*International Journal of Computational and Experimental Science and ENgineering (IJCESEN)*

*Vol. 10-No.2 (2024) pp. 126-133*
*http://www.ijcesen.com*

***ISSN: 2149-9144***



**Review Article**

# A review of metaheuristic optimization techniques in text classification

**Cem KAYA[1,2], Zeynep Hilal KİLİMCİ[3,*], Mitat UYSAL[4], Murat KAYA[5]**

[1] Department of Computer Engineering, Dogus University, 34775, İstanbul-Türkiye
[2] Tübitak Bilgem, , 41470, Kocaeli-Türkiye
**Email:** cem.kaya@tubitak.gov.tr - **ORCID:** 0009-0000-8247-5410

[3] Department of Information Systems Engineering, Kocaeli University, 41000, Kocaeli-Türkiye
* **Corresponding Author Email:** zeynep.kilimci@kocaeli.edu.tr - **ORCID:** 0000-0003-1497-305X

[4] Department of Software Engineering, Dogus University, 34775, İstanbul-Türkiye
**Email:** muysal@dogus.edu.tr - **ORCID:** 0000-0001-9713-2525

[5] Department of Computer Programming, Acıbadem University, 34752, İstanbul-Türkiye
**Email:** murat.kaya@acıbadem.edu.tr

**Article Info:**

**DOI:** 10.22399/ijcesen.295
**Received :** 31 January 2024
**Accepted :** 30 April 2024

**Keywords**

Metaheuristic optimization
Text classification
Feature selection

**Abstract:**

Metaheuristic algorithms, inspired by natural phenomena and human-based strategies, offer versatile approaches to navigate diverse search spaces and adapt to dynamic environments. These algorithms, including evolutionary algorithms, swarm intelligence, bio-inspired methods, human-based approaches, and plant-inspired techniques, have found applications across diverse domains such as engineering, finance, healthcare, logistics, and telecommunications. In the text classification domain, metaheuristic techniques have emerged as powerful tools to enhance the accuracy, efficiency, and robustness of classification systems. By optimizing feature subsets, fine-tuning model parameters, and addressing challenges such as feature selection, dimensionality reduction, class imbalance, and noisy data, metaheuristic algorithms provide flexible solutions that adapt to various text datasets and tasks. This review paper comprehensively explores recent advancements in metaheuristic applications in text classification across six categories. From evolutionary-based methods to swarm-based approaches, bio-inspired techniques to physics/chemistry-based strategies, human-based methods to plant-based algorithms, researchers have leveraged diverse metaheuristic techniques to push the boundaries of text classification. Through a systematic analysis of recent research studies, this review provides insights into the strengths, limitations, and future directions of metaheuristic optimization in the context of text classification.

## 1. Introduction

In recent years, metaheuristic optimization has found applications across diverse domains such as engineering [1], finance [2], healthcare [3], logistics [4], and telecommunications [5], where traditional optimization methods may struggle to find optimal solutions within reasonable timeframes. These techniques have been instrumental in solving real-world problems characterized by non-linearity, high dimensionality, uncertainty, and stochasticity, offering practitioners powerful tools to tackle complex optimization challenges efficiently.

The integration of metaheuristic techniques into the text classification domain has led to significant advancements in areas such as sentiment analysis, document categorization, information retrieval, and topic modeling. These methods, inspired by natural phenomena, human-based strategies, and physics/chemistry principles, offer novel solutions to complex optimization problems inherent in text analysis. From evolutionary algorithms to swarm intelligence, bio-inspired techniques to human-based strategies, and even plant-based approaches, researchers have explored diverse avenues to enhance the accuracy, efficiency, and robustness of text classification systems.

Evolutionary-based metaheuristic methods, such as Genetic Algorithms (GA), have shown promise in optimizing feature selection and model parameters,

leading to significant improvements in classification accuracy. Hybrid models combining GA with deep learning architectures like Convolutional Neural Networks (CNN) have demonstrated remarkable performance, particularly in languages like Arabic, where traditional methods face challenges. Similarly, swarm-based approaches, including Particle Swarm Optimization (PSO) and Intelligent Water Drops (IWD), have proven effective in optimizing feature selection and enhancing sentiment analysis accuracy, especially in social media data. Optimizers (GWO) are among the methods that have shown promise in optimizing text classification tasks, surpassing conventional algorithms in terms of performance and efficiency. Furthermore, physics/chemistry-based approaches like Efficient Binary Black Hole (EBBH) algorithm have demonstrated efficient exploration of feature subsets, leading to competitive performance in text classification tasks.

