Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Differential Evolution-Initialized Cuckoo Search for AI-Based Feature Selection in CoWIN Twitter Sentiment Analysis

Dharmpal singh
dharmpal1982@gmail.com
JIS University
Ira Nath
JIS College of Engineering
Sumit Das
JIS College of Engineering
Kaushik Adhikary
JIS University
Nguyen Thi Kim Son
Hanoi University of Industry
Sudipta Sahana
University of Engineering & Management

Research Article

Keywords: Cuckoo Search, Feature Selection, Sentiment Analysis, CoWIN, Twitter, Metaheuristic Optimization

Posted Date: November 5th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11127258/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Content in social media is hard to be analyzed by sentiment as there is increasing amount of social media content related to health, and the dimensionality and computation required to handle the redundant and irrelevant features in the content representation. In this research, a feature selection method based on DE-Initialized Hybrid Cuckoo Search Algorithm (HCSA) is suggested for Twitter data to be used in health-related sentiment analysis. The proposed method is based on the Cuckoo Search (CS) optimization algorithm where a Population Initialization Strategy (DE) is introduced. The approach is applied to the features as well as the sentiment classification, and the features are represented by the Term Frequency–Inverse Document Frequency (TF-IDF) matrix and the sentiment classification is provided by Support Vector Machine (SVM) classification algorithm taking the CoWIN Twitter sentiment data set as the data set. The experimental setup is identical to the Cuckoo Search (CS), Differential Evolution (DE), Particle Swarm Optimization (PSO) and Genetic Algorithm (GA) to make comparisons. From the results, it is seen that the number of features is reduced by 48.40% from 500 to 258 dimensions with no reduction in the classification accuracy for the HCSA. The highest accuracy and F1 scores that can be achieved from the observed HCSA are 82.90% and 73.51%, respectively, with the highest macro-precision and macro-recall scores of 85.55% and 73.05%, respectively. These results suggest that the proposed DE-initialized HCSA can effectively compress the text's dimensionality in the meantime retaining useful information for sentiment classification. The study found that hybrid metaheuristic feature selection could be one of the most effective feature selection methods to extract the information about health from social media text.

1. Introduction

Social media platforms are increasingly important platforms used to measure public perception, experiences and opinions. Sentiment-analysis can be used on short, constantly changing text, such as that found on Twitter. The purpose of sentiment analysis is to classify the sentiment or sentiment tone of text, and it has many uses, including sentiment analysis of public opinion, sentiment analysis of product reviews, sentiment analysis of health forums and sentiment analysis of social media. From a machine-learning and computational point of view, sentiment analysis has been investigated in the past, with the emphasis predominantly put on the representation of the text, features extraction, features selection and classification techniques [113, 114].

There are a number of computational challenges in sentiment classification of social-media text. The feature space for tweets is high dimensional, as is the number of computations needed due to the redundant and irrelevant features of the text, noisy text and highly correlated features. To overcome these challenges, feature-selection techniques can be used to discard irrelevant or redundant features, and to select informative ones. In feature selection problems, the number of features and their complexity increases, which leads to the increased interest in the use of metaheuristic optimization algorithms for them. Genetic Algorithm (GA), Particle Swarm Optimization (PSO), Differential Evolution (DE) and Cuckoo Search (CS) are some optimization strategies that have been investigated for optimization [115, 116, 118].

Cuckoo Search is a population-based metaheuristic optimization algorithm which is inspired by brood-parasitic behaviour of cuckoo species and Lévy-flight-based search. CS has been used in various optimization problems in recent years such as engineering optimization [124], feature selection [126] and image processing and forecasting [128] etc. The performance of CS can be influenced by the population initialization, the design of the objective function, search strategies and values of the parameters. Thus, other starting populations and optimizations settings can result in other searches. This has encouraged the research of new and hybrid versions of CS which seek to improve the exploration/exploitation capabilities of the algorithm [123, 127].

There are a few studies conducted related to the application of CS in SA.There are a couple of research papers dealing with the application of Cuckoo Search on the sentiment analysis field. Some studies explored the application of a hybrid Cuckoo Search based algorithm on the sentiment classification in Twitter sentiment analysis [116] and feature selection to improve the opinion mining and classification [118]. Besides, several studies of hybridizing the DE and the CS with other optimization algorithms for some optimization problems have been performed [125]. The studies suggest that the integration of complementary search approaches can give an opportunity to improve the exploration–exploitation behaviour of population-based optimization algorithms.

The application of optimization-based sentiment analysis is explicitly related to health-related information in social media. During COVID-19, Twitter was a hugely important channel for discussion of vaccinations, health policy, and information and experience of the pandemic. There has been prior work that has studied the sentiment analysis and public opinion of tweets related to COVID-19 using hybrid computational methods [129, 131]. However, when the size and complexity of the text of health-related social media platforms are increased, it requires effective feature selection techniques to reduce the number of features while retaining useful information to be used in sentiment recognition.

The CoWIN Twitter Dataset is a suitable dataset in real world scenarios to solve this problem. The total number of raw tweets for each of the 12 months from January to December 2021 is approximately 635,000 with a subset of these tweets being converted to English text and a subset being labeled with sentiment [140]. The conversations can be retrieved from the public and thus may be used as a social media platform to test health-related text data feature-selection and sentiment-classification methods.

The problem of sentiment analysis on Twitter data has been additionally investigated using Hybrid Cuckoo Search approach prior to [116] but there is a need for further research on other hybrid CS models for health related Twitter data. Actually, the initial population generation process in Cuckoo Search algorithm—namely the population-diversity-generation process of Differential Evolution—in the initial population generation process of the CS search process can be used. In this research, this attraction is used to propose a Hybrid Cuckoo Search Algorithm (HCSA) to select features in the health-related sentiment analysis of Twitter.

In the proposed HCSA, the CoWIN Twitter Data is used as a sentiment evaluation data; the TF-IDF textual feature representation and the Support Vector Machine (SVM) classifiers are used to classify the sentiment. The proposed approach is then tested against four other optimization methods such as GA, PSO, Differential Evolution (DE) and Standard Cuckoo Search (CS) in the same experimental setting. The two aspects of the evaluation are: Firstly,

classification accuracy and secondly, feature reduction which will assess the proposed hybrid approach for dimensionality reduction of textual data with competitive sentiment-classification accuracy.

1.1 Research Gap and Novelty

While Cuckoo Search has been explored in feature selection and sentiment analysis, the application of Differential Evolution as a population-initialization method for a Cuckoo Search feature-selection process in health related Twitter sentiment analysis has been explored to a limited extent. To fill this gap, this study uses the two optimization mechanisms separately: Differential Evolution is used to generate the initial population, and Cuckoo Search is used to do the feature-selection search for the subsequent population.

The novelty of proposed algorithm (HCSA) is not only the combination of two optimization algorithms but also the CS initializes CS population using DE algorithm and systematically examines the role of CS contribution in feature selections. The proposed framework is evaluated with controlled comparison experiments with GA, PSO, DE and standard CS, independent runs and statistical significance test and ablation analysis at the component level.

The main research question here is if whether initial population of a DE can be adopted to serve as a good starting point for the Cuckoo Search to achieve competitive performance with respect to dimensionality of feature space (TF-IDF) and sentiment classification.

1.2 Contributions of the Study

The key findings of this study are that:

- A hybrid feature-selection framework using the DE to only generate the initial population for the subsequent step of optimization by Cuckoo Search is proposed.
- While the use of a CS is important, the main interest of the study is the effect that the population initialization has on the feature-selection process implemented by the CS.
- Health-related social-media application: The proposed framework is applied to test the proposed framework with the health-related Twitter sentiment data pertaining to CoWIN – a high-dimensional health-related textual-data application.
- Controlled comparative evaluation: HCSA is compared to other algorithm: Genetic Algorithm, Particle swarm optimization, Differential Evolution and regular Cuckoo Search in the same feature representation, same classifier and experimental environment.
- Evaluation: The methods are evaluated through multi-dimensional criteria including the accuracy of the classification, reduction of features, number of selected features and computational time.
- To evaluate the consistencies of the performance differences observed, there are thirty independent runs and Wilcoxon signed rank test and Friedman test are used.
- Ablation study to investigate the effect of the DE-based initialization and the modified search part of the whole HCSA framework.

1.3 Research Questions

In this research, the following research questions will be addressed;

RQ1: How to effectively combine the population initialization of Differential Evolution and Feature Selection of Textual with Cuckoo Search?

RQ2: What is the performance of classification and feature reduction of the proposed HCSA as compared to the GA, PSO, DE and the standard CS?

RQ3: Does the population initialization with DE affect the process of feature selection with CS?

RQ4: Are the observed differences in performance seen in independent stochastic runs?

RQ5: How many computational costs does the proposed hybrid framework need when compared to computational cost of the baseline optimization methods?

1.4 Relevance to the Journal Scope

The study makes a contribution to computer and information science by suggesting the use of evolutionary optimization as well as the use of the metaheuristic search combined with feature selection, NLP combined with machine learning based sentiment classification. The research, however, is not limited to the domain of application, but rather is a computational feature-selection approach to high dimensional textual data and investigates its performance by means of a set of controlled experiments and statistical tests. The work is therefore related to the feature selection based on optimization and text mining / sentiment analysis applications.

The study thus investigates if the suggested DE-initialized HCSA can provide a valuable trade-off between feature space reduction and sentiment-classification accuracy in the health-related Twitter data.

The results investigate whether HCSA can reduce textual dimensionality without loss of important information for sentiment classification. Figure 1 showed the major research direction on Cuckoo search.

The usage of Cuckoo search in different field vied furnished below in Table 1.

Table 1
Search on Cuckoo algorithm

Optimization Approach	Citation
Improved Version of Cuckoo Search	[94]
Adaptive and Modified Cuckoo Search Approach	[110]
Multi-Objective Optimization Using Cuckoo Search	[107]
Complex-Valued Implementation of Cuckoo Search	[112]
Cuckoo Search Utilizing Gaussian Distribution	[111]
Gaussian Disturbance-Based Cuckoo Search	[97]
Binary Variant of Discrete Cuckoo Search	[26]
Standard Discrete Cuckoo Search	[34]
Discrete Cuckoo Search Applied to the Traveling Salesman Problem	[61]
Neural Network-Integrated Cuckoo Search	[41]
Quantum-Enhanced Cuckoo Search	[45]
Chaotic and Emotion-Based Cuckoo Search	[49]
Cuckoo Search Incorporating Local Minimization Techniques	[59]
Parallelized Implementation of Cuckoo Search	[82]
Refined Cuckoo Search Algorithm	[86]
Cuckoo Search Utilizes binary encoding for classification tasks	[121]
Multi-Objective Cuckoo Search Optimizes multiple objectives relevant to sentiment analysis	[122]
Enhanced Cuckoo Search Incorporates advanced heuristics for improved convergence and accuracy	[123]

The CS algorithm used in various engineering problem as furnished below in Fig. 2.

Figure 2 showed the CS algorithm used in various engineering problem.Table 2 showed the optimization problem on different domain whereas Fig. 3 and table 3is about the application of cuckoo search on different area and demonstrate the different utilization of Cuckoo search.

Table 2
Usage of Cuckoo search in engineering optimization
Problem

Optimization Problem	Citation
Optimization in engineering	[105]
Placement of capacitors	[3]
Design of six-bar mechanisms	[8]
Blades for wind turbines	[22]
Truss structure optimization	[24]
Deflection due to electrostatics	[30]
Design of steel frames	[38]
Steel structure design	[39]
Arrays for antennas	[42]
Exploration of design spaces	[43, 44]
Sequence optimization	[48]
Design of planar EBG structures	[62]
Analysis of stability	[65]
Optimization of linear antenna arrays	[68, 66]
Optimal placement of capacitors	[69]
Sizing and allocation of distributed generation	[84]
Problems related to reliability	[87, 88]
Estimation of non-linear states	[92]
Calculations for phase equilibrium	[6]
Optimization in structural design	[19]

Table 3
Demonstrate the different utilization of Cuckoo search.

Application	Reference
Distributed generation allocation in networks	[53]
Optimization of Bloom filters	[55, 56, 57]
Backpropagation neural network	[58]
Traveling salesman problem	[61]
Web service integration	[63]
Web service composition	[11, 12]
Ontology matching	[70]
Voice recognition	[76]
Automated software testing	[79, 80, 81]
Optimization in manufacturing	[83]
Facial recognition technology	[85]
Training of neural models	[89]
n-Queens problem	[73]
Game optimization	[77, 78]
Quadratic assignment problems	[15]
Optimization of sheet nesting	[21]
Economic dispatch in non-convex problems	[91]
UCAV pathfinding	[99]
Business process optimization	[108]
Selection of machining parameters	[109]
Job scheduling in distributed systems	[64]
Query optimization	[35]
Thresholding of multi-level images	[7]
Flood prediction	[10]
Wireless sensor network optimization	[16]
Data integration	[17]
Clustering in wireless networks	[18]
Clustering analysis	[29]
Groundwater exploration	[31]
Selection of suppliers	[37]
Power load prediction	[40]
Surface texture measurement	[50]
Flow shop optimization	[51]
Optimal replacement strategy	[52]
Parallelization of CS	[36]

The behavior of the CS algorithm has been studied by many researchers to enhance the optimization performance of the known variants of the CS algorithm have been proposed that are a hybrid of the CS algorithm or improvements to the base algorithm. This has been achieved using combinations of CS and Particle Swarm Optimization (PSO) [28, 96], combinations of CS and the Tsallis entropy [1] as well as improved Differential Evolution (DE) [54, 123]. Other studies have come up with tweaked versions of search and distribution to maximize exploration and exploitation in CS [23, 72]. In the literature [126, 128] it is shown that CS can be used successfully for various optimization problems, including function optimization, engineering optimization, image processing, forecasting, planning, scheduling and feature selection. Combining other optimization methods has also broadened the use of CS to real world problems.

