLE I. Related Studies On Techniques Used in Sentiment Analysis of Drug Reviews

Author	**Feature Extraction**
[45]	n-gram, word cloud
[44]	Count Vectorizer, Term frequency-inverse document frequency (TF-IDF)
[43]	Bag of words (BoWs), Term frequency-inverse document frequency (TF-IDF)
[42]	Word position encoding and embedding in vector representation
[19]	Not stated in paper.
[41]	Not stated in paper.
[40]	Yes, but did not state the specific type of feature
[39]	Not stated in paper.
[38]	Not stated in paper.
[37]	Bags of words (BoW), or term frequency-inverse document frequency (TF-IDF)
[36]	Part of speech tagging, n-gram, content words, function words
[35]	Bag of Word (BOW), Part of Speech (POS) tagging, Semantic group and descriptive words.
[34]	Not stated in paper

Feature Selection	Classification Technique	Evaluation
Not stated in paper.	LightGBM, XGBoost, and the CatBoost	Precision, recall, F1-score and support
Not stated in paper.	Artificial Neural Networks (ANN), Recurrent Neural Networks with Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU), Support Vector Machines (SVM), Logistic Regression (LR) and also Random Forests (RF)	Precision, Accuracy, Recall and F1-score
Glowworm Swarm Op timization + ML	Naïve Bayes (NB), K-Nearest Neighbour (KNN) and Support Vector Machine (SVM)	Precision, Accuracy, Recall and F1_Score
Not stated in paper.	SVM, RF, Naive Bayes, and RBFN	Recall, Precision, and F-Score
Not stated in paper.	CNN-NB, 3CRNN-NB, and GRU-NB. 3W3DT-DT, 3W3DT-NB, 3W3DT-KNN, and 3W3DT-RF	Recall, Precision, and F-Score
Not stated in paper.	Convolutional neural network (CNN), Weakly supervised model (WSM) and bidirectional long short-term memory (Bi-LSTM)	Precision, Accuracy, and F1-Score
Not stated in paper.	Probabilistic aspect mining model (PAMM)	Accuracy and Mean Pointwise Mutual Information (PMI)
Not stated in paper.	Radial basis function neural networks (RFN) Probabilistic neural network (PNN)	Recall, Precision, and F-Score
Not stated in paper.	Rule-based Linguistic	Precision, Accuracy, Recall and F1_Score
Fuzzy-Rough Quick Reduction (FRQR).	Random forest, Ripper, Decision tree and Naive Bayes	Accuracy, Performance of running independent hold-out test, time required to develop the model.
Pethe nguin Search Optimization (PeSOA)	Naïve Bayes (NB), K-Nearest Neighbour (KNN), support vector machine (SVM)	Precision, Accuracy, Recall and F1_Measure
Related Words (RW) Vector	Bidirectional Encoder Representation from Transformer (BERT)	Recall, Precision, and F-Measure
Not stated in paper	Bidirectional Encoder Representation from Transformer (BERT) with Long Short-Term Memory (LSTM)	Accuracy and F1_Score

IV. PARTICLE SWARM OPTIMIZATION

The study [21] first proposed Particle Swarm Optimization (PSO) in 1995. It is a stochastic optimization method for a population that takes its cue from the cooperative nature of flocking birds [22]. As stated by [6], PSO is an algorithm for population-based optimization that simulates the social interaction and communication among groups of animals, such as flocks of birds or schools of fish [39].

In PSO, each particle represents a candidate solution to the optimization problem. This set of particles will move through a search space in search of the optimal solution. Each particle has a position and a velocity, and the goal is to find the optimal position in the search space that minimises or maximises an objective function [23].

At each iteration of the algorithm, the particles will adjust their position and velocity based on their own best-known position (i.e., the best solution they have found so far) and the best-known position of the entire swarm (i.e., the best solution found by any particle in the swarm). This is done using a set of mathematical equations that determine each particle's new position and velocity. This process will continue until a stopping criterion is met, such as the maximum number of iterations or a satisfactory level of convergence. The final position of the particles represents the optimal solution to the optimization problem [23].

The PSO version employed in this study is represented by the equation in Fig. 1. Learning rates c1 and c2 are positive constants, r1 and r2 are random numbers in the range of 0 to 1, and Pi, Xi, and Vi represent the swarm's best previous location, the particle's current position, and the rate of change of position (velocity), respectively, for a D-dimensional problem space at iteration t. Additionally, w indicates the inertia weight, g represents the best particle among all particles in the population, and Vmax is the maximum velocity for the particles in each dimension. Vmax is applied, if the total number of elements on the right side of the equation exceeds a predetermined fixed value.

$$V_i(t+1) = w_i V_j(t) + c_1 .\, r_1(t).\, |P_i(t) - X_i(t)| + c_2 .\, r$$

$$X_i(t+1) = X_i(t) + V_i(t+1)$$

Fig. 1. Particle swarm optimization equation.

V. Methodology

This section presents a summary of the methodology used to conduct sentiment analysis of drug reviews using PSO for feature selection, as illustrated in Fig. 2.

