Feature Selection Using Memetic Salp Swarm Algorithm with Symmetrical Uncertainty Ranking for Arabic Text Classification

Mohammed Ghassan Abdulkareem [1], Alhasan Amer Ibrahim [1], Ibrahem Amer Hammed [2], and G. Abdulkareem-Alsultan [3,*]

[1] Department of Management and Marketing of Oil and Gas, College of Industrial Management of Oil and Gas, Basrah University for Oil and Gas Basrah, Basrah, Iraq
[2] Ministry of Higher Education and Scientific Research, Bagdad, Iraq
[3] Catalysis Science and Technology Research Centre (PutraCat), Faculty of Science, Universiti Putra Malaysia, Serdang, Malaysia
Email: mohammed.alsultan@buog.edu.iq (M.G.A.); Alhassan.amer@buog.edu.iq (A.A.I.); Ibrahemamer032@gmail.com (I.A.H.); kreem.alsultan@yahoo.com (G.A.A.)
*Corresponding author

***Abstract*—Dealing with high-dimensional data has made it very challenging to find the best subset of chosen features because of the limitation in reducing the exponential growth of the search process. Text feature selection is a dimensionality reduction technique that tends to reduce the extra text features for better classification accuracy. Feature selection is an Nondeterministic Polynomial-time hard (NP-Hard) optimization. Additionally, many feature selection models overlook the interactions between or between features and the decision class. To produce the suggested Memetic Salp Swarm Algorithm (MSSA), the SSA is combined with Symmetrical Uncertainty (SU). The proposed technique intends to pinpoint the most valuable characteristics of Arabic text by decreasing computing complexity and enhancing classification accuracy. MSSA, drawing inspiration from the innate behaviour of salps and the concepts of memetic algorithms, effectively navigates the search space to locate the most optimum subsets of features. Symmetrical Uncertainty ranking enhances the selection process by precisely measuring the significance of the characteristics of the classification objective. The suggested strategy has been proven successful in Arabic text datasets through experimental findings, surpassing existing classification accuracy and feature subset size approaches.**

***Keywords*—optimization, feature selection, memetic, classification, computing complexity**

I. Introduction

Modern datasets include vast amounts of data, necessitating the creation of sophisticated algorithms to extract valuable insights. Data extraction/selection models are constructed based on specific data mining objectives, typically in regression, grouping, and classification. Preprocessing datasets generally achieve more efficient data analyses by reducing dataset size and adapting the dataset to suit the specified analysis technique best. Currently, the primary focus is on utilizing advanced analytic tools available to researchers since the size of an average dataset is increasing in terms of samples and characteristics [1]. Unsupervised feature selection enables constructing more efficient and accurate analytical models without requiring large amounts of pre-labelled data. To achieve this goal, techniques such as frequency analysis, word relationship analysis, segmentation and clustering methods are used to select the most critical and meaningful text features. Because of the increase of electronic text documents, knowledge discovery from text become one of the most challenging tasks. More than 80% of the online information is primarily from the textual type. Reducing the extra text dimensions is considered a higher-priority task [2].

Feature selection is one of the most essential cleansing methodologies concerning the text. Reducing noisy text data is necessary as this increases the accuracy of the machine learning results and will reduce the processing time of unnecessary information. Different methods can be used for the feature selection of text; the easiest is the one that uses a dictionary of the vocabulary that must be removed from the text, which is usually known as the stop words. However, while specific stop word lists such as the Natural Language Toolkit (NLTK) default English stop words list (containing 179 predefined terms) exist with defined sizes, there is no universally standardized or consistently applied list across studies, leading to variability that complicates reproducibility and comparison of results in text processing tasks. Besides, this method does not consider the statistical relationship between terms within the text. Moreover, stop words vary from text to text, a topic-dependent issue. Therefore, machine learning techniques can be helpful to methods in this regard. There are three kinds of methods: filter, wrapper, and hybrid. Filter methods are only one pass

Manuscript received December 13, 2024; revised February 4, 2025; accepted June 3, 2025; published September 23, 2025.

doi: 10.12720/jait.16.9.1352-1363

filter that can pick up the best features based on a specific statistical criterion. Feature selection filter techniques function by assessing features according to their performance metrics without regard to the data modelling methodology used. The modeling algorithms can then use the optimal features after determining them. These techniques can evaluate whole feature subsets or rank individual features. There are several categories into which the feature filtering measures which have been established may be roughly classified, such as information, distance, consistency, similarity, and statistical measurements [3].