Bio-inspired metaheuristic techniques draw inspiration from nature to tackle feature selection, dimensionality reduction, and clustering in text data. Ant Colony Optimization (ACO), Social Spider Optimization (SSO), and Grey Wolf Optimization(GWO).

Human-inspired metaheuristic approaches, such as Imperialist Competitive Algorithm (ICA) and Weighted Feature-based ICA (WFICA), leverage human-based strategies to optimize partitioning and feature selection in text data, showcasing superior performance compared to traditional clustering and classification methods. Finally, plant-inspired metaheuristic algorithms like Invasive Weed Optimization (IWO) and Flower Pollination Algorithm (FPA) offer efficient solutions for feature selection and classifier optimization, particularly in handling imbalanced data and enhancing educational and sentiment analysis tasks.

In this comprehensive review, we delve into recent advancements across these 6 diverse categories of metaheuristic applications in text classification. Through systematic analysis and evaluation, we aim to elucidate the contributions, strengths, and potential avenues for future research in this burgeoning field.

## 2. Metaheuristic Techniques in Text Classification

In this section, we explore the application of metaheuristic optimization techniques in the domain of text classification. Metaheuristic algorithms, characterized by their ability to efficiently explore large search spaces and find near-optimal solutions, offer a promising approach to address the complexities inherent in text analysis tasks. By leveraging principles inspired by nature, human behavior, and physics/chemistry, these techniques provide flexible and adaptive solutions that enhance the accuracy, efficiency, and robustness of text classification systems.

### 2.1 Evolutionary-based

In a recent study [6] conducted using an evolutionary-based type of metaheuristic method, the authors propose a hybrid classification model for Arabic text using CNN and GA. The combination of CNN with GA proves effective in enhancing classification accuracy and RMSE for Arabic text. Despite a longer computation time compared to the baseline due to GA execution in training/validation, the benefits are validated with two large datasets. This research claims that GA-CNN demonstrates excellent results, surpassing the baseline and an existing method.

Another recent study [7] introduces MORDC, a multi-objective feature selection method designed for text classification tasks. MORDC addresses the dual objectives of relevancy and redundancy by employing an evolutionary optimization technique. The algorithm utilizes two specific objectives: RDC for scoring the relevancy of text features to the target class and the correlation between selected features. These objectives guide the multi-objective evolutionary process in searching for a set of features with optimal relevancy and minimal redundancy. Comparative experiments involving three benchmark datasets and three well-known classifiers position MORDC as superior to other algorithms, including univariate and multivariate filter algorithms. The study suggests potential enhancements, including the application of local search operators to navigate the extensive solution space more efficiently for improved convergence speed. It also proposes exploring unsupervised metrics for solution evaluation and considers the prospect of incorporating the proposed multivariate criterion into many-objective optimization tasks that involve more than three objectives.

In the context of post-disaster scenarios, utilizing automated methods to filter pertinent disaster-related social media posts proves beneficial for both emergency teams and affected individuals seeking assistance [8]. While ensuring the accuracy of social media post classification is essential, the primary focus lies on prompt response and support, especially challenging when human labeling during

a disaster event is impractical. This research builds on prior domain adaptation approaches by introducing the GA for feature selection, GADA, during domain adaptation. The GADA framework aims to enhance performance and outperforms supervised approaches, yielding notable accuracy improvements. GADA not only reduces experiment training time through robust feature engineering but also exhibits limitations, such as being untested on datasets beyond English language datasets and operating solely in binary classification settings.

Another research [9] that uses Genetic Algorithm outlines a progressive clustering model that integrates an English text classification model based on GA-SVM (GA - Support Vector Machines) and a clustering model using GA-FCM (GA - Fuzzy C-Means). The authors details the GA-FCM algorithm's fundamental principles derived from the GA-SVM algorithm and explore its application in text classifications and diverse clusters. The conducted experiments indicate that the progressive clustering model, combining GA-SVM and GA-FCM, significantly enhances the efficiency and accuracy of English text classification. Furthermore, the combination of the genetic algorithm and support vector machine demonstrates efficacy in reducing classification noise and improving overall classification outcomes. During the grouping stage, the integration of the genetic algorithm and fuzzy grouping method proves effective in preventing convergence to local optima, yielding superior grouping results and determining the most suitable number of groups.