A. Data Cleaning and Pre-processing

Prior to analysing the collected text data, data cleaning and text pre-processing steps were implemented to obtain satisfactory results. Data pre-processing involved the removal of words that did not help convey the meaning of the text, or were deemed unnecessary, thereby, enhancing the efficacy of the search for valuable words in each sentence [4]. To create a standardised dataset, several pre-processing techniques were applied, including correcting misspelled words, tokenisation, removing stop words, segmentation, and part-of-speech tagging.

Datasets to be analysed need to undergo pre-processing because reviews are typically written by non-experts in the field of language and could contain various errors, such as spelling mistakes, incorrect use of capital letters, punctuation errors, and grammatical errors. These errors can negatively impact the accuracy and effectiveness of sentiment analysis models, which rely on clean and standardised datasets. By applying pre-processing techniques, the datasets can be cleaned and standardised to make them more suitable for sentiment analysis [4], [24].

In this study, drug review datasets were cleaned and pre-processed by excluding all blank reviews, HTML tags, and spacing. Capitalisation of words and punctuations were verified, words were checked for errors, such as grammar and spelling, and paragraphs were segmented into sentences. This process was performed using Microsoft Word and Microsoft Excel software. Fig. 3 shows an example of the reviews before and after undergoing the cleaning and pre-processing process.

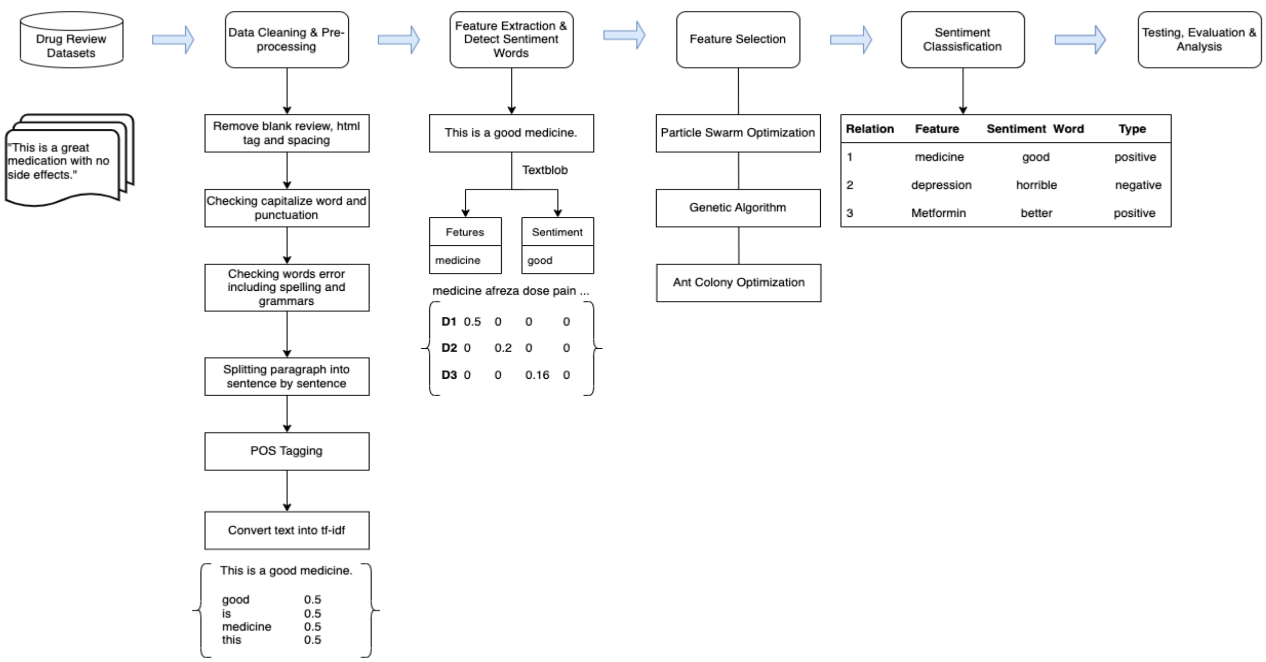

Fig. 2. The methodology of the study.

Sentence before preprocessing:

"I've experienced NO Side Efffects."

Sentence after preprocessing:

"I have experienced no side effect."

Fig. 3. An example of a cleaned and pre-processed review.

The figure shows a sentence that contains HTML code syntax “',” which represents the apostrophe symbol. This syntax was removed and replaced with the full-form word. There were also capitalisation errors, such as “NO” and “Side Effect,” as well as a spelling error in the word “Efffect,” with an extra “f”. After completing the cleaning process, part-of-speech tagging was conducted on each sentence to identify the features and sentiment words in the text.

B. Feature Extraction and Sentiment Word Identification

The identification of pertinent information from textual data through feature extraction is of utmost importance in sentiment classification, as it can inevitably affect the model’s performance [25]. The extraction of important terms from a dataset is known as feature extraction [24]. This approach seeks to select significant data that encompass the most important features of the text [25]. This study employed the TextBlob library, which is a Python-based tool that utilises the Natural Language Toolkit (NLTK) to conduct Natural Language Processing (NLP) tasks and to evaluate sentiment polarity. The library’s noun phrase extraction capabilities and sentiment analyser were used to extract all feature and sentiment words from the datasets.

C. Feature Selection

The first step in constructing a classification model is to detect relevant features within the dataset [2], [21]–[23]. The extracted features would then undergo data transformation from text data into feature vectors. Sentiment analysis is utilised for its emphasis on text documents, since a classifier cannot understand or interpret a text directly.