Several filter approaches are discussed in the literature) including gain Ratio, Correlation, Information Gain (IG), and Symmetrical Uncertainty (SU) [3]. Chi-square [4] and Relief and Relief-F [5, 6] are popular strategies and pertinent references for more information. It's crucial to remember that not every kind of data mining operation can profit from every filter option. Therefore, filters are also classified according to the type of job they are used for, such as clustering, regression, or classification [2]. The other category is wrapper-based methods, which are iterative. Meta-heuristic methods are the ideal choice for implementing wrapper-based methods. Such methods work on optimizing a specific objective function in the feature selection regarding the accuracy of categorizing the characteristics associated with the selected subset, which is the objective function that has been chosen. It has been established that metaheuristic-based algorithms are useful across various areas by offering realistic and achievable solutions within reasonable periods. To solve high-dimensional issues, they ensure that classification performance is optimized while minimizing the compute resources, storage needs, and feature count. Metaheuristic algorithms encompass a wide variety of techniques, such as ant colony optimization [7], genetic algorithms [8], memetic algorithm [9], particle swarm optimization [10], evolutionary-based algorithm [11], grey wolf optimizer [12], firefly [13], binary Jaya [14], dragonfly algorithm [15, 16], and others. Some examples of these algorithms are listed below.

As a third category of feature selection, hybrid methods are used to combine two feature selection types together to obtain the combined advantage of any two successful methods [17]. Thus, the combination of the Wrappers and Filters is considered successful in terms of obtaining better accuracies, i.e., selecting the most informative features subsets when compared with using each method alone. The Memetic algorithms are used to combine the wrappers as global search method with the filters as local searchers [18].

The main contribution of this study is to utilize the capabilities of Memetic Algorithms (MAs) inside the correlation-based memetic framework to efficiently and effectively address challenging optimization problems. We intend to showcase the robustness and flexibility of our proposed technique through the implementation of empirical assessments and experiments. The Salp Swarm Algorithm (SSA) has been included in the Modified Salp Swarm Algorithm (MSSA) to improve its feature selection efficacy in optimization tasks. The SSA is important in this integration, emulating the swarming behavior of salps to navigate the search space effectively. The SSA directs the search process by adjusting the placements of the salps according to the leader-follower mechanism, facilitating a balanced investigation and exploitation of features. Exploitation is an essential feature of optimization. This reduces the selection of the most pertinent characteristics from the dataset, decreases dimensionality, and enhances model performance. The grantee of the fast convergence can be achieved using the localized search. The integration guarantees that the algorithm circumvents local optima and converges on the optimal feature subset, hence improving the overall efficacy of the feature selection process. We aim to showcase the versatility of our method in effectively addressing diverse real-world scenarios. The suggested method is named Correlation based Memetic Salp Swarm Optimization Feature Selection (C-MSSO-FS). The model used is a wrapper-filter model the wrapper is represented by the Salp Swarm Optimization (SSO) optimization that conduct the global search of features while the local search is represented by the local ranking method which is an uncertainty correlation-based method.

In localized search with symmetrical uncertainty, the algorithm evaluates the relevance of each feature by calculating Symmetrical Uncertainty (SU), a normalized metric derived from mutual information. SU quantifies the degree of association between a feature and the target variable while compensating for the biases of entropy, ensuring a fair comparison across features of different cardinalities. Higher SU values signify that a feature provides more predictive information about the target, making it a strong candidate for inclusion in the optimal feature subset.