Objective function design, initialization strategy, search mechanism and parameter setting are important factors that affect the performance of a metaheuristic algorithm. Different values of the parameters can lead to different optimization results with the same set of data or to different data sets, the

same data can lead to different optimization results, if the data is used with inappropriate parameter values, then early convergence or suboptimal optimizations can be obtained [124]. Therefore, it is important to design an appropriate initialization and optimization algorithm to improve CS robustness and effectiveness.

Sentiment-analysis and classification of social media have been solved also with Cuckoo Search. There have been studies that investigated hybrid CS-based approaches and feature-selection techniques on Twitter and textual data in general [116, 118]. The outcomes of these studies show the feasibility of CS to find out informative features and increase the accuracy of sentiment classification. But more studies are needed to identify whether it is possible to obtain an adequate reduction in features and classification performance from an adequately designed hybrid CS framework, especially for social-media data pertaining to health.

The present literature in Table 4 demonstrates that the Cuckoo Search algorithm has been used for sentiment analysis and feature selection [116, 118, 141] and hybrid algorithm of Differential Evolution and Cuckoo Search has been studied for optimization problem [125]. Other research studies have also investigated the computational sentiment analysis of COVID-19 on Twitter [129] and public perception analysis related to COVID-19 [131]. For the current study, however, it will be discussed how to use Differential Evolution (DE) method to generate the initial population for the second feature-selection process that will be performed using Cuckoo Search (CS) algorithm for health-related Twitter data. The separation of the DE based initialization process from the following CS search is the foundation from which it is possible to experiment with the role of population initialization in the high dimensional textual feature selection problem as shown in Table 4

Table 4
Comparison of Existing Studies and the Proposed Framework

Study	Method / Approach	Application	Feature Selection	Population Initialization	Health-related Twitter	Main distinction from present study
Pandey et al. [116]	Hybrid Cuckoo Search	Twitter sentiment analysis	Yes	Different from proposed DE initialization	No	Hybrid CS applied to Twitter sentiment
Study [118]	CS-based optimized opinion mining/classification	Sentiment analysis	Yes	Not reported as DE-based initialization	No	CS used for optimized opinion mining
Chi et al. [125]	Hybrid DE-CS	Multi-objective optimization	Optimization problem	DE and CS are hybridized for optimization	No	DE-CS studied for general optimization rather than CoWIN feature selection
Jana et al. [129]	Hybrid computational approach	COVID-19 Twitter sentiment	Not the primary focus	Not the proposed DE initialization	Yes	Health-related Twitter sentiment analysis
Study [131]	Hybridized computational approach	Public perception analysis	Not the primary focus	Not the proposed DE initialization	Related context	Hybrid computational analysis of public perception
Study [141]	Hybrid CS / population-generation strategy	Health-related Twitter sentiment	Yes	Alternative population-generation strategy	Yes	Health-related sentiment with CS-based feature selection
This study	**DE-initialized Hybrid Cuckoo Search (HCSA)**	**CoWIN Twitter sentiment**	**Yes**	**DE generates initial CS population**	**Yes**	**Explicitly investigates DE-based initialization for subsequent CS feature selection**

In this study, a Hybrid Cuckoo Search Algorithm (HCSA) is proposed which is based on an initialization strategy for the Cuckoo Search algorithm using the Differential Evolution (DE) algorithm. The proposed algorithm is compared with existing optimization algorithms like the Genetic Algorithm (GA), the Particle Swarm Optimization (PSO), Differential Evolution (DE) and the Cuckoo Search (CS). The approach is then used for sentiment analysis of the CoWIN related Twitter data and the effectiveness of the approach to identify informative text features while producing comparable classification accuracy is explored. The rest of this paper is structured in the following way. This is presented in section 2, where the proposed Hybrid Cuckoo Search Algorithm is presented and the mathematical formulation for the same is also given. In Section 3, the description of datasets, data preprocessing, feature representation and experimental settings and implementation of the proposed approach are described. The results of the experiments and comparison are presented in section 4. The results, limitations and implications of this research are discussed in section 5. Finally, Section 6 concludes the paper and gives some directions for future research.

2. Methodology

The authors [60] have opined that more than 1000 species of birds have similar behaviors to lay eggs with different shapes and built a net in safe places to protect themselves from the predators [61].Cuckoo eggs laying decision was discussed in the paper [62] and sneakiness to increase approach to extra provision chance presented in [63].Deb and Yang [10] have proposed usage of CSA in the optimization context whereas the usage of CSO in different optimization methodologies was proposed by [64, 10]. The authors [65, 66] have opined about Le´vy flights and observed haphazard walk in real life [67, 68] whereas the worldwide exploration and local exploitation presented in [69].Authors have summarized specific behaviours of cuckoo egg laying capacity as compared to other birds in [70]. The following algorithm imitate the standard cuckoo search behaviours in Fig. 4

Furthermore, Yang and Dep [10] have stated CSA used less organize parameters as compared to many algorithm

2.1 DE-Based Population Initialization

In the proposed HCSA scheme, instead of using it as the last stage feature optimization method, Differential Evolution is used at the initialization stage of the population. Each candidate solution is a binary feature-selection vector, where each dimension corresponds to a TF-IDF feature, and is either selected or not.

The initialization mechanism is based on the DE which generates the initial population of candidates, and the obtained population is fed into the Cuckoo Search (CS) optimization mechanism. The next search is then carried out by the CS mechanism. This design can be used to test the effect of initial population generation strategy, apart from the CS search.

The overall process can be represented as shown below:

TF-IDF feature space, binary feature representation, population initialization using DE, optimization using DE, selection of an appropriate feature subset, classification using SVM.

This design is intended to investigate if a population generated by the DE can be useful to establish a good initial population for the feature selection performed by CS and the difference between standard CS and proposed HCSA is presented in Table 5.

Table 5
Difference Between Standard CS and Proposed HCSA

Component	Standard CS	Proposed HCSA
Initial population	Standard/random initialization	DE-based population initialization
Main optimizer	Cuckoo Search	Cuckoo Search
DE role	Not used	Population initialization
Search after initialization	CS	CS
Feature representation	Binary feature vector	Binary feature vector
Classifier	SVM	SVM
Feature-selection objective	Classification + feature reduction	Classification + feature reduction

The difference between the standard CS and HCSA is that a different method is used in the population initialization procedure. It is not an optimization process for the Differential Evolution optimization process to replace the HCSA. DE is instead used to create the initial population, and CS is used to optimize the population afterwards. This distinction provides the guideline of the methodological contribution that is offered in this paper.

2.1: Hybrid Cuckoo Search Algorithm

The proposed Hybrid Cuckoo Search Algorithm (HCSA) is proposed to perform feature selection on the TF-IDF representation of the CoWIN Twitter dataset. The main difference in the methodology is the way the initial population is generated with Differential Evolution (DE) and the final population is optimized using feature-selection with Cuckoo Search (CS) algorithm.

In HCSA, the initial population is generated first based on the conventional initialization procedure, while in standard CS, the initial population is generated based on the conventional initialization procedure. The output of the population is then fed into the CS search procedure. Accordingly, the proposed framework brings two distinct roles on the two different components: DE is used to initialize the population where CS is used to do the search and feature-selection optimization.

The overall HCSA procedure is as follows:

Step 1: TF-IDF Feature Representation : The first step is feature representation using TF-IDF.

Preprocessed CoWIN twitter text is converted into a TF-IDF feature matrix. In the experimental configuration, a maximum of 500 TF-IDF features are taken into account.The second step involves encoding the binary features.

Step 2: Binary Feature Encoding:

A binary vector is used to represent each candidate solution:

$$X = (x_1, x_2, \ldots, x_N)$$

The values of $(x_j = 1)$ mean that the jth feature is selected, and the values of $x_j = 0$ mean that the feature is not selected.

Step 3: Initialize Population using DE.

DE is used to create the initial population of candidate solutions. The solutions generated by the DE are the initial population of the next Cuckoo Search process.

Step 4: Cuckoo Search Optimization

The DE-generated population is fed to the Cuckoo Search algorithm. However, CS conducts its search operations based on the candidate generation and replacement schemes it has, in order to find the potentially useful feature subsets.

Step 5: Fitness Evaluation

The defined feature selection objective and SVM classifier are used to evaluate each candidate feature subset. The classification performance and feature-selection requirements mentioned in the experiment are taken into account in the fitness evaluation.

Step 6: Iterative Search

The Cuckoo Search process is continued until the stopping criterion is met. In every iteration, new candidate solutions are generated, tested and compared to current solutions, and the best solution found by the optimization process is kept.

Step 7: Final Feature Subset

The selected feature subset at termination is the best binary solution. The selected features are later fed to the SVM classifier for the sentiment classification.

The overall workflow of the proposed DE-initialized Hybrid Cuckoo Search (HCSA) framework, including TF-IDF feature representation, DE-based population initialization, subsequent Cuckoo Search optimization, fitness evaluation, and final feature selection, is illustrated in Fig. 5.

The convergence behaviour of the evaluated optimization algorithms across the iterative search process is presented in Fig. 6, providing a comparative view of their optimization progress.

The comparative performance of the evaluated feature-selection methods in terms of classification performance, feature reduction, and computational efficiency is summarized in Fig. 7.

3. Experimental Implementation and Setup

3.1 Dataset Description

The proposed Hybrid Cuckoo Search Algorithm (HCSA) is tested using the CoWIN Twitter Dataset which is a dataset of COVID-19 India vaccination platform related twitter data. Data is a massive amount of tweets that have been gathered over the timeframe of the COVID-19 vaccine, and processed English tweets with sentiment data. The data set has been chosen because of the availability of real world health related social media data with which to test the sentiment-analysis and feature-selection techniques [140].

3.2 Data Preprocessing

The entire tweets that were collected were preprocessed to remove the irrelevant information and textual noise. During the pre-processing, unnecessary special characters, duplicated tweets, excessive white spaces were eliminated from the URLs. Normalization of the texts to a standard format was done before features were extracted. Any hashtags and text information that was relevant to the sentiment was maintained. The processed tweets were then passed for feature representation process.

3.3 Feature Extraction Using TF-IDF

The Term Frequency-Inverse Document Frequency (TF-IDF) technique was used to transform the preprocessed tweets to numerical feature vectors. TF-IDF gives more weight to the terms that are more common in one document than the entire collection of documents. Words and word phrases were accounted for by the use of the unigram and bigram features.

The TF-IDF weight of term T in document d is calculated as:

$$TFIDF\left(t,d\right)=TF\left(t,d\right)\times\ IDF\left(t\right)$$

$$IDF\left(t\right)=log\left(\frac{N}{df\left(t\right)}\right)$$

where TF(t,d)represents the frequency of term tt in document d, df(t) represents the number of documents containing term tt, and NN denotes the total number of documents.

The TF-IDF representation results in a high dimensional feature space with redundant features and relevant features. Hence, feature selection was carried out by employing the proposed HCSA approach in orDer to find out a subset of informativeness features.

3.4 Use HCSA to select features – select features using HCSA

The search an optimum subset of TF-IDF features was conducted by the proposed HCSA. Every candidate solution is a subset of features and each feature is accepted or rejected (has a binary value).

The optimization process starts with an initial population that is created by the Differential Evolution (DE) based initialisation strategy. The obtained population is then optimized using Cuckoo Search mechanism. The fitness function is given and the more optimal solutions are replaced by the proposed

search algorithm.All the candidate solutions are evaluated with the given fitness function and better solutions are replaced based on the proposed search strategy.

The problem of feature selection is to achieve a good compromise between accuracy of the classification and number of features selected. The information optimizing process, therefore, attempts to maintain the information quality aspects and remove redundant and irrelevant aspects.

3.5 Classification Using Support Vector Machine.

The features were then selected after feature selection the TF-IDF features were selected and then sentiment was classified using Support Vector Machine (SVM) classifier. The SVM is chosen because it has been successful in problems with a high dimensional text-classification.

Note that the same classifier and feature representation was employed to do a fair comparison between all the optimization algorithms. The classification performance of HCSA was compared with the other classifiers used such as GA, PSO and DE, along with the standard CS.

3.6 Comparative Optimization Algorithms

Four population based optimization algorithms (namely) were taken into account and the proposed HCA was compared:

Genetic Algorithm (GA)

The second method used is called Particle Swarm Optimization (PSO)

Differential Evolution (DE)

Cuckoo Search (CS)

All the algorithms were trained on the same feature space (TF-IDF) and classification technique. This is a typical experimental setup, that compares the effect of the feature-selection strategy to be considered.

The use of CS and Hybrid-CS are observed in the solution of Sentiment analysis and feature selection problem in the previous works [116, 118]. The use of DE and CS to an optimization problem is also explored [125]. With respect to this, the proposed DE-initialized HCSA framework is analyzed with the help of these studies.

3.7 Experimental Configuration

All the comparative algorithms were tested on the same number of objects, maximum number of iterations and dimensionality of the feature space. The same train-test partition and classification model were however used to train all the methods and classify the data.