### 2.2 Swarm-based

In this category, a recent article [10] introduces a text sentiment analysis method for e-shopping product reviews using the chaotic Coyote Optimized Deep Belief Network (CCO-DBN) approach, aiming to mitigate challenges in sentiment analysis and improve accuracy. The simulation is implemented on the MATLAB platform. The proposed approach demonstrates higher accuracy compared to existing methods.

In another recent study [11], the author introduces an IWD based feature selection system for sentiment classification using the Turkish Twitter dataset against the ME classifier. Experimental results demonstrate that the IWD-based method selects superior features compared to the ReliefF filter-based method, leading to a significant increase in classification performance. The F-score for the Twitter dataset rises with the IWD-based feature selection, effectively reducing the number of features for high-dimensional data classification. The proposed method not only enhances classification accuracy but also reduces the time needed to classify the test dataset.

In another swarm-based metaheuristic research [12], the authors address the challenge of managing and extracting knowledge from the vast and unstructured data on the internet through text classification. Specifically focusing on Arabic text, the paper introduces the Optimal Configuration Determination for Arabic Text Classification method, utilizing the Particle Swarm Optimization algorithm. The proposed approach optimizes feature selection methods, machine learning classifiers, and a feature pool, considering a set of well-known feature selection methods, and classifiers. Experimental results demonstrate the method's success in achieving optimal configurations for Arabic text classification datasets, outperforming single classifiers and other state-of-the-art approaches.

A novel optimal deep learning based legal text summarization (ODL-LTS) technique for automatic legal text document summarization, employing TF-IDF and S-ROUGE measurements at two levels, is conducted in [13]. The summary generation is treated as a binary classification problem using the bidirectional gated recurrent neural network technique, with hyperparameters optimized by the glowworm swarm optimization (GSO) algorithm. The incorporation of GSO enhances efficiency in the summary generation process. Experimental validation against a dataset from Indian Supreme Court judgments demonstrates superior performance compared to other techniques.

### 2.3 Bio-Inspired based

In the Bio-Inspired category, a recent research [14] addresses the limitations of statistical feature selection approaches, emphasizing their failure to guarantee optimal reductions due to heuristic imperfections. Complete searches are unfeasible for medium-sized datasets, making stochastic approaches, such as ACO, promising for feature selection. The proposed ACO based technique is compared with other methods in text categorization, demonstrating competitive performance. The algorithm's effectiveness is showcased through experiments on the Reuters-21578 dataset, where it outperforms information gain and CHI methods in achieving better performance with fewer features.

In an alternative application of bio-inspired metaheuristic techniques[15], the authors tackle the formidable challenge posed by the curse of

dimensionality in the context of text document classification, a factor that significantly hampers overall system performance. The primary objective is to investigate the efficacy of the Social Spider Optimization (SSO) method in the context of feature selection and dimension reduction, aiming to augment the text classification capabilities of machine learning models. The study systematically assesses the impact of the SSO algorithm on various classifiers, encompassing Support Vector Machine, Logistic Regression, Random Forest, and Stochastic Gradient Descent, utilizing a standard open-source dataset. The experimental findings reveal a noteworthy enhancement in the classification performance of these classifiers, underscoring the potential of nature-inspired algorithms, such as SSO, to advance text classification proficiency across diverse domains like natural language processing and information retrieval.

In a related investigation[16], an enhanced binary GWO is employed for the classification of Arabic documents, serving to refine a feature selection methodology. The introduced technique is deployed across various learning models such as decision trees, K-nearest neighbor, Naïve Bayes, and Support Vector Machines classifiers, utilizing three distinct Arabic datasets. Outcomes reveal that the feature selection approach based on Support Vector Machines, enhanced by the proposed binary GWO optimizer with an elite-based crossover mechanism, exhibits superior performance in the classification of Arabic documents.

In [17], an innovative hybrid clustering approach, called the Spider Monkey Optimization algorithm using k-means, is introduced to analyze sentiments in tweets. The method utilizes k-means results to intelligently initialize the population in the proposed approach. The sander2 and Twitter datasets are employed for analysis, and the results are assessed for performance and accuracy. Comparative evaluations with state-of-the-art algorithms, including GA, PSO, SMO, and DE, demonstrate the accuracy of the proposed approach.