During the search process, SU values are used to guide the exploration of the feature space. Instead of performing a random or exhaustive search, the algorithm focuses on localized regions—neighborhoods in the feature space where high-SU features cluster together. These regions are considered more promising, as they are likely to contain feature combinations with higher discriminative power. By leveraging the SU values, the algorithm prioritizes evaluating feature subsets that include these high-relevance attributes, thus reducing computational complexity and enhancing efficiency. To refine the search, the algorithm may use clustering techniques to identify dense regions in the SU-ranked feature space. These clusters, or centroids, represent feature subsets that are not only individually strong but may also exhibit low redundancy and high complementarity. By iteratively exploring variations around these centroids, the algorithm aims to converge on optimal or near-optimal subsets that offer superior classification or prediction performance.

Moreover, this SU-guided localized search framework can be integrated with metaheuristic algorithms—such as Genetic Algorithms, Particle Swarm Optimization, or Ant Colony Optimization—to further enhance exploration and exploitation capabilities. These metaheuristics can use SU values as heuristic information or fitness components, effectively balancing relevance and redundancy in feature selection. The main contributions of this paper are listed below:

1. Suggesting a method that is capable of analyzing and clustering Arabic text using the optimization techniques.
2. Use three algorithmic combinations to make the best balance between exploration and exploitation in the search space for optimal solutions. The suggested algorithms are sequential global-first hybridization, sequential global-last hybridization, and interleaved hybridization sequentially.
3. Use the symmetrical uncertainty as it can effectively quantify the relevance and redundancy of features about a target variable.
4. Use the Salp Swarm Algorithm (SSA) due to its ability to explore large solution spaces effectively.

This paper is arranged as follows: in Section II, the related work is explained, while in Section III, the preliminaries are discussed in detail, including the salp optimization, the symmetrical uncertainty and the memetic algorithms. In Section IV the proposed method is explained. The dataset is explained in Section V, and in Sections VI and VII, the test results and the conclusions are explained, respectively.

II. Related Works

This section contains the related work to the proposed work in this paper. An analytical comparison has been done to show their strengths and weaknesses. The reviewed methods are wrapper and iterative methods that deploy the metaheuristic methods to resolve the feature selection problem. Most of those methods were based on the idea of the binary search using the metaheuristic methods.

The wrapper-based feature selection technique uses the learning algorithm to evaluate the utility of features. The wrapper approach facilitates an interaction between the classification process.

An algorithm for searching subsets. The programme incorporates a statistical resampling method subroutine, such as cross-validation. This methodology uses the target learning algorithm to measure the accuracy of different subsets of features. The wrapper strategy has proven superior in classification tasks because it can optimize classifier performance while handling real-world problems. However, the duration is extended throughout execution as the learning process must be invoked frequently.

Compared to the filter approach, they are more computationally demanding because of the recurrent learning phases and cross-validation. Swarm intelligence methods, as well as evolutionary computing-based methods, have had extended success in the domain of feature selection and dimensionality reduction [19]. The biology metaphorized method, the Particle Swarm Optimization (PSO) method, has been widely used with complex optimization problems such as the NP-hard optimization methods, including feature extraction and feature selection. Consequently, the power of the PSO is its capability to handle issues in discrete domains, which can be called discrete searching space. Thus, Li *et al.* [20] used a PSO model based on the logistic regression modelling in the settings of the dimensionality reduction.

Later on, Berlin and John [21] combined the catfish impact with the binary PSO to balance exploration and exploitation. Thus, Samal and Panigrahy [22] suggested the catfish-PSO to improve the standard PSO model. As another notable modification, Chegini *et al.* [23] enhanced the canonical PSO into an improved version of the Levy flight method, which has been used as a local search method. The global search coefficient represented by the weighting inertia factor and the mutations factor that controls the diversification as an enhancement factor has also been updated in this modified version of the PSO. Nonetheless, the improvement suggested incurred several disadvantages, such as adding more tuning parameter values compared to the older modified version. Besides, the computational time has notably been prolonged. Another drawback is that the parameter modulation becomes more complicated for the various applications.

The significant negativity of the PSO in this domain is its inability to progress as the particles were stacked in the hyper-cubic search space corner.