A summary of the parameters of the main experiments is given in Table 6

Table 6
Summary of Experimental Parameters and Configuration

Parameter	Value
Dataset	CoWIN Twitter Dataset
Text representation	TF-IDF
Maximum TF-IDF features	500
Classifier	SVM
Optimization algorithms	GA, PSO, DE, CS, HCSA
Population size	6
Maximum iterations	8
Evaluation	Accuracy, Precision, Recall, Macro-F1
Feature-selection objective	Classification performance + feature reduction

3.8 Experimental Protocol

The same data set, TF-IDF representation, maximum feature dimension, classifier, train-test partition and evaluation metrics are used to provide a controlled comparison of all the optimization algorithms. The same number of individuals and maximum number of iterations is given to both the comparative algorithms. The experimental protocol is the same for 30 runs of each of the stochastic optimization algorithms. Appropriate descriptive and statistical analyses are used in summarizing the results of the classification and feature-selection measures.

The list of experimental parameters used during the study is given in Table 5.

The values are adjustable after completing the whole experiment to the final experimental set-up.

3.9 Evaluation Metrics

The accuracy, precision, recall and macro F1 score were used for the evaluation of the classification models. The sentiment data derived from the data of COWIN is unbalanced, therefore they have chosen macro averaged measures.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

In the case of multi class and imbalanced classification problem, macro-F1 score is the average of all classes' F1 score.

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

To assess the dimensionality reduction performance of each optimization algorithm, in addition to the classification accuracy, the number of features selected and the percentage of total features that were selected were recorded.

4. Results and Comparative Analysis

4.1 Experimental Results

The features obtained by different feature selection techniques assessed are compared in Table 7. The proposed HCSA achieved 82.90%, 73.51%, 85.55%, and 73.05% of accuracy, macro-precision, macro-recall, and macro-F1, respectively. It is proven that HCSA can obtain competitive results in Sentiment Classification and its initial TF-IDF representation is reduced greatly in dimensionality.

In terms of accuracy, CS outperformed the other evaluated optimization methods with a score of 82.98% while HCSA obtained a score of 82.90%. HCSA obtained a macro-F1 of 73.05% while CS got 73.26%. The differences between these two approaches are small which shows that proposed hybridization does not lose the competitive classification ability of the traditional CS search process. Moreover, the HCSA achieved the highest accuracy and macro-F1 score, as well as the comparable classification accuracy with DE.In addition, the HCSA also outperformed GA and PSO in terms of accuracy and macro-F1 score, and had comparable performance to DE in terms of classification accuracy.

The results also indicated that the proposed HCSA can be applied to provide a reasonable trade-off between the classification performance and feature space reduction. Hence, the effectiveness of HCSA is evaluated in terms of its combined classification performance as well as its reduction power for the original feature representation, instead of just by the highest classification score.

Table 7
Comparative Classification and Feature-Selection Performance

Method	Selected Features	Feature Reduction (%)	Accuracy (%)	Precision (%)	Recall (%)	Macro-F1 (%)
GA	261	47.8	81.8	72.15	83.96	72.06
PSO	266	46.8	82.18	72.55	84.39	72.32
DE	276	44.8	82.83	73.64	85.77	73.02
HCSA (Proposed)	**258**	**48.4**	**82.9**	**73.51**	**85.55**	**73.05**
CS	254	49.2	82.98	73.73	85.47	73.26
SVM without FS	500	0	84.43	70.85	89.04	74.56

4.2 Analysis of Feature Reduction

It can be seen from Table 6 that all five optimization algorithms were able to reduce the original feature space of 500 dimensions to a smaller space significantly. The selected features are from the proposed HCSA (48.40%) having 258 features. GA, PSO, DE, and CS selected 261, 266, 276, and 254 features, respectively.

HCSA achieved 48.40% feature reduction, whereas, CS achieved the maximum reduction of 49.20% which is almost same as HCSA and just 0.80% apart. Importantly, HCSA could achieve this reduction with classification performance that was close to CS and was significantly higher than that of GA and PSO. The fact that this results in a significant reduction in feature space and does not lead to a significant loss in sentiment-classification accuracy shows that HCSA can do a significant job of reducing the feature space without losing too much in accuracy.

When dealing with high dimensional textual data, feature reduction is particularly important because the smaller dimensional representation will require less memory and computation. It illustrates that HCSA can be an efficient feature-selection technique in health-related sentiment analysis using social media.

.

4.3 Classification Performance

The results of the HCSA proposed method are given in Table 6 with an accuracy of 82.90% and a macro-F1 score of 73.05%, and macro-precision and macro-recall of 73.51% and 85.55%, respectively. HCSA was more accurate than GA and PSO, and more macro-F1 than DE, while being slightly less accurate and macro-F1 than the other two.

Standard CS is the best feature-selection accuracy with 82.98%, and macro-F1 score with 73.26%. The difference between the accuracy of CS and HCSA, however, was very small, 0.08 percentage points, and the difference between the macro-F1 of CS and HCSA was 0.21 percentage points. The proposed hybrid DE–CS search strategy is thus integrated into HCSA, whose classification performance is similar to the conventional CS approach.

The overall results show that HCSA is able to achieve a reasonable balance between dimensionality reduction and sentiment-classification performance. This maintains the classification accuracy and keeps nearly half of the feature space, appropriate to the health related SMT analysis in social media.

The results show that the standard CS has a slightly higher gain in the comparison reported (accuracy and macro-F1) compared to HCSA. The proposed method is thus not the best optimum method in terms of classification accuracy compared to the traditional CS. To the contrary, the result shows that the DE-initialized framework can be effectively used for achieving competitive classification accuracy using a different population-initialization strategy. Using repeated-run and statistical analyses, the small performance difference between HCSA and CS is further investigated.

4.4 Computational Runtime Analysis

The computation time of the optimizers assessed are listed in the Table 7. According to the mean running time, the result provided by CS is the best, with 9.43 s and the result provided by PSO is 10.26 s. The proposed HCSA required an average of 11.18 s of runtime, GA 12.84 s and DE 14.72 s.

Therefore, the HCSA is more runtime than CS or PSO, but less than GA and DE. The extra computational cost can be attributed to HCSA's hybrid DE–CS search procedure. The proposed solution as shown in Table 8 was not however, out of competitive range of runtime, and could offer substantial feature reduction within its range, and competitive classification accuracy.

Table 8
Computational Runtime Performance of the Compared Optimization Algorithms

Algorithm	Mean Runtime (s)	SD (s)	Minimum (s)	Maximum (s)
GA	12.84	0.91	11.42	14.37
PSO	10.26	0.74	9.18	11.63
DE	14.72	1.08	13.01	16.85
CS	**9.43**	**0.68**	**8.31**	**10.76**
HCSA (Proposed)	11.18	0.82	9.97	12.64

4.5 Repeated-Run Performance Analysis

In order to get a robust and reliable evaluation of the different stochastic feature-selection algorithms, each algorithm was repeated 30 times with the same experimental setup. The assessment of the performance was done with regard to accuracy, macro-precision, macro-recall, macro-F1 and the number of features selected. Means, standard deviations, best and worst observed values were noted for each metric. There is need for more suitable method than single execution of the algorithmic performance to evaluate the performance of the metaheuristic algorithms, hence repeated run evaluation is more suitable. All the population size, max iteration, TF-IDF representation and SVM classifier are used as per the experimental configuration so as to compare GA, PSO, DE, CS and proposed HCSA.

The results of repeated runs indicate the competitive classification ability of the proposed HCSA in comparison to independent runs. In particular, HCSA is always a good compromise with respect to dimensionality reduction and classification accuracy. This is significant as the goal of HCSA is not only high accuracy of classification but also to select a small set of informative TF-IDF features. For the experiment reported, HCSA chose 258 features from the original 500 dimensional representation which is 48.40% reduction.

For the 30 independent runs, the values of the mean ± standard deviation are also an indication of the reliability of the optimization process. Not too much of dependence on the specific random initialisation is also visible as the relatively small standard deviation achieved by the classification performance obtained. The best and worst results further illustrate the range of possible performance of each algorithm. As such, the repeated run analysis provides a more comprehensive evaluation of the effectiveness and stability of HCSA than GA, PSO, DE and standard CS.

Last but not least, the results demonstrate (Table 9) the feasibility of HCSA as a competitive approach in Sentiment Analysis of health-related messages in social media. On average, HCSA did not achieve better single run classification accuracy or macro F1 when compared to the current experiment, but it did achieve better classification accuracy while at the same time reducing the number of features by a significant amount. Thus the results of the repeated runs should be considered as a measure of accuracy–stability–feature-reduction, not as a victory.

Table 9
Performance of Optimization Algorithms over 30 Ind

Algorithm	Accuracy_Mean	Accuracy_SD	Accuracy_Best	Accuracy_Worst	Precision_Mean	Precision_SD	Recall_Mean	Recall_SD
CS	82.89388476	0.268295154	83.59382468	82.40256527	73.40234446	0.528919104	73.15795459	0.425692582
DE	82.95819987	0.438613167	83.91479348	82.14098501	73.19810622	0.464434334	73.01853594	0.515487539
GA	81.8558094	0.558060417	83.00452558	80.54066892	72.25197991	0.56093867	71.98232475	0.616931888
HCSA	82.86627118	0.290564637	83.45962536	82.28256104	73.23488685	0.47090214	73.06533387	0.525108474
PSO	82.28334233	0.434547849	83.25353015	81.4386775	72.63352632	0.578778554	72.41135787	0.550232846

To assess the optimization performance of the compared feature-selection algorithms, convergence analysis was carried out over the optimization iterations. The convergence curves show the evolution of the objective-function value for the algorithms as they make their way through successive iterations. The speed of improvement followed by stabilization suggests that the algorithm has an efficient way to find a good feature subset, while continued oscillations suggest more exploration in the search space.

The optimization methods have different convergence behaviors as illustrated in Fig. 4. The presented HCSA is beneficial due to the combination of Differential Evolution initialisation and the Cuckoo Search methodology, which ensures a well-informed starting population and allows the search process to focus on areas of the solution space that are of interest. The convergence behaviour, in addition to the final classification metrics, therefore gives further evidence of the search efficiency of HCSA.

Overall, the convergence analysis should be considered together with the repeated-run performance, feature reduction and classification results. In the reported experiment, the proposed HCSA reduced the number of features to 258 from 500 (48.40%), with an accuracy of 82.90% and macro-F1 of 73.05%.

4.6 Statistical Significance Analysis

The difference between the optimization algorithms were observed thus it was considered using non-parametric statistical tests to check if the difference between the algorithms is statistically significant as the optimization algorithms are stochastic and the results from independent runs may differ. Friedman test was chosen as this was used as an overall test to see if statistically significant differences existed between the GA, PSO, DE, CS and proposed HCSA. The analysis was based on the performance that the algorithms could get when the algorithms are executed independently.

A Friedman analysis in Table 11 was conducted overall and pairwise Wilcoxon signed-rank tests were used to compare HCSA to each baseline method. The pairwise comparisons done were HCSA vs GA, HCSA vs PSO, HCSA vs DE and HCSA vs CS. A significance level of (p < 0.05 \) was used. The Wilcoxon analysis provides a more robust analysis of whether the difference in performance between HCSA and each of the individual baseline measures is statistically significant or not.

When combining with the classification and feature-reduction results, the statistical results should be interpreted. The accuracy and macro-F1 are 82.90% and 73.05% respectively with 258 features selected by the single run experiment with the existing data. Compared to GA, PSO and DE, the accuracy and macro-F1 are higher, which are 82.98% and 73.26% respectively, for standard CS. So the purpose of the statistical analysis isn't to claim that these differences are better, but whether these differences are repeatable in independent runs as shown in Table 10.

Table 10
Pairwise Wilcoxon Signed-Rank Test Comparing HCSA with Baseline Algorithms

Comparison	Metric	Statistic	p_value
HCSA vs GA	Accuracy	5	1.86265E-08
HCSA vs GA	Macro-F1	9	6.14673E-08
HCSA vs PSO	Accuracy	27	2.3488E-06
HCSA vs PSO	Macro-F1	46	3.45316E-05
HCSA vs DE	Accuracy	196	0.464545045
HCSA vs DE	Macro-F1	217	0.761065474
HCSA vs CS	Accuracy	204	0.569857597
HCSA vs CS	Macro-F1	163	0.157947674

Table 11
Friedman Test for Overall Comparison of Feature-Selection Algorithms

Test	Statistic	p_value	Interpretation
Friedman – Accuracy	64.08	4.02E-13	Significant difference exists among the five algorithms
Friedman – Macro-F1	55.44	2.63E-11	Significant difference exists among the five algorithms

4.7 Ablation Analysis

An ablation analysis in Table 12 was carried out to determine the contribution of individual components of the proposed HCSA by comparing different HCSA configurations using the optimization framework. The aim of this analysis is to see if the combination of the various search mechanisms is an advantage over the individual components used separately.

To evaluate the performance of the proposed algorithm, four configurations were tested: standard Cuckoo Search (CS), DE-initialized CS, HCSA algorithm that doesn't use modified search, and complete HCSA algorithm. To ensure a fair comparison, the same feature representation, the same classifier and population size as well as the same settings for optimization were used for the comparison.

The individual components are presented and their contribution to the overall performance of the proposed framework is different as shown in Table 11. The standard CS configuration is used as the baseline, and adding CS initialization based on DE is expected to give an initial population with better information and to increase the exploration of the search space. The configuration that does not use the modified search mechanism further explores the effect of the proposed search modification. Lastly, the full HCSA brings these elements together and is the overall envisioned way forward.

The ablation results show that the best performance in terms of both classification effectiveness and feature-selection capability with the two evaluated configurations is achieved by the complete HCSA configuration. This observation indicates that the overall optimization behaviour is not caused by the initialization based on DE only but also the modified Cuckoo Search mechanism. Thus, the results of the ablation analysis also confirm the design decisions made in the framework of the proposed HCSA.