In the last research of this section, we present a study [18] which explores the application of ACO for optimized feature selection on both numerical and textual datasets. The proposed method demonstrates superiority over other approaches, exhibiting a significant performance across four machine learning models. Comparative analysis with multiple feature selection techniques, including TF-IDF, Bag of Words, and filter methods, highlights the challenges faced by text data and underscores the importance of ACO in addressing them. The numerical data outperforms textual data, emphasizing the advantages of numerical features for model performance.

### 2.4 Physics/Chemistry-Based

The initial study [19], introduces the Efficient Binary Black Hole (EBBH) algorithm designed for addressing high-dimensional text feature selection (FS) challenges in Semantic Textual Similarity (STS) resources. EBBH utilizes an innovative binary encoding framework to optimize search efficiency and diminish the number of features. The Population Fitness Adjustment (PFA) strategy ensures the initialization of a high-quality population, while the Adaptive Exploration Strategy (AES) dynamically adjusts parameters to maintain a balance between exploration and exploitation. Through experiments conducted on eight benchmark datasets, the introduced innovations enable EBBH to effectively explore feature subsets, exhibiting competitive performance in accuracy, selected features, fitness, and processing time in comparison to 12 alternative methods on STS datasets. While EBBH demonstrates excellence in accuracy and feature reduction, forthcoming research will explore alternative FS approaches to mitigate computational costs and investigate specific search strategies in conjunction with other classifiers. The proposed approach holds potential for extension to various domains beyond text feature selection.

In a recent study [20], the authors aim to enhance the accuracy of machine learning classifiers through improved feature selection. Two methods were implemented, and their performances were evaluated using three potential energy functions for classification problems. Comprehensive comparisons with existing feature selection techniques and five machine learning classifiers revealed the superiority of the proposed technique in most cases. For a KNN classifier, the proposed technique demonstrates significant improvements over existing methods.

### 2.5 Human-Based

In [21], the challenge of globally optimizing the partitioning of a set of documents into a specified number of clusters is addressed. The novel ICA is compared to k-means clustering and Particle Swarm Optimization (PSO) based clustering methods. A hybrid approach is formed by employing k-means clustering as an agent for ICA and PSO algorithms in the initial iteration. Experimental results on two document sets reveal that ICA yields superior solutions with high-quality F-measure compared to

k-means and PSO. While k-means converges faster, it may get stuck in local optima, and the hybrid ICA combines the global searching ability of ICA with the rapid convergence of k-means, overcoming drawbacks of both algorithms. The hybrid ICA demonstrates higher compact clustering compared to hybrid PSO or k-means alone.

A two-stage feature selection method named IGICA is introduced in [22]. The approach involves utilizing Information Gain in the first phase to eliminate redundant and irrelevant features, followed by the application of the ICA for feature selection in the second stage. The key advantage lies in the combination of filter and wrapper methods. The k-nearest neighbor classifier is employed for document classification, and the results demonstrate increased efficiency in document categorization compared to other methods.

In a recent research [23], a solution to enhance the SVM algorithm using the ICA. The proposed method employs ICA for feature selection and SVM for text classification. Various weighting schemes are utilized during the text feature extraction stage, assigning weights to extracted words to determine their impact as keywords. The key phase of ICA, the colonial competition phase, identifies colonies with high classification errors and features as weaker empires. The proposed method achieves efficient text classification by using a fraction of key features along with a classifier algorithm, such as SVM, and employs the RBF kernel function for mapping the nonlinear problem to a linear one. Evaluation on 12 datasets demonstrates the superiority of the proposed hybrid algorithm compared to standard SVM approaches.

Kavitha et al. [24] introduces a Weighted Feature-based Imperialist Competitive Algorithm (WFICA) combined with Ensemble Learning to classify imbalanced data effectively. Initial steps involve z-score normalization to enhance accuracy, followed by Synthetic Minority Oversampling Technique with Locally Linear Embedding for handling imbalanced data. Optimal features are then selected using WFICA. Ensemble Learning is used in the classification, including Improved Bidirectional Long Short-Term Memory, Enhanced Weighted Support Vector Machine, and k-Nearest Neighbour classifiers.

In another study [25] within this category, the authors introduce a model for sentiment identification and classification in social media, called SOADL-SAC. The model involves data preprocessing, feature vector generation using the glove technique, sentiment recognition and classification, and hyperparameter tuning. Experimental results demonstrate the superior performance of the SOADL-SAC model compared to recent state-of-the-art approaches on benchmark datasets.