Consequently, it is necessary to transform textual documents into a format recognisable by a computer, with the Vector Space Model being the preferred technique to accomplish this. Term frequency-inverse document frequency (TF-IDF) was used in this study, as it is one of the most basic methods of conveying features through term count [26]. This model employed a vector space representation of the documents in the dataset, wherein the dimensions of the vector corresponded to the features selected from the dataset [27]. The data objects can be expressed as feature vectors in the feature space. Feature selection will then identify and remove non-essential features from a set of features to improve the classification accuracy, while reducing the feature space dimensionality [28]. In this study, PSO was chosen as a feature selection method to obtain the optimum subset of features.

D. Sentiment Classification

Sentiment classification is a crucial subfield of sentiment analysis, which entails the identification and categorisation of emotions conveyed in a textual data, including reviews and tweets. The primary objective of sentiment classification is to classify texts into various categories based on their subjective information, which typically include positive and negative sentiments [29]. In this study, feature and sentiment words were obtained feature extraction and sentiment word identification section. They underwent manual reviews in conjunction with the datasets to ascertain their conformity with the contextual information presented in the data collection.

E. Testing and Evaluation

The testing and evaluation procedure involved measuring how well the feature and sentiment words were related to one another. This procedure was performed to evaluate the ability of the proposed algorithm to detect and acquire dependable features along with correct sentiments. The efficacy of the proposed method is evaluated based on the following four metrics: precision, recall, F1-score, and accuracy using the measures shown in Table II:

- True Positive (TP) denotes correctly identified positive reviews.
- True Negative (TN) denotes correctly identified negative reviews.
- False Positive (FP) denotes positive reviews incorrectly identified as negative.
- False Negative (FN) denotes negative reviews incorrectly identified as positive.

TABLE II. CONFUSION MATRIX

	Predicted Positive	**Predicted Negative**
Actual Positive	TP	FN
Actual Negative	FP	TN

$$\text{Precision} = TP \div (TP + FP) \quad (1)$$

$$\text{Recall} = TP \div (TP + FN) \quad (2)$$

$$F1 = (2 \times \text{Precision} \times \text{Recall}) \div (\text{Precision} + \text{Recall}) \quad (3)$$

$$\text{Accuracy} = (TP + TN) \div (TP + TN + FP + FN) \quad (4)$$

VI. EXPERIMENTS

This section summarises the drug review datasets obtained from the University of California Irvine (UCI) machine learning repository, along with the methodology employed for feature selection. Specifically, this study has introduced PSO as the primary feature selection technique, with Ant Colony Optimization (ACO) and Genetic Algorithm (GA) as the benchmarks.

A. Datasets

In this study, 1,229 drug reviews from druglib.com were analysed for several conditions, such as diabetes, high blood pressure, heart attack, etc. The datasets contained patient feedback on medications, information on associated conditions, and a patient rating of 10 stars, which reflected the overall patient happiness. Data were gathered by crawling the pharmaceutical review website, Drugs.com. The datasets utilised in this study were stratified into two distinct polarity levels based on review ratings, namely, positive (class 1, rating ≥ 5) and negative (class -1, rating ≤ 5), as previously defined in [30]. These datasets comprised a total of 719 positive reviews and 509 negative reviews, representing a valuable resource for training and evaluating sentiment classification models. These datasets were then divided into 10 sets. The first set includes 10% of the overall features, while the second set contains 20% and so on, as shown in Table III. Using a stratified random sampling method, 70% and 30% of the datasets were put aside for training and testing.

The distribution of the two classes is illustrated in Fig. 4, with approximately 41% of the reviews having a negative polarity (class -1) and 59% having a positive polarity (class 1).

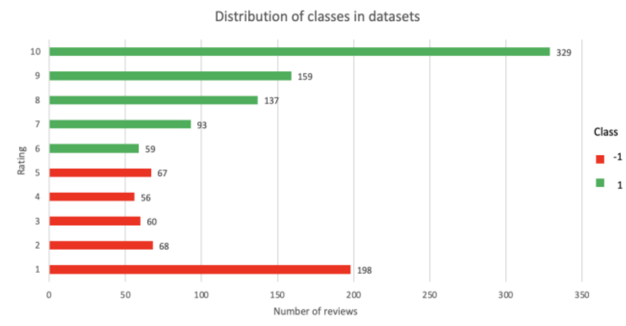

Fig. 4. Class distribution of drug review datasets.

B. Baseline Algorithms

Due to their proven effectiveness in improving the performance of sentiment analysis [31]–[33], the ACO and GA were chosen as the baseline algorithms for the PSO algorithm in this study. The aim was to compare the performance of PSO against these algorithms in identifying an optimal feature subset for the sentiment analysis of drug reviews.

TABLE III. SUMMARY OF TOTAL FEATURES AND REVIEWED SENTENCES IN EACH DRUG REVIEW DATASET

Dataset	Total Features	Total of Reviewed Sentences
Set 1	83	398
Set 2	167	583
Set 3	250	762
Set 4	334	923
Set 5	417	1030
Set 6	501	1135
Set 7	584	1231
Set 8	668	1329
Set 9	751	1417
Set 10	835	1506

C. Experimental Setups

The aim of the present study was to evaluate the efficacy of the proposed algorithm through a series of experiments designed to compare its performance with existing algorithms in the field. The experiments were conducted on a computer system equipped with an Apple M2 processor and 16 GB of RAM, which provided sufficient computational power and memory capacity to perform the required tasks. The high processing power and ample memory capacity of this machine facilitated the efficient execution of experiments and reliable collection of data. The proposed PSO algorithm was implemented using Python and was run on the Jupyter Notebook, while the ACO and GA algorithms were executed using an established algorithm in WEKA on the same machine.