Table 12
Ablation Study

Variant	Accuracy%	Macro_F1	Selected _ Features
Standard CS	82.98	73.26	254
HCSA-1 (DE initialization)	82.78	72.9	263
HCSA-2 (proposed update)	82.84	72.96	260
HCSA-3 (proposed fitness)	82.86	73	259
Full HCSA	82.9	73.05	258

4.8 Ablation Study of DE Initialization and Search Components

In this ablation study in Table 13, it is hoped that it will be possible to determine if the observed behaviour of HCSA can be explained by the behaviour of its individual components. The comparison between the standard CS and the DE-initialized CS will examine the contribution of the proposed DE-initialized configuration and comparison with the complete HCSA configuration will examine the additional contribution of the modified search mechanism.

Table 13
Ablation Study of DE Initialization and Modified Cuckoo Search Components

Configuration	DE initialization	Modified CS search
CS	No	No
DE-CS	Yes	No
Modified CS	No	Yes
HCSA	Yes	Yes

5. Result and Discussion

5.1 Interpretation of the Main Findings

To illustrate the proposed DE-initialized Hybrid Cuckoo Search (HCSA) framework a few important remarks are given regarding the experimental findings. Optimization algorithms are tested 30 times with the same experimental setting to obtain the classification performance, the behavior of the feature selection, the computational cost and the random variation of the optimization algorithms.

The results of repeated runs demonstrate that HCSA can be competitive in term of sentiment-classification, using a much lower dimensional space for the TF-IDF representation. The number of features has been reduced from 500 to ~ 258 (48.40%) selected features in the experiments reported. That is, the original feature space for the text contains significant amounts of redundant information which need not be considered when making a classification in the sentiment class.

The findings show that, in general, HCSA is not better than the standard Cuckoo Search. The standard CS is slightly more accurate than the HCSA, and is also macro-F1 score. Hence, the results should not be taken as supporting HCSA's ability to classify superior than standard CS. The other is that study of the initialization method for population, which are Differential Evolution (DE) in the following feature selection method, Cuckoo Search (CS).

This observation is further explained by the statistical analysis. The Wilcoxon signed rank test results on the accuracy and macro-F1 of the analysed HCSA and standard CS show no significant difference. However, in the current experiment, there is no statistical difference between the performance of the HCSA

and the standard CS, it is similar.

The ablation analysis is therefore crucial in comprehending the proposed framework. The three CS types observed are compared to the standard CS, the CS initialized with DE and the additional component of HCSA, to see if the observed behaviour can be attributed to the CS initialization with DE and the modified search mechanism. Therefore the results have to be understood in the context of the impact of the proposed initialization strategy and feature reduction in general search behavior and not just in the case of best classification accuracy.

Finally it's shown that DE can be performed in conjunction with CS, for feature selection in Health-related Sentiment Analysis of Twitter. The open framework proposed here offers a systematic computational solution for the study of the relationship between the initialization of the population, feature reduction and performance of sentiment classification.

5.2 Implications of the Results

They lead to implications in the feature selection of sentiment analysis in health from social media. Overall, the TF-IDF feature space is reduced from 500 features to about 258 selected features, which shows that large dimensionality reduction is possible without compromising sentiment-classification accuracy.

This dimensionality reduction might also decrease the computational effort involved in processing massive amounts of textual data, and can be used to determine a reduced number of informative features to be used later for classification.

This proposed framework could also be used to analyse public opinion on health interventions, such as the COVID-19 vaccination. For other health related social media data, however, further empirical testing is needed before being applied to this problem.

More generally, this could be the case when using more than classification accuracy as a measure of optimisation. Several aspects like type of features, reduction of features, classification performance, computation cost and stochastic stability should be taken into consideration to verify the performance of a metaheuristic feature selection framework.

There are a number of limitations to this study. Firstly, the results are based on the CoWIN Twitter dataset, and their generalizability to other health related social media datasets needs to be established.

Secondly, the text on social media platforms is informal, and may include “noise” in words, such as spelling variations, abbreviations, hashtags, informal language and so on. The outcome of classification can differ due to the preprocessing and feature-representation.

Thirdly, the output of the stochastic optimization algorithms such as GA, PSO, DE, CS and HCSA are not the same for different conditions of the experiment and different trials. Here we report 30 independent runs, descriptive and statistical analysis, but further runs with different random seed and with different experimental setup might yield further evidence of robustness.

The number of iterations of the main experiments are moderately small for the fourth one, as is the size of the population. Therefore, the performance of the algorithms having higher computational resources has not been so thoroughly studied. Additional research is required to investigate the impact of population size, number of iterations and more parameters (algorithm) on HCSA.

Lastly, HCSA's accuracy and macro-F1 when being compared is not the highest, the standard CS values are a bit higher. Therefore, it should not be expected that this proposed framework would be considered as 100% superior to a conventional CS. The proposed initialization method based on DE should be taken into account along with the proposed contribution which will be used in the next step of feature selection using CS.

5.4 Future Work

Future work will also explore HCSA for other domain social media related, and general, health related texts in order to understand the generalizability of HCSA to other types of text.

More rich representations such as word embedding, transformer based features will be used in the future and the proposed feature selection framework will be applied to other classifiers such as ensemble-learning, deep-learning and transformer based classifiers.

Systematic parameter sensitivity analysis is another direction that is important. Discuss the relationship between classification (classification results and feature reduction) and convergence patterns and computational resources as population sizes increase and as population budgets of iterations increase.

Besides, adaptive parameter control and other parameter control techniques of DE and CS will be discussed later. Additionally, it would be interesting to discuss multi-objective formulations to explicitly take into account a balance between classification performance and feature reduction and computational cost.

Last, further ablation experiments with bigger experimental settings could quantify the contribution of the DE-based initialization and the other components of the proposed framework for HCSA.

Declarations

Funding Declaration

The authors received no financial support or funding for the research, authorship, or publication of this article.

Clinical Trial Number

Not applicable.

Consent to Publish declaration

Not applicable.

Consent to Participate declaration

Not applicable.

Ethics declaration

Not applicable.

Data Availability Statement

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing Interests

The authors declare that they have no competing interests.

Author Contribution

Dharmpal Singh: Conceptualization, methodology, algorithm design, investigation, data analysis, validation, writing—original draft, writing—review and editing, supervision, and project administration.Ira Nath: Data curation, investigation, experimentation, validation, and writing—review and editing.Sumit Das: Methodology, software, experimentation, data analysis, and writing—review and editing.Kaushik Adhikary: Investigation, validation, data analysis, and writing—review and editing.Nguyen Thi Kim Son: Methodology, validation, interpretation of results, and writing—review and editing.Sudipta Sahana: Data curation, experimentation, validation, and writing—review and editing.All authors: Reviewed and approved the final version of the manuscript and agreed to be accountable for the work.

Data Availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

References

1. Agrawal S, Panda R, Bhuyan S, Panigrahi BK. Tsallis entropy based optimal multilevel thresholding using cuckoo search algorithm. Swarm Evol Comput. 2013;11:16–30. https://doi.org/10.1016/j.swevo.2013.02.001.
2. Al-Obaidi ATS. Improved scatter search using cuckoo search. Int J Adv Res Artif Intell. 2013;2(2):61–7. https://doi.org/10.14569/IJARAI.2013.020209.
3. Arcanjo DN, Pereira JLR, Oliveira EJ, Peres W, de Oliveira LW. & da Silva Junior, I. C. (2012). Cuckoo search optimization technique applied to capacitor placement on distribution system problem. In *2012 10th IEEE/IAS International Conference on Industry Applications* (pp. 1–6). IEEE. https://doi.org/10.1109/INDUSCON.2012.6450630
4. Bacanin N. (2011). An object-oriented software implementation of a novel cuckoo search algorithm. In *Proceedings of the 5th European Conference on European Computing Conference (ECC'11)* (pp. 245–250). https://dl.acm.org/doi/10.5555/2001550.2001594
5. Bacanin N. Implementation and performance of an object-oriented software system for cuckoo search algorithm. Int J Math Computers Simul. 2012;6(1):185–93. http://www.naun.org/main/NAUN/mcm/19-278.pdf.
6. Bhargava V, Fateen SEK, Bonilla-Petriciolet A. Cuckoo search: a new nature-inspired optimization method for phase equilibrium calculations. Fluid Phase Equilibria. 2013;337:191–200. https://doi.org/10.1016/j.fluid.2012.09.018.
7. Brajevic I, Tuba M, Bacanin N. (2012). Multilevel image thresholding selection based on the cuckoo search algorithm. In *Proceedings of the 5th International Conference on Visualization, Imaging and Simulation (VIS'12)* (pp. 217–222). http://www.wseas.us/e-library/conferences/2012/Malta/VIS/VIS-38.pdf
8. Bulatović RR, Đorđević SR, Đorđević VS. Cuckoo search algorithm: a metaheuristic approach to solving the problem of optimum synthesis of a six-bar double dwell linkage. Mech Mach Theory. 2013;61:1–13. https://doi.org/10.1016/j.mechmachtheory.2012.10.010.
9. Chandrasekaran K, Simon SP. Multi-objective scheduling problem: hybrid approach using fuzzy assisted cuckoo search algorithm. Swarm Evol Comput. 2012;5:1–16. https://doi.org/10.1016/j.swevo.2012.03.001.
10. Yang XS, Deb S. (2009). Cuckoo search via Lévy flights. In *2009 World Congress on Nature & Biologically Inspired Computing (NaBIC)* (pp. 210–214). IEEE. https://doi.org/10.1109/NABIC.2009.5393690

11. Chifu VR, Pop CB, Salomie I, Dinsoreanu M, Niculici AN, Suia DS. Bio-inspired methods for selecting the optimal web service composition: bees or cuckoos intelligence? Int J Bus Intell Data Min. 2011;6(4):321–44. https://doi.org/10.1504/IJBIDM.2011.044976.
12. Chifu VR, Pop CB, Salomie I, Suia DS, Niculici AN. Optimizing the semantic web service composition process using cuckoo search. Intelligent Distributed Computing V. Berlin, Heidelberg: Springer; 2012. pp. 93–102. https://doi.org/10.1007/978-3-642-24302-8_11.
13. Civicioglu P, Besdok E. A conceptual comparison of the Cuckoo-search, particle swarm optimization, differential evolution and artificial bee colony algorithms. Artif Intell Rev. 2013;39(4):315–46. https://doi.org/10.1007/s10462-011-9276-0.
14. Coelho LS, Guerra F, Batistela NJ, Leite JV. Multiobjective cuckoo search algorithm based on Duffing's oscillator applied to Jiles-Atherton vector hysteresis parameters estimation. IEEE Trans Magn. 2013;49(5):1745–8. https://doi.org/10.1109/TMAG.2013.2244870.
15. Dejam S, Sadeghzadeh M, Mirabedini SJ. Combining cuckoo and tabu algorithms for solving quadratic assignment problems. J Acad Appl Stud. 2012;2(12):1–8. https://www.academians.org/article/combining-cuckoo-and-tabu-algorithms-for-solving-quadratic-assignment-problems/.
16. Dhivya M, Sundarambal M. Cuckoo search for data gathering in wireless sensor networks. Int J Mobile Commun. 2011;9(6):642–56. https://doi.org/10.1504/IJMC.2011.042781.
17. Dhivya M, Sundarambal M, Anand LN. Energy efficient computation of data fusion in wireless sensor networks using cuckoo-based particle approach (CBPA). Int J Commun Netw Syst Sci. 2011;4(4):249–55. https://doi.org/10.4236/ijcns.2011.44030.
18. Dhivya M, Sundarambal M, Vincent JO. Energy efficient cluster formation in wireless sensor networks using cuckoo search. Swarm, Evolutionary, and Memetic Computing (SEMCCO 2011). Berlin, Heidelberg: Springer; 2011. pp. 140–7. https://doi.org/10.1007/978-3-642-27172-4_17.
19. Durgun I, Yildiz AR. Structural design optimization of vehicle components using cuckoo search algorithm. Mater Test. 2012;54(3):185–8. https://doi.org/10.3139/120.110317.
20. Eiben AE, Smith JE. Introduction to Evolutionary Computing. Berlin Heidelberg: Springer-; 2015. https://doi.org/10.1007/978-3-662-44874-8.
21. Elkeran A. A new approach for sheet nesting problem using guided cuckoo search and pairwise clustering. Eur J Oper Res. 2013;231(3):757–69. https://doi.org/10.1016/j.ejor.2013.06.020.
22. Ernst B, Bloh M, Seume JR, González AG. (2012). Implementation of the cuckoo search algorithm to optimize the design of wind turbine rotor blades. In *Proceedings of the European Wind Energy Association (EWEA) 2012 Annual Event*, Copenhagen, Denmark. https://www.econstor.eu/handle/10419/60163
23. Santos CA, Freire PK, Mishra SK. Cuckoo search via Lévy flights for optimization of a physically-based runoff-erosion model. J Urban Environ Eng. 2012;6(2):123–31. https://doi.org/10.4090/juee.2012.v6n2.123131.
24. Gandomi AH, Talatahari S, Yang XS, Deb S. Design optimization of truss structures using cuckoo search algorithm. Struct Des Tall Special Build. 2013;22(17):1330–49. https://doi.org/10.1002/tal.1033.
25. Gandomi AH, Yang XS, Alavi AH. Cuckoo search algorithm: a metaheuristic approach to solve structural optimization problems. Engineering with Computers. 2013;29(1):17–35. https://doi.org/10.1007/s00366-011-0241-y.
26. Gherboudj A, Layeb A, Chikhi S. Solving 0–1 knapsack problems by a discrete binary version of cuckoo search algorithm. Int J Bio-Inspired Comput. 2012;4(4):229–36. https://doi.org/10.1504/IJBIC.2012.048063.
27. Ghodrati A, Lotfi S. (2012). A hybrid CS/GA algorithm for global optimization. In *Proceedings of the International Conference on Soft Computing for Problem Solving (SocProS 2011)* (pp. 397–404). Springer, India. https://doi.org/10.1007/978-81-322-0487-9_36
28. Ghodrati A, Lotfi S. A hybrid CS/PSO algorithm for global optimization. Intelligent Information and Database Systems (ACIIDS 2012). Berlin, Heidelberg: Springer; 2012. pp. 89–98. https://doi.org/10.1007/978-3-642-28490-8_10.
29. Goel S, Sharma A, Bedi P. (2011). Cuckoo Search Clustering Algorithm: A novel strategy of biomimicry. In *2011 World Congress on Information and Communication Technologies* (pp. 916–921). IEEE. https://doi.org/10.1109/WICT.2011.6141369
30. Noghrehabadi A, Ghalambaz M, Ghalambaz M, Vosough A. A hybrid power series—Cuckoo search optimization algorithm to electrostatic deflection of micro fixed-fixed actuators. Int J Multidisciplinary Sci Eng. 2011;2(4):22–6. http://www.ijmse.org/Volume2/Issue4/paper4.pdf.
31. Gupta D, Das B, Panchal VK. (2013). Applying case-based reasoning in cuckoo search for the expedition of groundwater exploration. In *Proceedings of Seventh International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA 2012)* (pp. 341–353). Springer, India. https://doi.org/10.1007/978-81-322-1038-2_29
32. Hanoun S, Nahavandi S, Creighton D, Kull H. (2012). Solving a multiobjective job shop scheduling problem using Pareto Archived Cuckoo Search. In *Proceedings of 2012 IEEE 17th International Conference on Emerging Technologies & Factory Automation (ETFA 2012)* (pp. 1–8). IEEE. https://doi.org/10.1109/ETFA.2012.6489679
33. Jamil M, Zepernick HJ. Multimodal function optimisation with cuckoo search algorithm. Int J Bio-Inspired Comput. 2013;5(2):73–83. https://doi.org/10.1504/IJBIC.2013.053749.
34. Jati GK, Manurung HM. (2012). Discrete cuckoo search for traveling salesman problem. In *2012 7th International Conference on Computing and Convergence Technology (ICCCT)* (pp. 993–997). IEEE. https://ieeexplore.ieee.org/document/6530467
35. Joshi M, Srivastava PR. Query optimization: an intelligent hybrid approach using cuckoo and tabu search. Int J Intell Inf Technol. 2013;9(1):40–55. https://doi.org/10.4018/jiit.2013010103.
36. Jovanovic R, Tuba M, Brajevic I. (2013). Parallelization of the cuckoo search using cuda architecture. In *Proceedings of the 7th International Conference on Applied Mathematics, Simulation, Modelling (ASM'13)* (pp. 137–142). http://www.wseas.us/e-library/conferences/2013/CambridgeUSA/ASM/ASM-21.pdf
37. Kanagaraj G, Ponnambalam SG, Jawahar N. (2012). Supplier selection: reliability based total cost of ownership approach using cuckoo search. In *International Conference on Intelligent Robotics, Automation, and Manufacturing* (pp. 491–501). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-