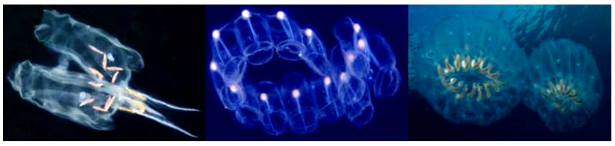
The proposed approach by Jain and Dharavath [24] incorporates a Memetic Salp Swarm Optimization Algorithm (MSSOA), which has been converted into a binary MSSOA. The purpose of this modification is to ascertain the optimal number of attributes that would enable the highest possible level of classification accuracy. The proposed method achieves better results than five metaheuristic feature selection algorithms Binary Salp Swarm Algorithm (BSSA), Binary Particle Swarm Optimization (BPSO), Binary Moth-Flame Optimization (BMFO), Binary Cuckoo Optimization Algorithm (BCOA), and Improved Binary Harris Hawks Optimization (IBHHO) when evaluated on benchmark datasets from the University of California, Irvine, achieving improved classification accuracy and reduced feature size. Applied to maize, rice, and grape plants, the proposed algorithm achieves classification accuracies of 90.6%, 67.9%, and 91.6%, respectively, with peak accuracies reaching 93.6%, 79.1%, and 95%.

The Genetic Algorithm (GA) is another biological-based method widely used as an iterative method in feature selection, which is based on repeating the iteration and wrapping around the solution to the end of the optimization process. Tao *et al.* [25] suggested a GA-based method used to handle dimensionality reduction (SVM), which has been used with the genetic algorithm to resolve this problem. The main focus of that research is the features subsets and SVM parameters simultaneously while keeping the accuracy intact. It was noticed that the classification accuracy had been remarkably enhanced via the decrement of the feature's subset number. However, there is still a lag hidden in the Girds technique. A combination of the Genetic algorithm and the Ant Colony Optimization method has been suggested by Liang and Wang [26] for an efficient feature selection. The recommended method is used for Protein prediction. Moreover, Hosseini and Zade [27] suggested modifying the standard Genetic Algorithm (GA); their proposed method is Modified Genetic Algorithm (MGA). This method used a transfer learning technique that deployed the Deep Neural Nets (DNN) model. This model was helpful in the prediction of demand for many basic resources in outpatient settings. Consequently, combining both algorithms had some promising results compared to each single algorithm used solely. The expected overhead was not considered significant when it compromised the accuracy obtained. In this regard, the processing power was not high for this combination.

Ultimately, numerous approaches exhibit a deficiency in flexibility, as they are frequently customized for particular objectives and encounter difficulties adapting to diverse domains or data types. The suggested method seeks to tackle these issues by augmenting accuracy, strengthening generalization skills, minimizing computing expenses, and providing increased flexibility for various applications.

III. Salp Optimization Method

Salps are part of the Salpidae family and have a translucent barrel-shaped body. Their tissues closely resemble those of jellyfish. They move similarly to jellyfish, using water propulsion to move forward. Fig. 1(a) displays the morphology of a salp. Research on this organism is in its early stages because of the challenges of accessing and maintaining their natural habitats in laboratory settings. One of the salps' most intriguing behaviours, highlighted in the research, is their swarming behaviour. Salps frequently aggregate in deep ocean environments to create a collective structure called a salp chain. Fig. 1(b) depicts this chain. The primary cause of this behaviour remains unclear. However, some experts suggest it is done to improve mobility through quick, coordinated adjustments and foraging.



(a) (b) (c)

Fig. 1. Illustrates the form and composition of salp swarms found in the deep ocean. It includes (a) an individual salp, (b) a single chain of salps, and (c) double chains of salps.

The mathematical modeling of salp chains involves initially categorizing the population into leaders and followers. The Salp at the head of the chain is the leader, while the other salps are followers. The leader of the salps commands the swarm, while the following follow either directly or indirectly Fig. 1(c). Like other swarm-based methods, salps' positions are determined in an n-dimensional search space, where n represents the number of variables in a specific issue. Thus, the coordinates of all salps are recorded in a two-dimensional matrix named *x*. It is assumed that there is a food supply named *F* in the search area, which is the objective of the swarm. An equation is suggested to update the leader's position. Salps are translucent, barrel-shaped animals that move by pushing water through their gelatinous bodies as they contract. They belong to the family Salpidae. To get phytoplankton to eat, they internally filter the water. The flow of water through their body helps them propel forward. Fig. 2 shows a picture of a single salp. One notable characteristic of salps is their propensity to swarm. Salps frequently form swarms in deep ocean conditions termed salp chains, as seen in Fig. 2. Individual salps, also known as blastozooids, stay attached while swimming and eating, progressively growing in size. While the precise cause of this behavior is yet unknown, several scientists hypothesize that synchronized movement and feeding may improve locomotor efficiency (Eq. (1)).