TABLE IV. PSO PARAMETER SETTING

Parameter	Value
Acceleration Coefficients, c1 & c2	2
Inertial weight, w	1
Population Size	20
Iteration	20
Fitness	1

Table IV lists the PSO parameters that have been employed to identify the optimal parameter configuration that would yield a sentiment analysis model with superior performance based on the test set. The testing was performed to evaluate the ability of the PSO algorithm to produce a feature subset that was both significant and non-redundant, aiming to determine the most optimal feature set. The effectiveness of the proposed algorithm was measured using several performance metrics, namely, precision, recall, accuracy, and F-score, as elaborated in testing and evaluation section.

VII. ANALYSIS OF RESULTS

This section presents the outcomes of utilising the PSO algorithm for feature selection of a drug review dataset. Its performance was compared with the performance of the ACO and GA. The purpose of this research was to evaluate the effectiveness of PSO in identifying a relevant and non-redundant feature subset that could improve the performance of the sentiment analysis model. To achieve this objective, accuracy, precision, recall, and F-score were used as performance metrics to assess the quality of the outputs generated by each algorithm.

Table V presents the experimental outcomes of the proposed PSO algorithm, along with the compared algorithms (ACO and GA) based on the performance metrics (precision, recall, F-score, and accuracy). These results indicated that PSO outperformed ACO and GA in most of the datasets, with an average of 49.3% for precision, 73.6% for recall, 59% for F-score, and 57.2% for accuracy.

Fig. 5 to Fig. 9 show the performance evaluation of PSO compared to ACO and GA for feature selection in sentiment analysis based on precision, recall, F-score, and accuracy. The results showed that PSO outperformed both ACO and GA in terms of the average precision, recall, F-score, and accuracy. The average precision of PSO was 49.3%, which was higher than ACO with 46.4% and GA with 46.2%. Similarly, the average recall of PSO was 73.6%, which was higher than ACO with 71.5% and GA with 70.5%.

The experimental results listed in Table VI show that PSO has selected a higher number of features on average (197) compared to ACO (111) and GA (112). The table also shows that PSO was more efficient in selecting relevant features, as evidenced by its superior performance in terms of precision, recall, F-score, and accuracy. Therefore, the higher number of selected features by PSO can be attributed to its ability to identify and retain more relevant features, which ultimately improved its classification performance.

TABLE V. THE RESULTS OF ACO, GA, AND PSO BASED ON PERFORMANCE METRICS (PRECISION, RECALL, F-SCORE, AND ACCURACY) ON DRUG REVIEW DATASETS

Dataset	Precision			Recall			F-Score			Accuracy		
	ACO	GA	PSO	ACO	GA	PSO	ACO	GA	PSO	ACO	GA	PSO
Set 1	42.9	46.7	**49.4**	60.0	68.3	**71.7**	50.0	55.4	**58.5**	53.8	53.6	**55.8**
Set 2	**48.7**	43.6	45.3	**75.3**	64.1	73.8	**59.2**	51.9	56.1	**57.9**	51.6	53.4
Set 3	**49.2**	43.8	47.1	**75.3**	68.4	70.2	**59.5**	53.4	56.4	**58.4**	51.2	57.2
Set 4	44.5	49.2	**51.2**	70.1	71.3	**75.7**	54.4	58.2	**61.1**	53.4	54.9	**56.9**
Set 5	48.4	44.5	**48.9**	**74.8**	72.0	72.5	**58.7**	55.0	58.4	54.2	50.8	**56.8**
Set 6	44.5	46.2	**49.8**	72.6	72.2	**75.6**	55.2	56.3	**60.1**	52.4	52.7	**57.7**
Set 7	50.2	**51.8**	51.7	71.6	73.7	**74.1**	59.1	60.8	**60.9**	56.8	58.9	**58.9**
Set 8	45.1	43.2	**52.4**	72.4	68.3	**73.1**	55.5	52.9	**61.0**	54.0	51.9	**58.9**
Set 9	43.3	45.5	**48.2**	70.9	73.6	**74.9**	53.8	56.2	**58.7**	51.3	55.2	**57.6**
Set 10	47.5	47.4	**49.0**	72.4	72.9	**74.8**	57.3	57.5	**59.2**	54.4	54.9	**59.1**
Average	**46.4**	**46.2**	**49.3**	**71.5**	**70.5**	**73.6**	**56.3**	**55.8**	**59.0**	**54.7**	**53.6**	**57.2**