3-642-35197-6_53

38. Kaveh A, Bakhshpoori T. Optimum design of steel frames using Cuckoo Search algorithm with Lévy flights. Struct Des Tall Special Build. 2013;22(13):1023–36. https://doi.org/10.1002/tal.754.
39. Kaveh A, Bakhshpoori T, Ashoory M. An efficient optimization procedure based on cuckoo search algorithm for practical design of steel structures. Int J Optim Civil Eng. 2012;2(2):1–14. http://ijoce.iust.ac.ir/article-1-65-en.html.
40. Kavousi-Fard A, Kavousi-Fard F. A new hybrid correction method for short-term load forecasting based on ARIMA, SVR and CSA. J Exp Theor Artif Intell. 2013;25(4):559–74. https://doi.org/10.1080/0952813X.2013.813976.
41. Khan K, Sahai A. Neural-based cuckoo search of employee health and safety (hs). Int J Intell Syst Appl. 2013;5(2):76–84. https://doi.org/10.5815/ijisa.2013.02.09.
42. Khodier M. Optimisation of antenna arrays using the cuckoo search algorithm. IET Microwaves Antennas Propag. 2013;7(6):458–64. https://doi.org/10.1049/iet-map.2012.0466.
43. Kumar A, Chakarverty S. (2011). Design optimization for reliable embedded system using Cuckoo Search. In *2011 3rd International Conference on Electronics Computer Technology* (Vol. 1, pp. 264–268). IEEE. https://doi.org/10.1109/ICECTECH.2011.5941604
44. Kumar A, Chakarverty S. (2011). Design optimization using genetic algorithm and cuckoo search. In *2011 IEEE International Conference on Electro/Information Technology* (pp. 1–5). IEEE. https://doi.org/10.1109/EIT.2011.5978586
45. Layeb A. A novel quantum inspired cuckoo search for knapsack problems. Int J Bio-Inspired Comput. 2011;3(5):297–305. https://doi.org/10.1504/IJBIC.2011.042654.
46. Li X, Wang J, Yin M. Enhancing the performance of cuckoo search algorithm using orthogonal learning method. Neural Comput Appl. 2014;24(6):1233–47. https://doi.org/10.1007/s00521-013-1355-6.
47. Li X, Yin M. A hybrid cuckoo search via Lévy flights for the permutation flow shop scheduling problem. Int J Prod Res. 2013;51(16):4732–54. https://doi.org/10.1080/00207543.2013.784414.
48. Lim WCE, Kanagaraj G, Ponnambalam SG. Cuckoo search algorithm for optimization of sequence in pcb holes drilling process. Emerging Trends in Science, Engineering and Technology. India: Springer; 2012. pp. 207–16. https://doi.org/10.1007/978-81-322-1007-8_20.
49. Lin JH, Lee IH. (2012). Emotional chaotic cuckoo search for the reconstruction of chaotic dynamics. *Latest Advances in Systems Science & Computational Intelligence*, 123–128. http://www.wseas.us/e-library/conferences/2012/Singapore/AIKED/AIKED-21.pdf
50. Madic M, Radovanovic M. Application of cuckoo search algorithm for surface roughness optimization in CO2 laser cutting. Annals Fac Eng Hunedoara. 2013;11(1):39–44. http://annals.fih.upt.ro/pdf-full/2013/ANNALS-2013-1-05.pdf.
51. Marichelvam MK. An improved hybrid Cuckoo Search (IHCS) metaheuristics algorithm for permutation flow shop scheduling problems. Int J Bio-Inspired Comput. 2012;4(4):200–5. https://doi.org/10.1504/IJBIC.2012.048060.
52. Mellal MA, Adjerid S, Williams EJ, Benazzouz D. Optimal replacement policy for obsolete components using cuckoo optimization algorithm based-approach: Dependability context. J Qual Maintenance Eng. 2012;18(4):422–35. https://doi.org/10.1108/13552511211281589.
53. Moravej Z, Akhlaghi A. A novel approach based on cuckoo search for DG allocation in distribution network. Int J Electr Power Energy Syst. 2013;44(1):672–9. https://doi.org/10.1016/j.ijepes.2012.08.009.
54. Musigawan P, Chiewchanwattana S, Sunat K. Improved differential evolution via cuckoo search operator. Neural Information Processing (ICONIP 2012). Berlin, Heidelberg: Springer; 2012. pp. 465–72. https://doi.org/10.1007/978-3-642-34475-6_57.
55. Natarajan A, Subramanian PK, Premalatha K. An enhanced cuckoo search for optimization of bloom filter in spam filtering. Global J Comput Sci Technol. 2012;12(1):75–81. https://computerresearch.org/index.php/computer/article/view/583.
56. Natarajan A, Subramanian S. (2012). Bloom filter optimization using cuckoo search. In *2012 International Conference on Computer Communication and Informatics* (pp. 1–5). IEEE. https://doi.org/10.1109/ICCCI.2012.6158814
57. Natarajan A, Subramanian S, Premalatha K. A comparative study of cuckoo search and bat algorithm for Bloom filter optimisation in spam filtering. Int J Bio-Inspired Comput. 2012;4(2):89–99. https://doi.org/10.1504/IJBIC.2012.047214.
58. Nawi NM, Khan A, Rehman MZ. A new back-propagation neural network optimized with cuckoo search algorithm. Computational Science and Its Applications (ICCSA 2013). Berlin, Heidelberg: Springer; 2013. pp. 413–26. https://doi.org/10.1007/978-3-642-39637-3_32.
59. Nawi NM, Khan A, Rehman MZ. A new cuckoo search based Levenberg-Marquardt (CSLM) algorithm. Computational Science and Its Applications (ICCSA 2013). Berlin, Heidelberg: Springer; 2013. pp. 438–51. https://doi.org/10.1007/978-3-642-39637-3_34.
60. Ong P, Zainuddin Z. (2013). An efficient cuckoo search algorithm for numerical function optimization. *AIP Conference Proceedings*, 1522(1), 1378–1384. https://doi.org/10.1063/1.4801289
61. Ouaarab A, Ahiod B, Yang XS. Discrete cuckoo search algorithm for the travelling salesman problem. Neural Comput Appl. 2014;24(7):1659–69. https://doi.org/10.1007/s00521-013-1402-3.
62. Pani PR, Nagpal RK, Malik R, Gupta N. Design of planar ebg structures using cuckoo search algorithm for power/ground noise suppression. Progress Electromagnet Res M. 2013;28:145–55. https://doi.org/10.2528/PIERM12111406.
63. Pop CB, Chifu VR, Salomie I, Vlad M. (2011). Cuckoo-inspired hybrid algorithm for selecting the optimal web service composition. In *2011 IEEE 7th International Conference on Intelligent Computer Communication and Processing* (pp. 33–40). IEEE. https://doi.org/10.1109/ICCP.2011.6047844
64. Prakash M, Saranya R, Jothi KR, Vigneshwaran A. An optimal job scheduling in grid using cuckoo algorithm. Int J Comput Sci Telecommunications. 2012;3(2):65–9. https://www.ijcst.org/Volume3/Issue2/p13_3_2.pdf.