$$x_j^1 = \begin{cases} F_j + c_1((ub_j - lb_j)c_2 + lb_j) & c_3 \geq 0 \\ F_j - c_1((ub_j - lb_j)c_2 + lb_j) & c_3 < 0 \end{cases} \quad (1)$$

x_j^1 represents the leader's position in the j^{th} dimension, F_j is the food source's location in the j^{th} dimension, ub_j is the upper limit of the j^{th} dimension, lb_j is the lower bound of the j^{th} dimension, c_1, c_2, and c_3 are random integers.

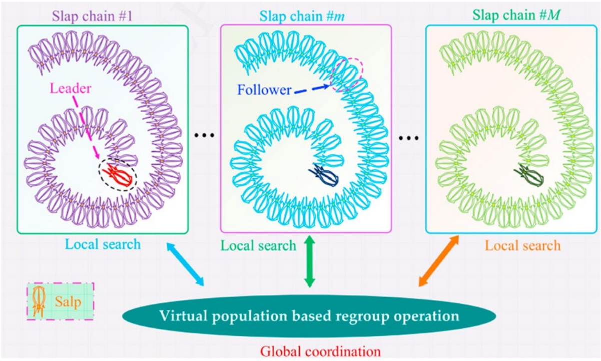


Fig. 2. Optimization framework of MSSA.

Eq. (2) demonstrates that the leader adjusts its location about the food supply. The coefficient c_1 is crucial in SSA since it balances exploration and exploitation.

$$c_1 = 2e^{-\left(\frac{4l}{L}\right)^2} \quad (2)$$

where *"l"* represents the current iteration and *"L"* represents the maximum number of iterations. The parameters c_2 and c_3 are randomly generated numbers within the interval [0, 1]. They determine whether the next point in the j^{th} dimension should go towards positive or negative infinity and the magnitude of the step. The followers' location is updated using the following equations based on Newton's law of motion (Eq. (3)).

$$x_j^i = \frac{1}{2}at^2 + v_0 t \quad (3)$$

where $i \geq 2, x_j^i$ displays the location of the i^{th} follower salp in the j^{th} dimension at time t, with v_0 representing the beginning velocity. $a = \frac{v_{final}}{v_0}$ where $v = \frac{x - x_0}{t}$, The disparity between iterations in optimization is equal to 1, and with $v_0 = 0$, the equation may be stated as follows (Eq. (4)):

$$x_j^i = \frac{1}{2}(x_j^i + x_j^{i-1}) \tag{4}$$

where x_j^i the position of i^{th} follower salp in j^{th} dimension.

IV. The Proposed Memetic Design

The paper presents a new method that utilizes the Memetic Salp Swarm Algorithm (MSSA) to improve the search capability of the Salp Swarm Algorithm (SSA). The enhancement is accomplished by using the memetic computing paradigm and specifically integrating the symmetrical uncertainty ranker filter technique. The suggested method outlines three architectural frameworks: sequential global-first hybridization, sequential global-last hybridization, and interleaved. Every architectural design provides distinct benefits and is crucial in enhancing the optimization procedure in evolutionary algorithms. The paper attempts to clarify the effectiveness of MSSA and add to the growing field of metaheuristic optimization approaches through a thorough investigation.

MSSA lacks explicit mathematical proofs but its convergence behavior agrees with empirical observations in swarm and memetic algorithms. For instance, SSA's convergence towards a global optimum is numerically guaranteed under decreasing stochasticity (c_1 rightarrow 0) and large iteration [20]. Additional studies could strictly establish MSSA's convergence by using Lyapunov stability requirements or leveraging Markov chain theory, similar to PSO and GA convergence studies [10, 27]. Additionally, drawing from strict frameworks like No-Free-Lunch theorems may place MSSA's exploration-exploitation trade-offs [7].