TABLE VI. AVERAGE PERFORMANCE AND SELECTED FEATURES FOR ACO, GA, AND PSO ON DRUG REVIEW DATASETS

Dataset	Precision			Recall			F-Score			Accuracy			Number of features		
	ACO	GA	PSO	ACO	GA	PSO	ACO	GA	PSO	ACO	GA	PSO	ACO	GA	PSO
Set 1	42.9	46.7	49.4	60.0	68.3	71.7	50.0	55.4	58.5	53.8	53.6	55.8	8	9	26
Set 2	48.7	43.6	45.3	75.3	64.1	73.8	59.2	51.9	56.1	57.9	51.6	53.4	30	38	56
Set 3	49.2	43.8	47.1	75.3	68.4	70.2	59.5	53.4	56.4	58.4	51.2	57.2	29	38	97
Set 4	44.5	49.2	51.2	70.1	71.3	75.7	54.4	58.2	61.1	53.4	54.9	56.9	73	83	135
Set 5	48.4	44.5	48.9	74.8	72.0	72.5	58.7	55.0	58.4	54.2	50.8	56.8	86	94	201
Set 6	44.5	46.2	49.8	72.6	72.2	75.6	55.2	56.3	60.1	52.4	52.7	57.7	99	98	244
Set 7	50.2	51.8	51.7	71.6	73.7	74.1	59.1	60.8	60.9	56.8	58.9	58.9	147	169	249
Set 8	45.1	43.2	52.4	72.4	68.3	73.1	55.5	52.9	61.0	54.0	51.9	58.9	228	183	322
Set 9	43.3	45.5	48.2	70.9	73.6	74.9	53.8	56.2	58.7	51.3	55.2	57.6	175	165	343
Set 10	47.5	47.4	49.0	72.4	72.9	74.8	57.3	57.5	59.2	54.4	54.9	59.1	239	246	302
Average	**46.4**	**46.2**	**49.3**	**71.5**	**70.5**	**73.6**	**56.3**	**55.8**	**59.0**	**54.7**	**53.6**	**57.2**	**111**	**112**	**197**

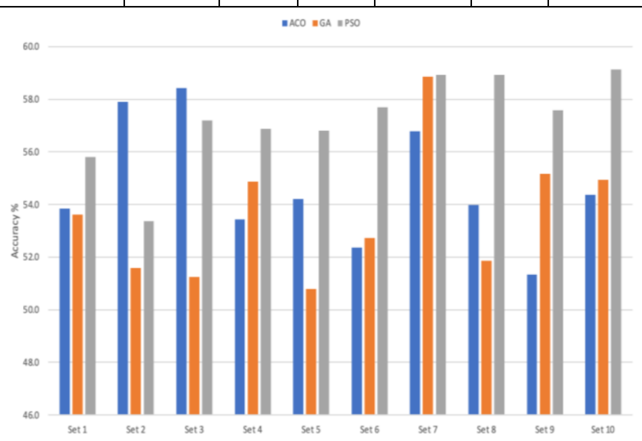

Fig. 5. Precision results for ACO, GA, and PSO on drug review datasets.

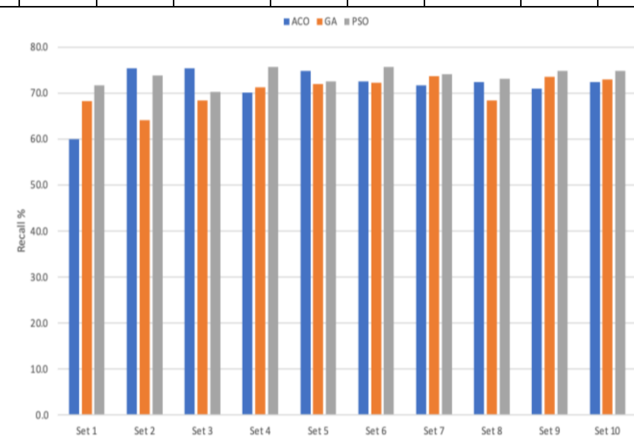

Fig. 6. Recall results for ACO, GA, and PSO on drug review datasets.

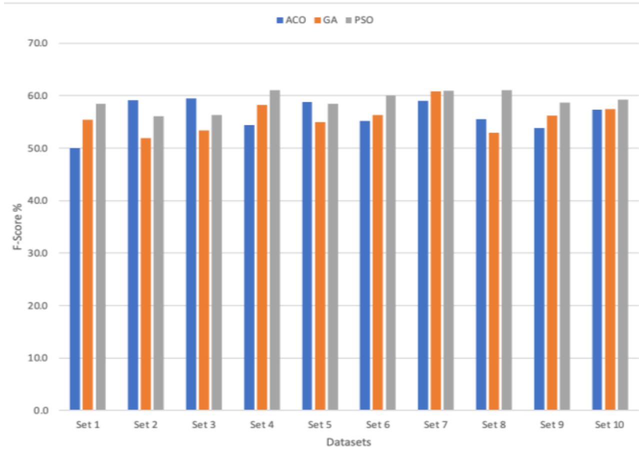

Fig. 7. F-score results for ACO, GA, and PSO on drug review datasets.