The research [26] emphasizes the impact of spelling correction in text processing, combined with natural language processing, on sentiment classification accuracy. It introduces the RSTLBO algorithm, a combination of rough set theory (RST) and teaching-learning based optimization (TLBO), for feature selection in sentiment classification. The combination of text processing and RSTLBO outperforms text processing alone, showcasing the importance of effective feature selection. Statistical analysis confirms the significance of RSTLBO, demonstrating its superiority over other methods. The study underscores the crucial role of text processing in sentiment analysis and the positive impact of feature selection on classification accuracy.

In [27], the authors introduce a novel approach using nature-based optimization, specifically Biogeography Based Optimization (BBO), as a feature selection technique combined with an ensemble classifier for optimal text classification. The BBO-based model is compared with other state-of-the-art algorithms on various datasets, demonstrating superior performance in terms of Accuracy, Precision, Recall, and F-measure compared to GA and PSO. The model is validated on a real-time airline dataset and is considered a competitive method for practical text classification optimization.

### 2.6 Plant-Based

In this family of metaheuristic research, The HAN classifier [28], integrated with the IWTSO algorithm, is developed for text classification. The process involves pre-processing, feature extraction, and feature selection to enhance classification performance. The applications of text classification include spam detection, language detection, sentiment analysis, speech recognition, topic labeling, and intent detection.

In [29], a novel Chinese text clustering algorithm is introduced by leveraging the search capabilities of the IWO algorithm, coupled with the low computational complexity of K-Means. The algorithm is designed to optimize the evaluation function, which involves the ratio of cluster separation to tightness within clusters. Various

experimental datasets are utilized to assess the algorithm's performance. The findings indicate that the proposed IWO-KM1 algorithm outperforms all other methods considered in the research when applied to the experimental data.

In another plant-based research [30], the authors utilize FPA and Ada-Boost for the classification of text documents, focusing on educational documents in the Ada-Boost model. Preprocessing involves balancing class sizes through word frequency recognition. Feature selection employs FPA, and after selecting properties, documents undergo training using the Ada-Boost model. The proposed model demonstrates higher accuracy compared to KNN-K-Means and NB-K-Means models.

One last research [31] in this category examines the Natural Language Processing method, specifically spelling correction in text preprocessing, in comparison to conventional text preprocessing techniques. Additionally, the study proposes the use of FPA algorithms for feature selection to enhance sentiment classification performance. The approach demonstrated promising results in accuracy compared to the baseline model.

## 4. Conclusions

This review paper has highlighted the significant impact of metaheuristic optimization techniques on advancing the field of text classification. From evolutionary algorithms to swarm intelligence, bio-inspired methods to human-based strategies, and even plant-based approaches, researchers have leveraged diverse metaheuristic techniques to enhance the accuracy, efficiency, and robustness of classification systems.

Through a comprehensive analysis of recent research studies, we have observed the versatility and effectiveness of metaheuristic algorithms in addressing key challenges in text classification, such as feature selection, dimensionality reduction, class imbalance, and noisy data. These techniques have demonstrated promising results across various domains, including sentiment analysis, document categorization, information retrieval, and topic modelling.

## Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

## References

[1]Abdel-Basset, M., Mohamed, R., Jameel, M., & Abouhawwash, M. (2023). Nutcracker optimizer: A novel nature-inspired metaheuristic algorithm for global optimization and engineering design problems. *Knowledge-Based Systems*, 262, 1-36. https://doi.org/10.1016/j.knosys.2022.110248

[2]Mousapour Mamoudan, M., Ostadi, A., Pourkhodabakhsh, N., Fathollahi-Fard, A. M., & Soleimani, F. (2023). Hybrid neural network-based metaheuristics for prediction of financial markets: a case study on global gold market. *Journal of Computational Design and Engineering*, 10(3), 1110-1125. https://doi.org/10.1093/jcde/qwad039

[3]Savanović, N., Toskovic, A., Petrovic, A., Zivkovic, M., Damaševičius, R., Jovanovic, L., ... & Nikolic, B. (2023). Intrusion detection in healthcare 4.0 internet of things systems via metaheuristics optimized machine learning. *Sustainability*, 15(16), 1-28. https://doi.org/10.3390/su151612563

[4]Abed, A. M., Seddek, L. F., & AlArjani, A. (2023). Enhancing Two-Phase Supply Chain Network Distribution via Three Meta-Heuristic Optimization Algorithms Subsidized by Mathematical Procedures. *Journal of Advanced Manufacturing Systems*, 22(03), 445-476. https://doi.org/10.1142/S0219686723500221