The suggested method has the advantage of balancing both the explorative and the exploitative sides to ensure better optimization performance. However, combining both approaches can lead to an extended processing time from the architectural point of view. Still, this method can be used suitably with different optimization processes, which are not limited to feature selection. The main two steps are applying the local and global searches in various sequences, and both are explained below:

1. Local search within each chain: MSSA comprises multiple parallel salp chains, each representing a swarm of salps. Each salp chain independently conducts a local search based on the SSA search strategy during each iteration.
2. Global coordination within the virtual population: The MSSA population functions solely as hosts for memes, where a meme serves as a unit of cultural evolution [28]. Memes are selected for enhanced communicability among the hosts (e.g., the salps in MSSA). Additionally, the physical attributes of individual salps remain unchanged during global coordination. Consequently, all salps are treated as part of a virtual population. This virtual population is then reorganized into multiple new salp chains, facilitating global coordination among different salp chains.

Global cooperation is required within this population. Memes are chosen based on their increased capacity to be communicated among the hosts, such as the salps in MSSA. Furthermore, the physical characteristics of individual salps stay constant during global coordination. Therefore, all salps are regarded as members of a simulated population. The virtual population is rearranged into several new salp chains, promoting global coordination across distinct ones. The global search was expounded upon in the part before this one. Detailed explanations of the combination approaches utilized in this work's design are provided in this part. The acronym Correlation-Based Search Mechanism (CBSM) refers to the Filter search-based feature selection technique, which is utilized in the implementation of the k-means clustering mechanism. Feature selection is accomplished by using the Salp search iterative approach, which is referred to as the global search-based method of feature selection (G-SSA). In addition, we have the Memetic Salp Swarm Algorithm (M-SSA) technique, which is a clustering method that makes use of memetic optimization technology [29]. In this study, two methods of M-SSA are presented: sequential hybridizations and interleaved hybridizations to be specific. CBSM (GS(X_t), LS(X_t)) equals X_{t+1}.

where:

X_t represents the population of solutions at iteration t.

X_{t+1} represents the updated population of solutions at iteration t+1.

$GS(X_t)$ denotes the application of the global search operator on the current population X_t.

$LS(X_t)$ denotes the application of the local search operator on the current population X_t.

CBSM (A, B) represents the correlation-based selection mechanism applied to solutions from sets A and B.

This model iterates until a termination criterion is met, such as reaching a maximum number of iterations or achieving a desired level of solution quality.

The convergence properties of the MSSA algorithm originate from its memetic hybridization of global exploration (SSA) and exploitation locally (SU-based filtering). SSA ensures exploration due to leader-follower dynamics, where leader position update (Eqs. (1) and (2)) balances stochasticity and decay to push the swarm towards optimas globally. Local SU filtering tunes solutions with respect to higher relevance features and accelerates convergence through reduced dimensionality of search spaces. Theoretical guarantees of convergence for swarm-inspired approaches like SSA are usually based on probabilistic completeness, i.e., iterative population updates ensure asymptotic convergence subject to parameter conditions (e.g., diminishing exploration coefficient (c_1) [20, 21].

The Memetic Salp Swarm Algorithm (MSSA) adopts the memetic computing framework to achieve the search capabilities of the Salp Swarm Algorithm (SSA). In M-SSA, the main goal of the exploration phase is to fully explore the solution space and stop the problem from settling on bad solutions too quickly, which is a common problem in traditional SSA [30]. Thus, the proposed method can be summarized to be arranged in two designs in which the sequential or an interleaved design are used as will be explained in the following subsections.

A. Sequential Global-First Hybridization

Two modules are incorporated in the Sequential Hybridization, the local search and the M-SSA modules. Fig. 3 is a flowchart of the sequential hybridization where the global search wrapper is applied first and later the local search is applied.