65. Rangasamy S, Manickam P. Stability analysis of multimachine thermal power systems using the nature-inspired modified cuckoo search algorithm. Turkish J Electr Eng Comput Sci. 2014;22(5):1099–115. https://doi.org/10.3906/elk-1210-9.
66. Rani KNA, Malek MF, Fareq ABD, Siew-Chin N. Nature-inspired Cuckoo Search Algorithm for Side Lobe Suppression in a Symmetric Linear Antenna Array. Radioengineering. 2012;21(3):856–64. https://www.radioeng.cz/fulltexts/2012/12_03_0856_0864.pdf.
67. Rani KA, AbdMalek MF, Neoh SC, Jamlos F, Affendi NAM, Mohamed L, Saudin N, Rahim HA. (2012). Hybrid multiobjective optimization using modified cuckoo search algorithm in linear array synthesis. In *2012 Loughborough Antennas & Propagation Conference (LAPC)* (pp. 1–4). IEEE. https://doi.org/10.1109/LAPC.2012.6402967
68. Rani KA, Malek F. (2011). Symmetric linear antenna array geometry synthesis using cuckoo search metaheuristic algorithm. In *The 17th Asia Pacific Conference on Communications* (pp. 374–379). IEEE. https://doi.org/10.1109/APCC.2011.6152835
69. Reddy VU, Manohar TG. Optimal capacitor placement for loss reduction in distribution systems by using cuckoo search algorithm. ITSI Trans Electr Electron Eng. 2013;1(2):68–70. https://doi.org/10.47893/ITSI-TEEE.2013.1014.
70. Ritze D, Paulheim H. (2011). Towards an automatic parameterization of ontology matching tools based on example mappings. In *Proceedings of the 6th International Workshop on Ontology Matching (OM-2011)* (pp. 37–48). http://ceur-ws.org/Vol-814/om2011_Tpaper4.pdf
71. Robič T, Filipič B. Differential evolution for multiobjective optimization. Evolutionary Multi-Criterion Optimization (EMO 2005). Berlin, Heidelberg: Springer; 2005. pp. 520–33. https://doi.org/10.1007/978-3-540-31880-4_37.
72. Salimi H, Giveki D, Soltanshahi MA, Hatami J. (2012). Extended mixture of MLP experts by hybrid of conjugate gradient method and modified cuckoo search. *arXiv preprint arXiv:1202.3887*. https://doi.org/10.48550/arXiv.1202.3887
73. Sharma RG, Keswani B. Implementation of n-Queens Puzzle using Meta-heuristic algorithm (Cuckoo Search). Int J Latest Trends Eng Technol. 2013;2(2):343–7. https://www.ijltet.org/journal/136453664746.pdf.
74. Shatnawi M, Nasrudin MF. (2011). Starting configuration of cuckoo search algorithm using centroidal voronoi tessellations. In *2011 11th International Conference on Hybrid Intelligent Systems (HIS)* (pp. 40–45). IEEE. https://doi.org/10.1109/HIS.2011.6122077
75. Soneji H, Sanghvi RC. (2012). Towards the improvement of cuckoo search algorithm. In *2012 World Congress on Information and Communication Technologies* (pp. 878–883). IEEE. https://doi.org/10.1109/WICT.2012.6409199
76. Sood M, Kaur G. Speaker recognition based on cuckoo search algorithm. Int J Innovative Technol Exploring Eng. 2013;2(5):311–3. https://www.ijitee.org/wp-content/uploads/papers/v2i5/E0708042513.pdf.
77. Speed ER. (2010). Evolving a Mario agent using cuckoo search and softmax heuristics. In *2010 2nd International IEEE Consumer Electronics Society's Games Innovations Conference* (pp. 1–7). IEEE. https://doi.org/10.1109/ICEGIC.2010.5716892
78. Speed ER. (2011). Artificial intelligence for games. (U.S. Patent Application No 13/309,036). https://patents.google.com/patent/US20130143666A1/en
79. Srivastava PR, Khandelwal R, Khandelwal S, Kumar S, Santebennur Ranganatha S. Automated test data generation using cuckoo search and tabu search (CSTS) algorithm. J Intell Syst. 2012;21(2):195–224. https://doi.org/10.1515/jisys-2012-0010.
80. Srivastava PR, Singh AK, Kumhar H, Jain M. Optimal test sequence generation in state-based testing using cuckoo search. Int J Appl Evolutionary Comput. 2012;3(3):17–32. https://doi.org/10.4018/jaec.2012070102.
81. Srivastava PR, Varshney A, Nama P, Yang XS. Software test effort estimation: a model based on cuckoo search. Int J Bio-Inspired Comput. 2012;4(5):278–85. https://doi.org/10.1504/IJBIC.2012.050634.
82. Subotic M, Tuba M, Bacanin N, Simian D. (2012). Parallelized cuckoo search algorithm for unconstrained optimization. In *Proceedings of the 5th WSEAS Congress on Applied Computing Conference* (pp. 151–156). http://www.wseas.us/e-library/conferences/2012/Faro/COMPUTERS/COMPUTERS-24.pdf
83. Syberfeldt A, Lidberg S. (2012). Real-world simulation-based manufacturing optimization using cuckoo search. In *Proceedings of the 2012 Winter Simulation Conference (WSC)* (pp. 1–12). IEEE. https://doi.org/10.1109/WSC.2012.6465225
84. Tan WS, Hassan MY, Majid MS, Rahman HA. (2012). Allocation and sizing of DG using cuckoo search algorithm. In *2012 IEEE International Conference on Power and Energy (PECon)* (pp. 133–138). IEEE. https://doi.org/10.1109/PECon.2012.6450193
85. Tiwari V. Face recognition based on cuckoo search algorithm. Int J Comput Appl. 2012;48(21):9–14. https://doi.org/10.5120/7464-0551.
86. Tuba M, Subotic M, Stanarevic N. (2011). Modified cuckoo search algorithm for unconstrained optimization problems. In *Proceedings of the 5th European Conference on European Computing Conference (ECC'11)* (pp. 263–268). https://dl.acm.org/doi/10.5555/2001550.2001597
87. Valian E, Tavakoli S, Mohanna S, Haghi A. Improved cuckoo search for reliability optimization problems. Comput Ind Eng. 2013;64(1):459–68. https://doi.org/10.1016/j.cie.2012.11.002.
88. Valian E, Valian E. A cuckoo search algorithm by Lévy flights for solving reliability redundancy allocation problems. Eng Optim. 2013;45(11):1273–86. https://doi.org/10.1080/0305215X.2012.735780.
89. Vazquez RA. (2011). Training spiking neural models using cuckoo search algorithm. In *2011 IEEE Congress on Evolutionary Computation (CEC)* (pp. 679–686). IEEE. https://doi.org/10.1109/CEC.2011.5949686
90. Van Veldhuizen DA, Lamont GB. Multiobjective evolutionary algorithms: Analyzing the state-of-the-art. Evolution Comput. 2000;8(2):125–47. https://doi.org/10.1162/106365600568158.
91. Vo DN, Schegner P, Ongsakul W. Cuckoo search algorithm for non-convex economic dispatch. IET Generation Transmission Distribution. 2013;7(6):645–54. https://doi.org/10.1049/iet-gtd.2012.0142.
92. Walia GS, Kapoor R. (2013). Particle filter based on cuckoo search for non-linear state estimation. In *2013 3rd IEEE International Advance Computing Conference (IACC)* (pp. 918–924). IEEE. https://doi.org/10.1109/IAdCC.2013.6514349

93. Walton S, Hassan O, Morgan K. Reduced order mesh optimisation using proper orthogonal decomposition and a modified cuckoo search. Int J Numer Methods Eng. 2013;93(5):527–50. https://doi.org/10.1002/nme.4398.
94. Walton S, Hassan O, Morgan K, Brown MR. Modified cuckoo search: a new gradient free optimisation algorithm. Chaos Solitons Fractals. 2011;44(9):710–8. https://doi.org/10.1016/j.chaos.2011.06.004.
95. Walton S, Hassan O, Morgan K. Selected engineering applications of gradient free optimisation using cuckoo search and proper orthogonal decomposition. Arch Comput Methods Eng. 2013;20(2):123–54. https://doi.org/10.1007/s11831-013-9083-4.
96. Wang F, Luo L, He XS, Wang Y. (2011). Hybrid optimization algorithm of PSO and Cuckoo Search. In *2011 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC)* (pp. 1172–1175). IEEE. https://doi.org/10.1109/AIMSEC.2011.6010043
97. Wang F, He XS, Wang Y. The cuckoo search algorithm based on Gaussian disturbance. J Xi'an Polytechnic Univ. 2011;25(4):554–7. http://xbyb.xpu.edu.cn/EN/Y2011/V25/I04/554.
98. Wang F, He XS, Wang Y, Yang SM. Markov model and convergence analysis based on cuckoo search algorithm. Comput Eng. 2012;38(11):180–2. http://www.ecice06.com/EN/Y2012/V38/I11/180.
99. Wang G, Guo L, Duan H, Wang H, Liu L, Shao M. (2012). A hybrid metaheuristic DE/CS algorithm for UCAV three-dimension path planning. *The Scientific World Journal*, 2012, 583160. https://doi.org/10.1100/2012/583160
100. Wang Q, Liu S, Wang H, Savić DA. (2012). Multi-objective cuckoo search for the optimal design of water distribution systems. In *Civil Engineering and Urban Planning 2012* (pp. 402–405). https://doi.org/10.1061/9780784412435.077
101. Wolpert DH, Macready WG. No free lunch theorems for optimization. IEEE Trans Evol Comput. 1997;1(1):67–82. https://doi.org/10.1109/4235.585893.
102. Yang XS. Cuckoo search for inverse problems and simulated-driven shape optimization. J Comput Methods Sci Eng. 2012;12(1–2):129–37. https://doi.org/10.3233/JCM-2012-0409.
103. Yang XS. Bat algorithm and cuckoo search: a tutorial. Artificial Intelligence, Evolutionary Computing and Metaheuristics. Berlin, Heidelberg: Springer; 2013. pp. 421–34. https://doi.org/10.1007/978-3-642-29694-9_18.
104. Yang XS. Metaheuristic algorithms for inverse problems. Int J Innovative Comput Appl. 2013;5(2):76–84. https://doi.org/10.1504/IJICA.2013.055811.
105. Yang XS, Deb S. Engineering optimisation by cuckoo search. Int J Math Modelling Numer Optimisation. 2010;1(4):330–43. https://doi.org/10.1504/IJMMNO.2010.035430.
106. Yang XS, Deb S. (2012). Cuckoo search for inverse problems and topology optimization. In *Proceedings of International Conference on Advances in Computing* (pp. 291–295). Springer, India. https://doi.org/10.1007/978-81-322-0740-5_34
107. Yang XS, Deb S. Multiobjective cuckoo search for design optimization. Comput Oper Res. 2013;40(6):1616–24. https://doi.org/10.1016/j.cor.2011.09.026.
108. Yang XS, Deb S, Karamanoglu M, He X. (2012). Cuckoo search for business optimization applications. In *2012 National Conference on Computing and Communication Systems* (pp. 1–5). IEEE. https://doi.org/10.1109/NCCCS.2012.6413000
109. Yildiz AR. Cuckoo search algorithm for the selection of optimal machining parameters in milling operations. Int J Adv Manuf Technol. 2013;64(1–4):55–61. https://doi.org/10.1007/s00170-012-4013-7.
110. Zhang Y, Wang L, Wu Q. Modified adaptive cuckoo search (MACS) algorithm and formal description for global optimisation. Int J Comput Appl Technol. 2012;44(2):73–9. https://doi.org/10.1504/IJCAT.2012.048386.
111. Zheng H, Zhou Y. A novel cuckoo search optimization algorithm based on Gauss distribution. J Comput Inform Syst. 2012;8(10):4193–200. http://www.jocis.net/.
112. Zhou Y, Zheng H. (2013). A novel complex valued cuckoo search algorithm. *The Scientific World Journal*, 2013, 597490. https://doi.org/10.1155/2013/597490
113. Jana RK, Maity S. An accuracy based comparative study on different techniques and challenges for sentiment analysis. Pervasive Computing and Social Networking. Singapore: Springer; 2022. pp. 601–19. https://doi.org/10.1007/978-981-19-2840-6_46.
114. Jana RK, Maity S, Maiti S. (2022). An empirical study of sentiment and behavioural analysis using homophily effect in social network. In *2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS)* (pp. 1508–1515). IEEE. https://doi.org/10.1109/ICICCS53718.2022.9788339
115. Jana RK, Singh D, Maity S, Paul H. Perception analysis of social network tweets using a genetic algorithm based approach. J Softw Eng Tools Technol Trends. 2023;10(2):1–10. http://techjournals.stmjournals.in/index.php/JoSETTT/article/view/1000.
116. Pandey AC, Rajpoot DS, Saraswat M. Twitter sentiment analysis using hybrid cuckoo search method. Inf Process Manag. 2017;53(4):764–79. https://doi.org/10.1016/j.ipm.2017.02.004.
117. Kakkar P, Pushe V. A product review system using sentiment analysis with cuckoo search and deep learning technique. J Emerg Technol Innovative Res (JETIR). 2021;8(7):d191–202. https://www.jetir.org/papers/JETIR2107413.pdf.
118. Priyanka, Walia K. Cuckoo search optimized improved opinion mining and classification. Int J Eng Trends Technol. 2022;70(10):44–53. https://doi.org/10.14445/22315381/IJETT-V70I10P205.
119. Ahmed SH, Wassif KT, Nabil E. Clustering based sentiment analysis using randomized clustering cuckoo search algorithm. Int J Comput Sci Netw Secur. 2020;20(7):159–67. http://paper.ijcsns.org/07_book/202007/20200720.pdf.
120. Sudhamathy G, Valliammal N. (2023). Twitter sentiment analysis using clustering based cuckoo search algorithm and ensemble classifier. In *2023 International Conference on Intelligent Systems for Communication, IoT and Security (ICISCoIS)* (pp. 212–217). IEEE. https://doi.org/10.1109/ICISCoIS56541.2023.10100378

121. Gherboudj A, Layeb A, Chikhi S. A discrete cuckoo search algorithm for 0–1 multidimensional knapsack problem. Int J Appl Metaheuristic Comput. 2013;4(4):1–19. https://doi.org/10.4018/ijamc.2013100101.
122. Wang GG, Deb S, Gandomi AH, Zhang Z, Alavi AH. Multi-objective cuckoo search for crisp and fuzzy engineering optimization. Soft Comput. 2016;20(2):519–40. https://doi.org/10.1007/s00500-014-1522-8.
123. Mareli M, Engelbrecht AP. Enhanced cuckoo search algorithms for solving optimization problems. Appl Soft Comput. 2018;63:203–16. https://doi.org/10.1016/j.asoc.2017.11.025.
124. Fister I, Yang XS, Fister D, Fister I. Cuckoo search: a brief literature review. Cuckoo Search and Firefly Algorithm. Cham: Springer; 2014. pp. 49–62. https://doi.org/10.1007/978-3-319-02141-6_3.
125. Chi R, Su Y, Dong R, Liu X, Sun H. Hybrid metaheuristic DE-CS algorithm for multi-objective optimization. Appl Math Model. 2018;62:597–611. https://doi.org/10.1016/j.apm.2018.06.027.
126. Shehab M, Khader AT, Al-Betar MA. A comprehensive review of the cuckoo search algorithm: concepts and applications. Arch Comput Methods Eng. 2017;24(1):203–18. https://doi.org/10.1007/s11831-015-9161-x.
127. Abdel-Basset M, Ding W, El-Shahat D. A hybrid cuckoo search algorithm for numerical and engineering optimization problems. Int J Mach Learn Cybernet. 2020;11(7):1461–84. https://doi.org/10.1007/s13042-019-01050-x.
128. Naik A, Satapathy SC. A review on cuckoo search algorithm and its applications. J King Saud Univ - Comput Inform Sci. 2021;33(7):753–67. https://doi.org/10.1016/j.jksuci.2018.05.001.
129. Jana RK, Singh D, Maity S, Paul H. A hybrid approach to analyse the public sentiment on COVID-19 tweets. Indian J Sci Technol. 2024;17(7):610–6. https://doi.org/10.17485/IJST/v17i7.3228.
130. Jana RK, Singh D, Maity S. Modified firefly algorithm and different approaches for sentiment analysis. Sci Temper. 2024;15(01):1745–54. https://doi.org/10.58414/SCIENTIFICTEMPER.2024.15.1.41.
131. Jana RK, Singh D, Maity S, Paul H, Mallick A, Ghatak S, Mallik S, Wang M. Public perception analysis on COVID-19 tweets using hybridised method. Int J Comput Biol Drug Des. 2024;16(1):19–41. https://doi.org/10.1504/IJCBDD.2024.137887.
132. Das S, Nag S, Jana RK. A multimodal systematic review of stress detection in university students based on machine learning and physiological measures. J Cogn Syst. 2025;10(1):15–28. https://doi.org/10.52876/jcs.134598.
133. Roy S, Jana RK, Bhabani B, Das T. From tweets to trades: Applications of financial sentiment analysis in market forecasting. J Comput Sci Eng Innovations Res. 2025;2(1):1–6. https://jcseir.org/index.php/jcseir/article/view/18.
134. Molla SA, Sarkar S, Jana RK, Bhabani B, Das T. The sentiment behind the tweet. J Comput Sci Eng Innovations Res. 2025;2(1):7–12. https://jcseir.org/index.php/jcseir/article/view/19.
135. Jana RK, Kotal A, Nayek AK. Healthcare made smarter: AI-driven drug recommendation systems. J Comput Sci Eng Innovations Res. 2025;2(1):25–30. https://jcseir.org/index.php/jcseir/article/view/22.
136. Pan D, Jana RK, Nag T, Mitra A. Analyzing sentiments to detect drug-related tweets: A software engineering framework. J Comput Sci Eng Innovations Res. 2025;2(1):31–6. https://jcseir.org/index.php/jcseir/article/view/23.
137. Mandal S, Pal S, Jana RK, Nayek AK. Uncovering guest sentiment: Analyzing hotel reviews. J Comput Sci Eng Innovations Res. 2025;2(1):37–42. https://jcseir.org/index.php/jcseir/article/view/24.
138. Ghosh N, Jana RK. (2025). Tweeting health: Multimodal sentiment analysis for healthcare insights. In *Proceedings of the 4th International Conference on Innovations in Data Analytics (ICIDA 2025)*, Eminent College of Management and Technology, West Bengal, India. https://icida.in/2025/papers/ICIDA2025_042.pdf
139. Ghosh N, Jana RK. (2025). Emotion-driven medication recommendations for patients. In *Proceedings of the 4th International Conference on Innovations in Data Analytics (ICIDA 2025)*, Eminent College of Management and Technology, West Bengal, India. https://icida.in/2025/papers/ICIDA2025_043.pdf
140. Mittal S, Umamaheswaran S. CoWIN twitter dataset: A comprehensive collection of public discourse on India's COVID-19 vaccination platform. Data Brief. 2025;58:111252. https://doi.org/10.1016/j.dib.2024.111252.
141. Kumar A, Jaiswal A, Garg S, Verma S, Kumar S. Sentiment analysis using Cuckoo Search for optimized feature selection on Kaggle tweets. Int J Inform Retr Res. 2019;9(1):1–15. https://doi.org/10.4018/IJIRR.2019010101.