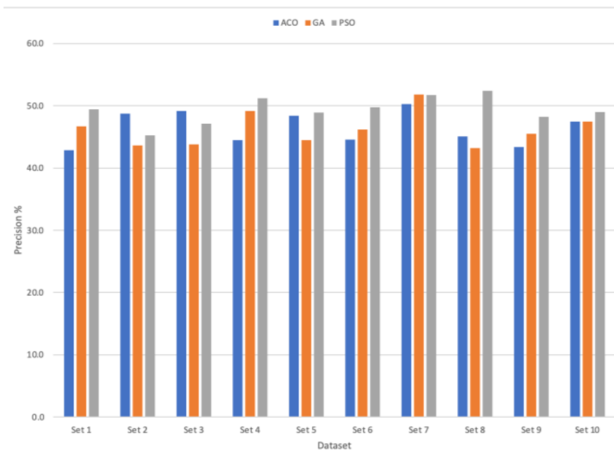

Fig. 8. Accuracy results for ACO, GA, and PSO on drug review datasets.

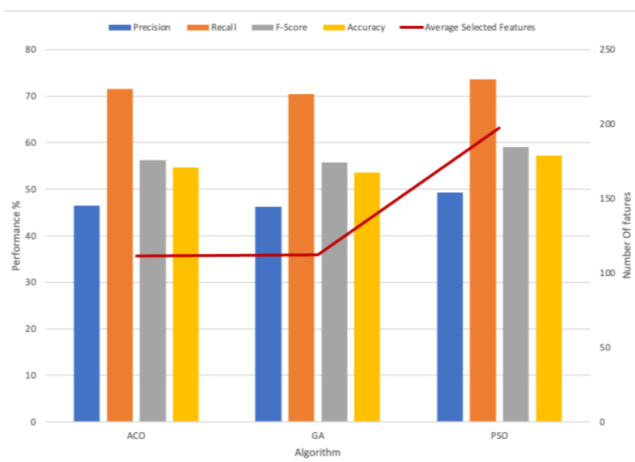

Fig. 9. Average performance and selected features of ACO, GA, and PSO on drug review datasets.

This study has shown that PSO was a highly effective algorithm for feature selection in sentiment analysis. The results showed that PSO outperformed the other two algorithms, demonstrating the highest level of performance. In these experiments, the parameters were set up appropriately to ensure the validity of the results. However, these experiments were not intended to determine the optimality of the selected parameter values for PSO in the feature selection process of sentiment analysis. Consequently, further research is needed to investigate the optimal parameter values for PSO in this specific application domain. Additional experiments are also required to obtain accurate parameter settings for PSO in feature selection for sentiment analysis to enhance its performance and accuracy in this context.

The experimental results indicated that PSO possessed the capability to produce high-quality feature subsets, which increased the classification accuracy of the sentiment analysis model. This algorithm explored the vast search space of possible feature subsets to identify the most relevant and discriminative features, while discarding redundant or irrelevant ones. The global optimization capabilities of PSO ensured that the algorithm was able to converge the optimal solution, which further enhanced the quality of the selected feature subset. Therefore, the empirical evidence has shown that PSO was a highly effective method for feature selection, since it was capable of producing feature subsets that could improve the classification accuracy of a variety of applications.

VIII. Conclusion

This paper presents a novel approach that utilised the PSO algorithm for feature selection in drug reviews, with the objective of improving the performance of the sentiment analysis model. To verify the efficacy of the proposed approach, three feature selection methods were employed to compare the results obtained from drug review datasets. The results showed that PSO provided a better classification performance than GA and ACO in most datasets. PSO was both simple to use and well-known for accelerating convergence, since all particles learned from the global best, which was the current best value achieved by any particle in its neighbourhood. The results have also indicated that these techniques could yield favourable outcomes in acquiring the optimal feature subset. However, in a complex optimization, PSO could suffer from premature convergence. There is a scope for further research into the use of PSO for sentiment analysis in other domains, such as social media and customer reviews. Additionally, the combination of PSO with other optimization techniques, such as genetic algorithms or k-nearest neighbour algorithms could further improve the accuracy of sentiment analysis algorithms.

Acknowledgment

The authors gratefully acknowledge Universiti Pertahanan Nasional Malaysia and the Ministry of Education Malaysia, as well as the Fundamental Research Grant Scheme (FRGS) for supporting this research project through grant number R0144-FRGS/1/2022/ICT06/UPNM/02/1.

References

[1] S. Noferesti and M. Shamsfard, “Resource construction and evaluation for indirect opinion mining of drug reviews,” *PLoS One*, vol. 10, no. 5, pp. 1–25, 2015, doi: 10.1371/journal.pone.0124993.

[2] G. H. John, R. Kohavi, and K. Ppeger, “Irrelevant Features and the Subset Selection Problem,” *Machine Learning: Proceedings Of The Eleventh International Conference*, pp. 121–129, 1994.

[3] A. Dasgupta, A. Banerjee, A. G. Dastidar, A. Barman, and S. Chakraborty, “A Study and Analysis of a Feature Subset Selection

Technique Using Penguin Search Optimization Algorithm," in *Data Science*, CRC Press, 2019, pp. 51–68. doi: 10.1201/9780429263798-3.

[4] M. R. Y. 3 Siti Rohaidah Ahmad 1,*, Azuraliza Abu Bakar 2 and 1, "A review of feature selection techniques in sentiment analysis," *Intelligent Data Analysis*, pp. 159–189, 2019.

[5] S. R. Ahmad, A. A. Bakar, and M. R. Yaakub, "Metaheuristic algorithms for feature selection in sentiment analysis," in *Proceedings of the 2015 Science and Information Conference, SAI 2015*, Institute of Electrical and Electronics Engineers Inc., Sep. 2015, pp. 222–226. doi: 10.1109/SAI.2015.7237148.