[5]Alkanhel, R., El-kenawy, E. S. M., Elsheweikh, D. L., Abdelhamid, A. A., Ibrahim, A., & Khafaga, D. S. (2023). Metaheuristic Optimization of Time Series Models for Predicting Networks Traffic. *Computers Materials & Continua*, 75(1), 427-442. https://doi.org/10.32604/cmc.2023.032885

[6]Alsaleh, D., & Larabi-Marie-Sainte, S. (2021). Arabic Text Classification Using Convolutional Neural Network and Genetic Algorithms. *IEEE Access*, 9, 91670-91685. https://doi.org/10.1109/ACCESS.2021.3091376

[7]Labani, M., Moradi, P., & Jalili, M. (2020). A multi-objective genetic algorithm for text feature selection using the relative discriminative criterion. *Expert Systems with Applications*, 149, 1-21. https://doi.org/10.1016/j.eswa.2020.113276

[8]Dwarakanath, L., Kamsin, A., & Shuib, L. (2023). A Genetic Algorithm based Domain Adaptation Framework for Classification of Disaster Topic Text Tweets. *Int. Arab J. Inf. Technol*., 20(1), 57-65. https://doi.org/10.34028/iajit/20/1/7

[9]Jin, Q. (2023). Genetic algorithm and support vector machine application in English text classification for intelligent teaching. *Soft Computing*, 27(16), 1-12. https://doi.org/10.1007/s00500-023-09084-x

[10]Mohana, R. S., Rajathi, K., Kousalya, K., & Yuvaraja, T. (2022). Text sentiment analysis on E-shopping product reviews using chaotic coyote optimized deep belief network approach. *Concurrency and Computation: Practice and Experience*, 34(19), 1-16. https://doi.org/10.1002/cpe.7039

[11]Parlar, T., & Sarac, E. (2019). IWD based feature selection algorithm for sentiment analysis. *Elektronika ir Elektrotechnika*, 25(1), 54-58. https://doi.org/10.5755/j01.eie.25.1.22736

[12]Alhaj, Y. A., Dahou, A., Al-qaness, M. A. A., Abualigah, L., Abbasi, A. A., Almaweri, N. A. O., Elaziz, M. A., & Damaševičius, R. (2022). A novel text classification technique using improved particle swarm optimization: A case study of Arabic language. *Future Internet*, 14(7), 1-18. https://doi.org/10.3390/fi14070194

[13]Vaissnave, V., & Deepalakshmi, P. (2023). Modeling of automated glowworm swarm optimization based deep learning model for legal text summarization. *Multimedia Tools and Applications*, 82(17), 17175–17194. https://doi.org/10.1007/s11042-022-14171-6

[14]Hosseinzadeh Aghdam, M., Ghasem-Aghaee, N., & Basiri, M. E. (2009). Text feature selection using ant colony optimization. *Expert Systems with Applications*, 36(3, Part 2), 6843–6853. https://doi.org/10.1016/j.eswa.2008.08.022

[15]Al-Anzi, F. S., & Sarath, S. (2023). Social Spider Optimization for Text Classification Enhancement. In C. Kahraman, I. U. Sari, B. Oztaysi, S. Cebi, S. Cevik Onar, & A. Ç. Tolga (Eds.), *Intelligent and Fuzzy Systems, 532-539*. https://doi.org/10.1007/978-3-031-39774-5_59

[16]Chantar, H., Mafarja, M., Alsawalqah, H., et al. (2020). Feature selection using binary grey wolf optimizer with elite-based crossover for Arabic text classification. *Neural Computing & Applications*, 32(21), 12201–12220. https://doi.org/10.1007/s00521-019-04368-6

[17]Shekhawat, S. S., Shringi, S., & Sharma, H. (2021). Twitter sentiment analysis using hybrid Spider Monkey optimization method. *Evolutionary Intelligence*, 14(4), 1307–1316. https://doi.org/10.1007/s12065-019-00334-2

[18]Gite, S., Patil, S., Dharrao, D., Yadav, M., Basak, S., Rajendran, A., & Kotecha, K. (2023). Textual Feature Extraction Using Ant Colony Optimization for Hate Speech Classification. *Big Data and Cognitive Computing*, 7(1), 1-23. https://doi.org/10.3390/bdcc7010045