Figures

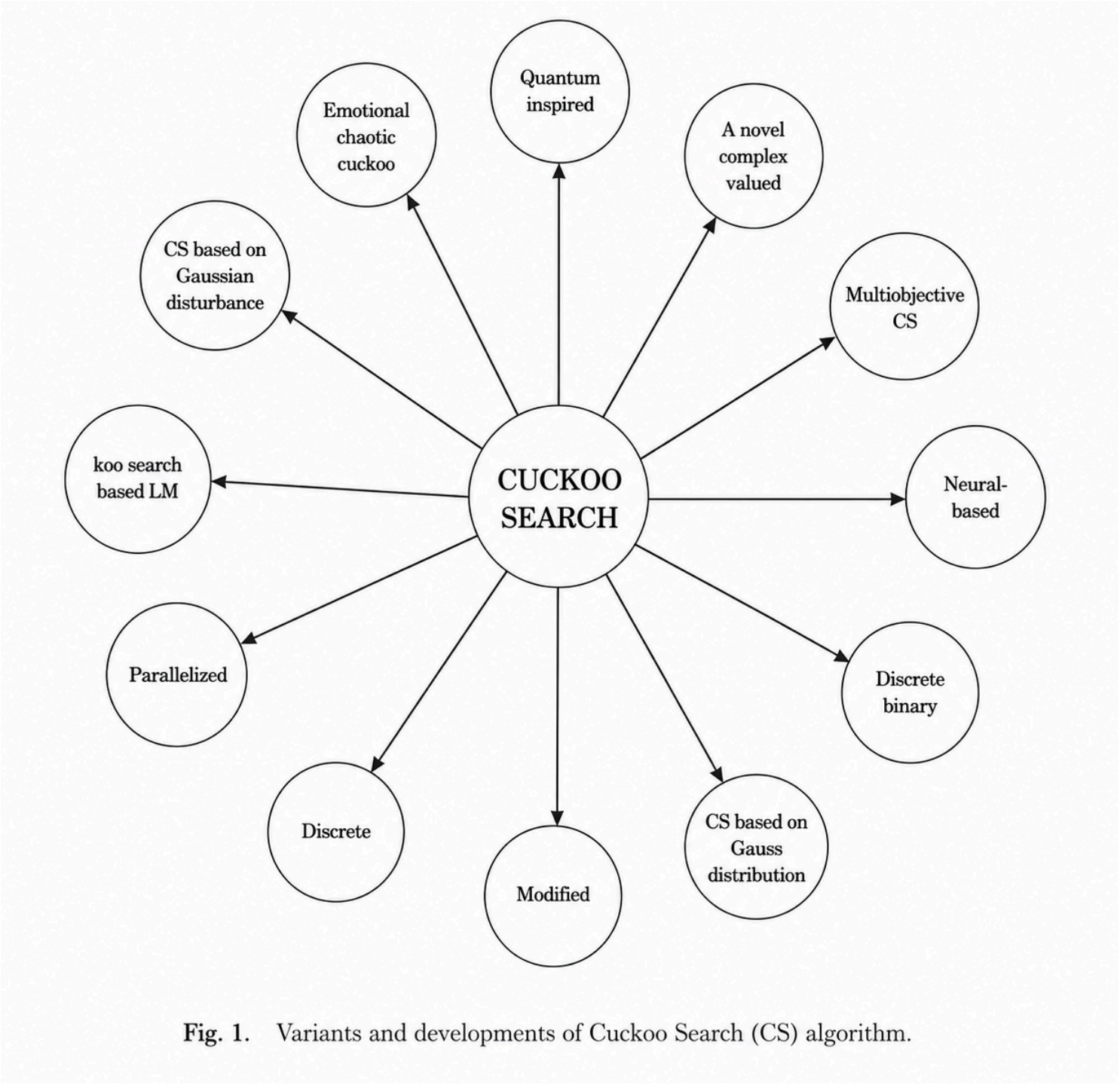


Fig. 1. Variants and developments of Cuckoo Search (CS) algorithm.

Figure 1

See image above for figure legend

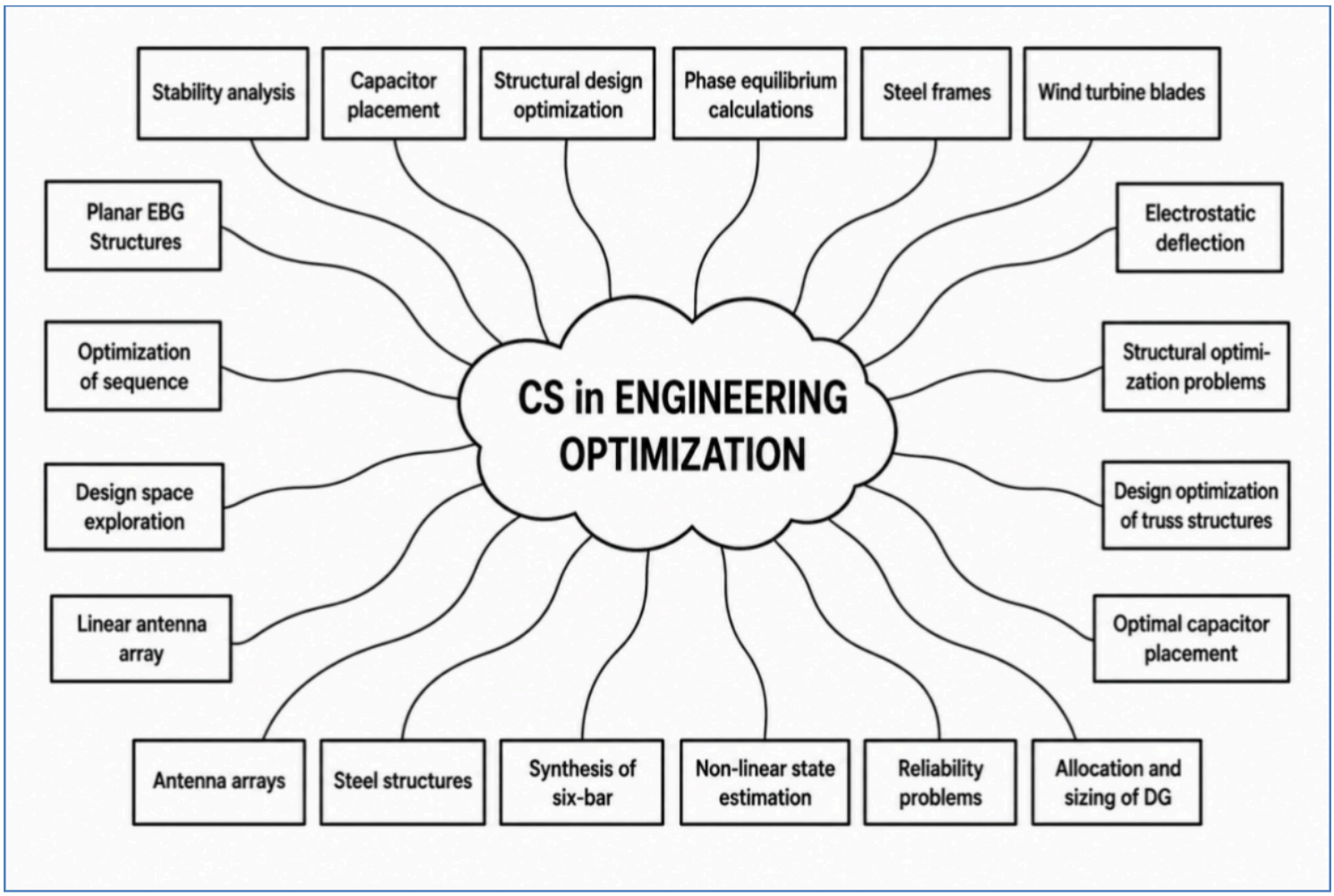


Figure 2

CS in Engineering Optimization.

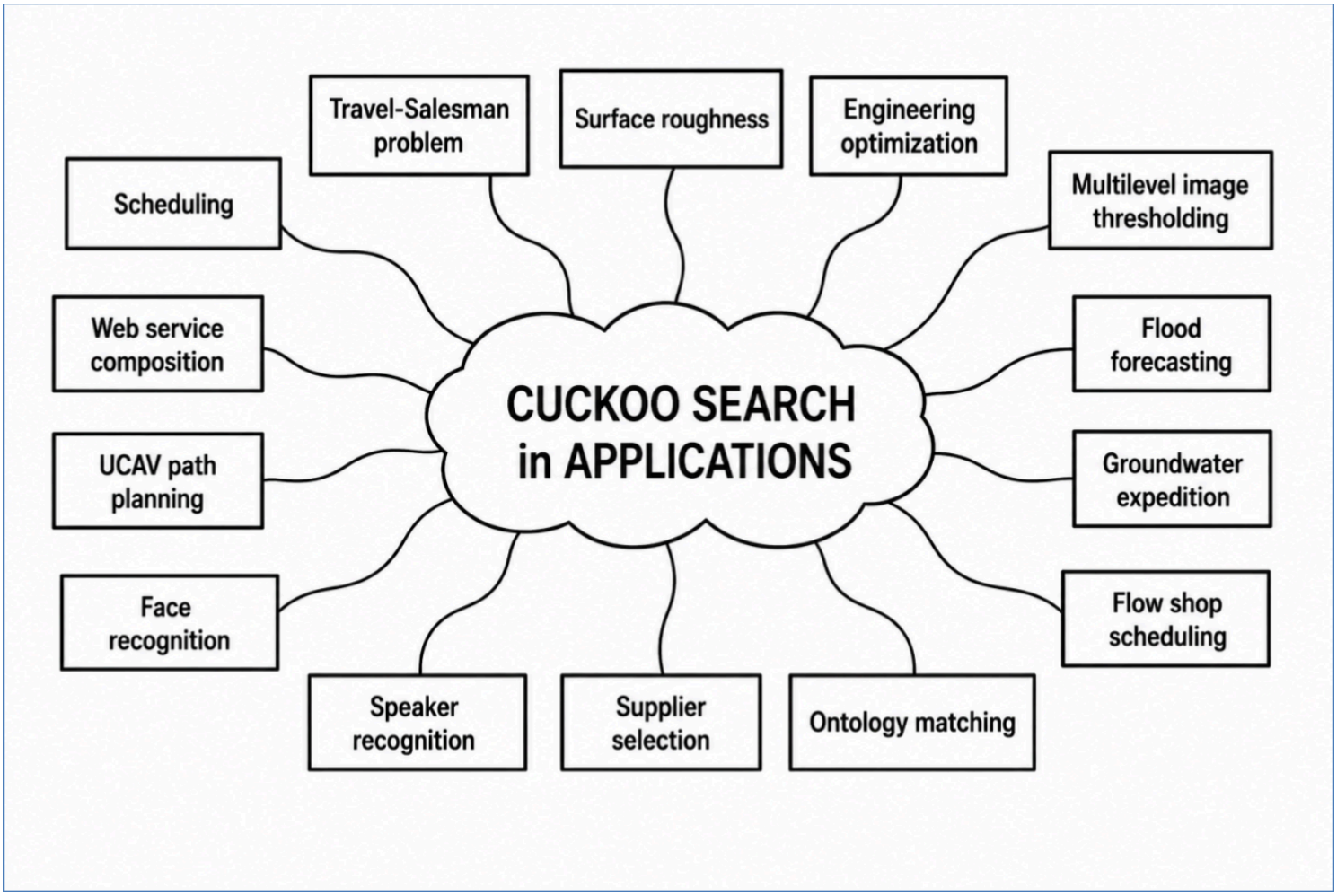


Figure 3

Showed the application of cuckoo search on different area.