[6] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, IEEE, pp. 1942–1948. doi: 10.1109/ICNN.1995.488968.

[7] D. Yazdani, B. Nasiri, A. Sepas-Moghaddam, and M. R. Meybodi, "A novel multi-swarm algorithm for optimization in dynamic environments based on particle swarm optimization," *Applied Soft Computing Journal*, vol. 13, no. 4, pp. 2144–2158, 2013, doi: 10.1016/j.asoc.2012.12.020.

[8] M. Ghosh, R. Guha, I. Alam, P. Lohariwal, D. Jalan, and R. Sarkar, "Binary genetic swarm optimization: A combination of ga and pso for feature selection," *Journal of Intelligent Systems*, vol. 29, no. 1, pp. 1598–1610, Jan. 2020, doi: 10.1515/jisys-2019-0062.

[9] A. de Campos, A. T. R. Pozo, and E. P. Duarte, "Parallel multi-swarm PSO strategies for solving many objective optimization problems," *J Parallel Distrib Comput*, vol. 126, pp. 13–33, Apr. 2019, doi: 10.1016/j.jpdc.2018.11.008.

[10] H. M.Harb and A. S. Desuky, "Feature Selection on Classification of Medical Datasets based on Particle Swarm Optimization," *Int J Comput Appl*, vol. 104, no. 5, pp. 14–17, 2014, doi: 10.5120/18197-9118.

[11] L. Shang, Z. Zhou, and X. Liu, "Particle swarm optimization-based feature selection in sentiment classification," *Soft comput*, vol. 20, no. 10, pp. 3821–3834, 2016, doi: 10.1007/s00500-016-2093-2.

[12] P. Moorthi, "Hybrid optimization for feature selection in opinion mining," 2018. [Online]. Available: www.sciencepubco.com/index.php/IJET.

[13] V. A. Kharde and S. S. Sonawane, "Sentiment Analysis of Twitter Data: A Survey of Techniques," 2016. [Online]. Available: http://ai.stanford.

[14] K. Denecke and Y. Deng, "Sentiment analysis in medical settings: New opportunities and challenges," *Artif Intell Med*, vol. 64, no. 1, pp. 17–27, May 2015, doi: 10.1016/j.artmed.2015.03.006.

[15] D. Cavalcanti and R. Prudêncio, "Aspect-based opinion mining in drug reviews," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10423 LNAI, no. August, pp. 815–827, 2017, doi: 10.1007/978-3-319-65340-2_66.

[16] J. Serrano-Guerrero, J. A. Olivas, F. P. Romero, and E. Herrera-Viedma, "Sentiment analysis: A review and comparative analysis of web services," *Inf Sci (N Y)*, vol. 311, pp. 18–38, Aug. 2015, doi: 10.1016/j.ins.2015.03.040.

[17] P. H. Prastyo, I. Ardiyanto, and R. Hidayat, "A Review of Feature Selection Techniques in Sentiment Analysis Using Filter, Wrapper, or Hybrid Methods," in *Proceedings - 2020 6th International Conference on Science and Technology, ICST 2020*, Institute of Electrical and Electronics Engineers Inc., 2020. doi: 10.1109/ICST50505.2020.9732885.

[18] J. Zheng, P. Wen, X. Ji, X. Lyu, and Y. Yang, "Helpfulness Prediction of Online Drug Reviews," in *2021 IEEE International Conference on Consumer Electronics and Computer Engineering, ICCECE 2021*, Institute of Electrical and Electronics Engineers Inc., Jan. 2021, pp. 528–537. doi: 10.1109/ICCECE51280.2021.9342308.

[19] M. E. Basiri, M. Abdar, M. A. Cifci, S. Nemati, and U. R. Acharya, "A novel method for sentiment classification of drug reviews using fusion of deep and machine learning techniques," *Knowl Based Syst*, vol. 198, Jun. 2020, doi: 10.1016/j.knosys.2020.105949.

[20] S. Rohaidah Ahmad, N. Moziyana Mohd Yusop, A. Mohd Asri, M. Fahmi Muhamad Amran, U. Pertahanan Nasional Malaysia, and S. Besi Kuala Lumpur, "A Review of Feature Selection Algorithms in Sentiment Analysis for Drug Reviews," 2021. [Online]. Available: www.askapatient.com.

[21] R. Eberhart and J. Kennedy, "A New Optimizer Using Particle Swarm Theory."

[22] M. Ghosh, R. Guha, I. Alam, P. Lohariwal, D. Jalan, and R. Sarkar, "Binary genetic swarm optimization: A combination of ga and pso for feature selection," *Journal of Intelligent Systems*, vol. 29, no. 1, pp. 1598–1610, Jan. 2020, doi: 10.1515/jisys-2019-0062.

[23] Chenye Qiu, *A Multi-swarm Particle Swarm Optimization with an Adaptive Regrouping Strategy for Feature Selection*.

[24] S. R. Priya and M. Devapriya, "Twitter Sentiment Analysis with Diabetic Drugs Using Machine Learning Techniques with Glowworm Swarm Optimization Algorithm." [Online]. Available: www.ijert.org.

[25] M. Wankhade, A. C. S. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," *Artif Intell Rev*, 2022, doi: 10.1007/s10462-022-10144-1.