[19]Wu, X., Fei, M., Wu, D., Zhou, W., Du, S., & Fei, Z. (2023). Enhanced Binary Black Hole algorithm for text feature selection on resources classification. *Knowledge-Based Systems*, 274, 1-23. https://doi.org/10.1016/j.knosys.2023.110635

[20]Rao, P. C. S., Kumar, A. J. S., Niyaz, Q., Sidike, P., & Devabhaktuni, V. K. (2021). Binary chemical reaction optimization based feature selection techniques for machine learning classification problems. *Expert Systems with Applications*, 167, 1-12. https://doi.org/10.1016/j.eswa.2020.114169

[21]Küçükdeniz, T., & Büyüksaatçı, S. (2014, June 4-6). Imperialist Competitive Algorithm Compared With Particle Swarm Optimization And K-Means On Document Clustering. *International Conference on Engineering and Applied Sciences Optimization (OPTI 2014), Kos Island-Greece.* https://www.ses.gr/opti-2014-1st-announcement/

[22]Mojaveriyan, M., Ebrahimpour-komleh, H., & jalaleddin Mousavirad, S. (2016). IGICA: A hybrid feature selection approach in text categorization. *International Journal of Intelligent Systems and Applications*, 8(3), 42-47. https://doi.org/10.5815/ijisa.2016.03.05

[23]Asheghi Dizaji, Z., Asghari Aghjehdizaj, S., & Soleimanian Gharehchopogh, F. (2020). An Improvement in Support Vector Machines Algorithm with Imperialism Competitive Algorithm for Text Documents Classification. *Journal of Soft Data Processing*, 17(1), 117-130. https://doi.org/10.29252/jsdp.17.1.117

[24]Kavitha, D. (2021). Weighted Feature Based Imperialist Competitive Algorithm With Ensemble Learning For Imbalanced Data Classification. *Turkish Journal of Computer and Mathematics Education* (TURCOMAT), 12(7), 1037–1050. https://www.turcomat.org/index.php/turkbilmat/article/view/2713

[25]Alghamdi, H. M., Hamza, S. H., Mashraqi, A. M., & Abdel-Khalek, S. (2022). Seeker Optimization with Deep Learning Enabled Sentiment Analysis on Social Media. *Computers, Materials & Continua*, 73(3), 1-15. https://doi.org/10.32604/cmc.2022.031732

[26] Muhammad, A., Abdullah, S., & Sani, N. S. (2021). Optimization of Sentiment Analysis Using Teaching-Learning Based Algorithm. *Computers, Materials & Continua*, 69(2), 1-17. https://doi.org/10.32604/cmc.2021.018593

[27]Khurana, A., & Verma, O.P. (2020). Novel approach with nature-inspired and ensemble techniques for optimal text classification. *Multimedia Tools and Applications*, 79(21), 23821-23848. https://doi.org/10.1007/s11042-020-09013-2

[28]Singh, G., Nagpal, A., & Singh, V. (2023). Optimal feature selection and invasive weed tunicate swarm algorithm-based hierarchical attention network for text classification. *Connection Science*, 35(1), 1-25. https://doi.org/10.1080/09540091.2023.2231171

[29]Fan, C., Zhang, T., Yang, Z., & Wang, L. (2015, August 10-14). A Text Clustering Algorithm Hybriding Invasive Weed Optimization with K-Means. *IEEE 12th Intl Conf on Ubiquitous Intelligence and Computing and 2015 IEEE 12th Intl Conf on Autonomic and Trusted Computing and 2015 IEEE 15th Intl Conf on Scalable Computing and Communications and Its Associated Workshops (UIC-*

*ATC 2015), Beijing-China.* https://doi.org/10.1109/UIC-ATC-ScalCom-CBDCom-IoP.2015.241

[30]Majidpour, H., & Soleimanian Gharehchopogh, F. (2018). An improved flower pollination algorithm with AdaBoost algorithm for feature selection in text documents classification. *Journal of Advances in Computer Research*, 9(1), 29-40. https://sanad.iau.ir/journal/acr/Article/653945?jid=653945

[31]Muhammad Iqbal, A. L., Yaakub, M. R., & Ibrahim, S. A. (2022). Flower Pollination Algorithm for Feature Selection in Tweets Sentiment Analysis. *International Journal of Advanced Computer Science and Applications*,13(5), 429-436. https://doi.org/10.14569/IJACSA.2022.013055