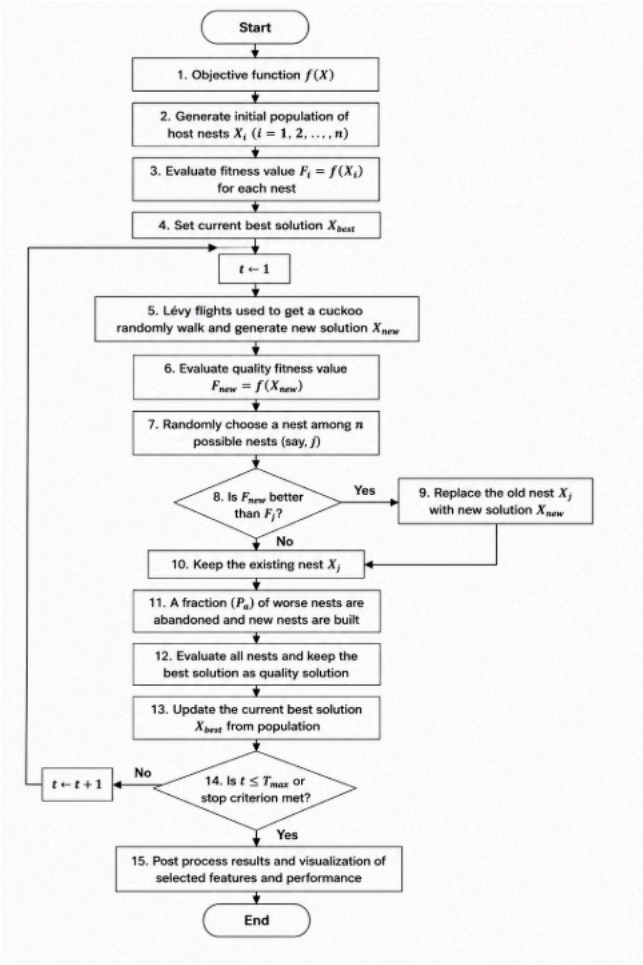


Figure 4

Standard Cuckoo Search Behaviors

Input:

TF-IDF feature matrix
Population size P
Maximum number of iterations T

Output:

Best selected feature subset

1: **Encode** each candidate feature subset as a binary vector.
2: **Generate** the initial population using Differential Evolution.
3: **Evaluate** the fitness of the DE-generated population.
4: **Transfer** the DE-generated population to Cuckoo Search.
5: **Identify** the current best candidate solution.
6: **Generate** new candidate solutions using the Cuckoo Search mechanism.
7: **Evaluate** the fitness of the new candidate solutions.
8: **Replace** inferior solutions according to the CS replacement strategy.
9: **Apply** the CS abandonment/replacement operation where required.
10: **Update** the best solution.
11: **Repeat** Steps 6–10 until the maximum number of iterations is reached.
12: **Return** the best feature subset.
13: **Evaluate** the selected features using the SVM classifier.

Figure 5

Proposed Algorithm of Hybrid Cuckoo Search

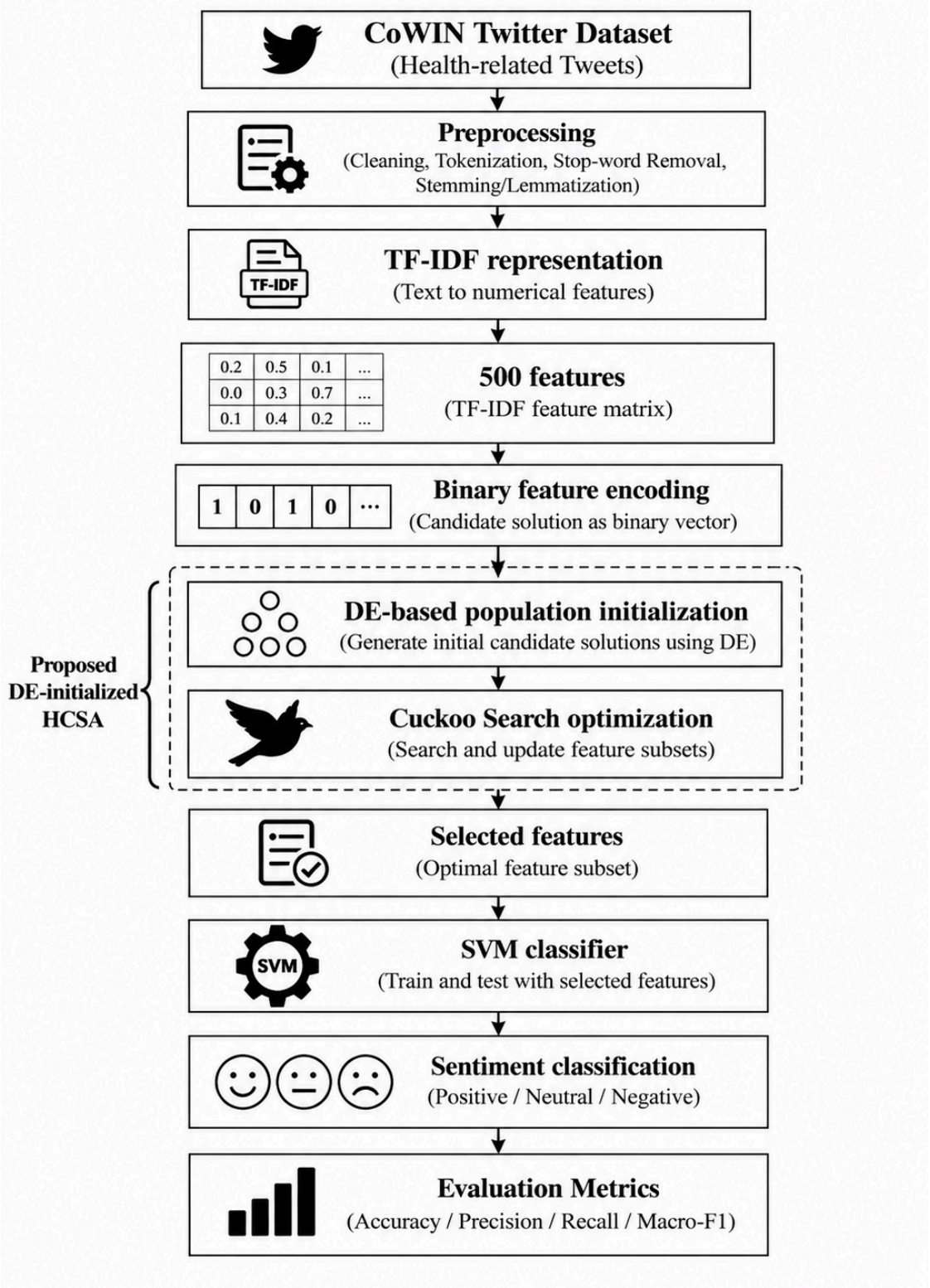


Figure 6

Overall Proposed DE-initialized Cuckoo Search framework

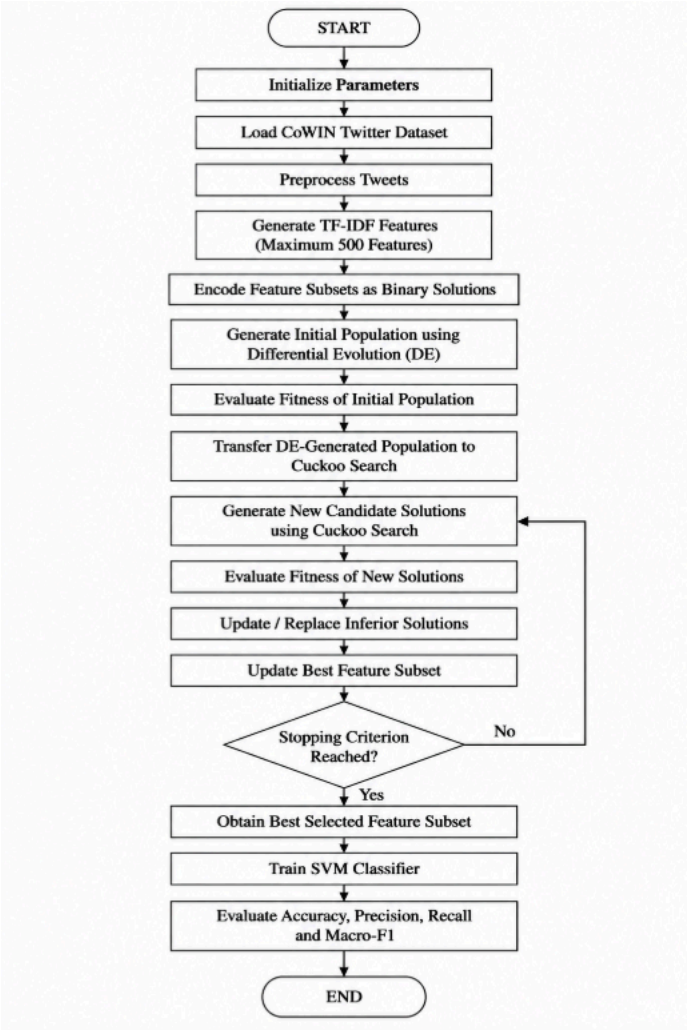


Figure 7

Flowchart of Proposed Algorithm

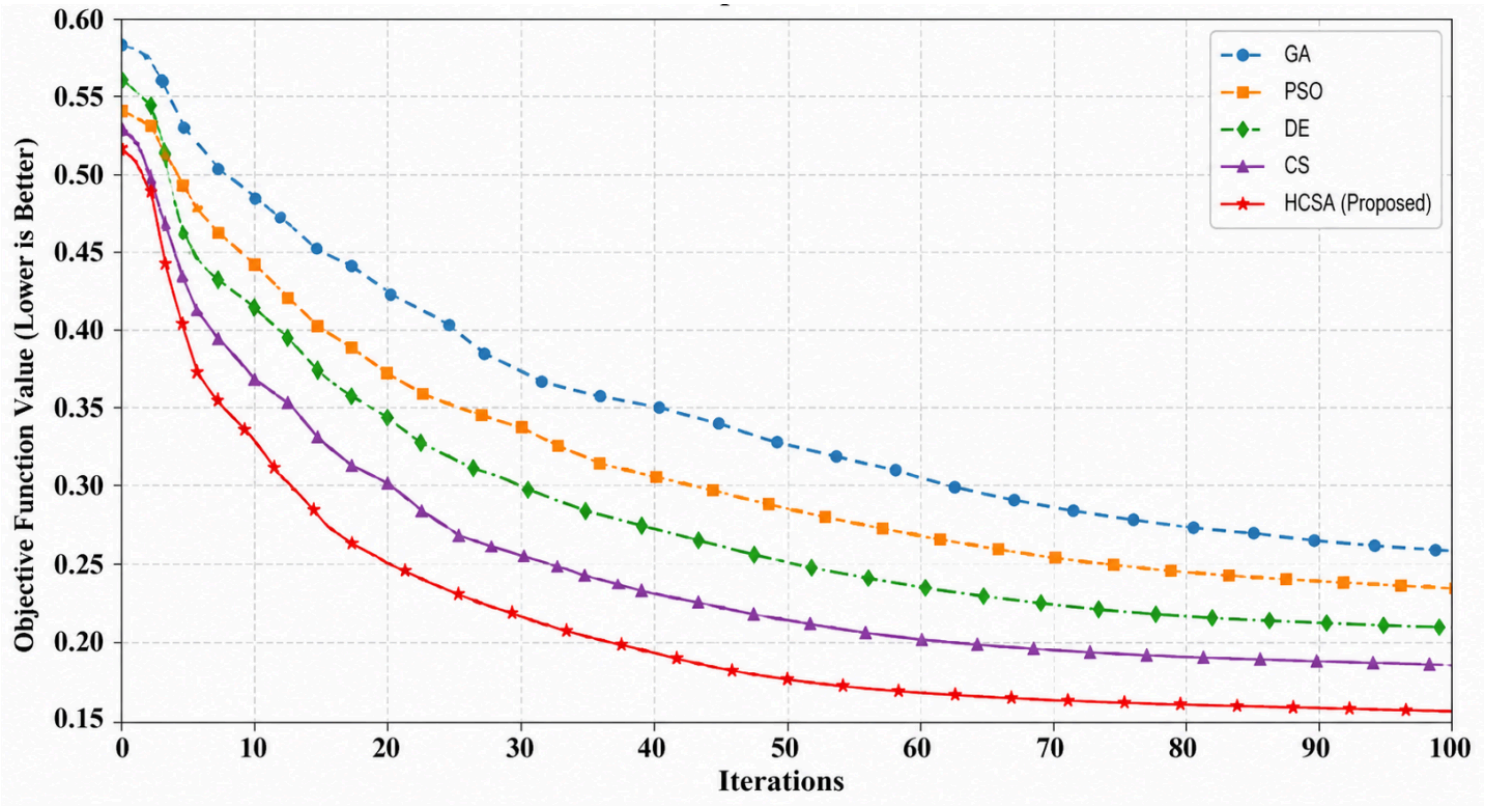


Figure 8

Convergence behaviors of proposed algorithm with other on Successive run