[26] A. Sharma, J. Lyons, A. Dehzangi, and K. K. Paliwal, "A feature extraction technique using bi-gram probabilities of position specific scoring matrix for protein fold recognition," *J Theor Biol*, vol. 320, pp. 41–46, Mar. 2013, doi: 10.1016/j.jtbi.2012.12.008.

[27] S. R. Ahmad, A. Abu Bakar, and M. R. Yaakub, "A review of feature selection techniques in sentiment analysis."

[28] A. Kumar, R. Khorwal, and S. Chaudhary, "A survey on sentiment analysis using swarm intelligence," *Indian J Sci Technol*, vol. 9, no. 39, 2016, doi: 10.17485/ijst/2016/v9i39/100766.

[29] B. Liu, "Sentiment Analysis and Opinion Mining," Morgan & Claypool Publishers, 2012.

[30] F. Gräßer, H. Malberg, S. Kallumadi, and S. Zaunseder, "Aspect-Based sentiment analysis of drug reviews applying cross-Domain and cross-Data learning," *ACM International Conference Proceeding Series*, vol. 2018-April, pp. 121–125, 2018, doi: 10.1145/3194658.3194677.

[31] R. Katarya and A. Yadav, "A comparative study of genetic algorithm in sentiment analysis," in *Proceedings of the 2nd International Conference on Inventive Systems and Control, ICISC 2018*, Institute of Electrical and Electronics Engineers Inc., Jun. 2018, pp. 136–141. doi: 10.1109/ICISC.2018.8399051.

[32] S. R. Ahmad, A. A. Bakar, and M. R. Yaakub, "Ant colony optimization for text feature selection in sentiment analysis," *Intelligent Data Analysis*, vol. 23, no. 1, pp. 133–158, 2019, doi: 10.3233/IDA-173740.

[33] D. Eleyan, "Ant Colony Optimization Based Feature Selection in Rough Set Theory Customer Satisfaction and Awareness, A case study of South African Commercial vs Islamic Banks View project Software engineering View project," 2016. [Online]. Available: https://www.researchgate.net/publication/260986337.

[34] G. Abuka, J. Ranganathan, Z. Dong, and Y. Gu, "Text Summarization and Sentiment Analysis of Drug Reviews: A Transfer Learning Approach," 2023.

[35] M. Imani and S. Noferesti, "Aspect extraction and classification for sentiment analysis in drug reviews," *J Intell Inf Syst*, vol. 59, no. 3, pp. 613–633, Dec. 2022, doi: 10.1007/s10844-022-00712-w.

[36] T. Anuprathibha and C. S. Kanimozhiselvi, "Penguin search optimization based feature selection for automated opinion mining," *International Journal of Recent Technology and Engineering*, vol. 8, no. 3, pp. 648–653, 2019, doi: 10.35940/ijrte.B2629.098319.

[37] T. Chen, P. Su, C. Shang, R. Hill, H. Zhang, and Q. Shen, "Sentiment Classification of Drug Reviews Using Fuzzy-rough Feature Selection," *IEEE International Conference on Fuzzy Systems*, vol. 2019-June, pp. 1–6, 2019, doi: 10.1109/FUZZ-IEEE.2019.8858916.

[38] J. C. Na, W. Y. M. Kyaing, C. S. G. Khoo, S. Foo, Y. K. Chang, and Y. L. Theng, "Sentiment classification of drug reviews using a rule-based linguistic approach," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7634 LNCS, pp. 189–198, 2012, doi: 10.1007/978-3-642-34752-8_25.

[39] V. Gopalakrishnan and C. Ramaswamy, "Patient opinion mining to analyze drugs satisfaction using supervised learning," *Journal of Applied Research and Technology*, vol. 15, no. 4, pp. 311–319, 2017, doi: 10.1016/j.jart.2017.02.005.

[40] P. Gawande and S. Gore, "Extraction of Aspects from Drug Reviews Using Probabilistic Aspect Mining Model," *International Journal of*

Science and Research (IJSR), vol. 4, no. 11, pp. 4–7, 2015, doi: 10.21275/v4i11.nov151059.

[41] Z. Min, "Drugs Reviews Sentiment Analysis using Weakly Supervised Model," *Proceedings of 2019 IEEE International Conference on Artificial Intelligence and Computer Applications, ICAICA 2019*, pp. 332–336, 2019, doi: 10.1109/ICAICA.2019.8873466.

[42] S. Liu and I. Lee, "Extracting features with medical sentiment lexicon and position encoding for drug reviews," *Health Inf Sci Syst*, vol. 7, no. 1, 2019, doi: 10.1007/s13755-019-0072-6.

[43] S. R. Priya, "Twitter Sentiment Analysis with Diabetic Drugs Using Machine Learning Techniques with Glowworm Swarm Optimization Algorithm," vol. 9, no. 07, pp. 62–68, 2020.

[44] S. Vijayaraghavan and D. Basu, "Sentiment Analysis in Drug Reviews using Supervised Machine Learning Algorithms," *ArXiv*, 2020.

[45] Sumit Mishra, "Drug Review Sentiment Analysis using Boosting Algorithm," *International Journal of Trend in Scientific Research and Development (IJTSRD)*, vol. Volume 5, no. Issue 4, May-June 2021, 2021.