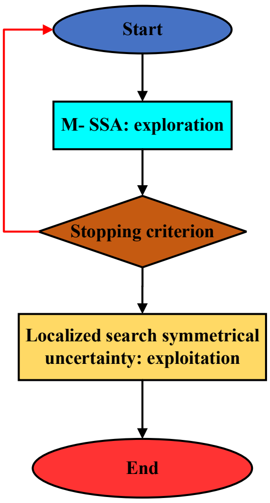


Fig. 3. Sequential global-first hybridization.

SU's entropy-based ranking of features removes redundancy and focuses the search effort in directions of high value within the solution space. SU prevents irrelevant features at the initial stage, hence mitigating the risk of local optima traps and computational cost, essentially lowering the effective dimensionality of the problem. This is in accordance with filter-wrapper convergence concepts, where feature subset quality is a direct function of fewer iteration cycles [3, 17]. The synergy of SU as a local search operator ensures MSSA to satisfy monotonic improvement conditions, a key condition for convergence in hybrid optimization [18].

In localized search utilizing Symmetrical Uncertainty (SU), SU assesses the pertinence of characteristics to the target class and the redundancy among features. SU facilitates the search by emphasizing traits with high relevance and little repetition [31]. During the search process, characteristics are aggregated according to their SU values, and the optimal centroids denoting the most informative features are chosen. These centroids optimize relevance to the target class while reducing overlap with other attributes. This method guarantees that the search prioritizes identifying the most significant characteristics for the classification problem, enhancing overall feature selection efficacy [32].

Sequential and interleaved hybridization strategies (Section IV) directly influence convergence rate. Sequential global-first hybridization encourages broad exploration, delaying refinement but not premature stagnation, whereas global-last hybridization employs local optima prematurely and possibly with nonoptimal convergence. Interleaved hybridization dynamically balances between these two stages, as efficient memetic systems that combine global stochastic search with deterministic local refinements should have polynomial-time convergence in feature selection [18, 22]. Experimental results verify this, with higher F-scores for interleaved MSSA than for isolated SSA or SU.

First, the M-SSA identifies the promising areas in the search space while the k-means local search will try to find the best centroids location within the area identified by the M-SSA. The outcomes of the SSA module is utilized as the first step in the optimization process seed of K-means module. The advantage of this approach is to ensure exploration of the search space for leading to a global optimum. However, it has slower convergence, as local refinement is delayed until after broad exploration [33].

B. Sequential Global-Last Hybridization

In this hybridization scheme the local search is applied first as a one step for the exploitation and then the M-SSA module is applied [30]. Fig. 4 is a flowchart of the sequential global-last hybridization.

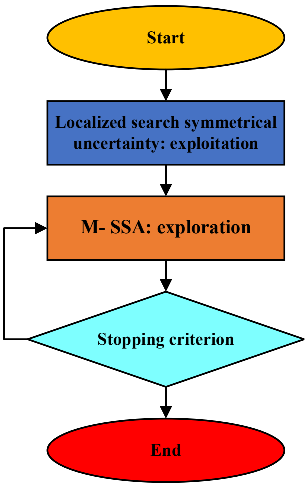


Fig. 4. Sequential global-last hybridization.

C. Interleaved Hybridization

In this scheme, the local search algorithm is combined inside the SSA. In specific, the local searcher utilizes the optimal vector from population as the starting point after every iteration. Fig. 5 shows the interleaved hybridization.

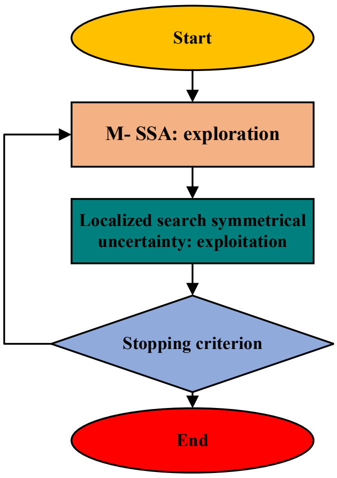


Fig. 5. Interleaved hybridization.

The population will be updated in the case where locally adjusted solutions have better fitness values compared to the ones that are in the entire population and such process continuous till stopping criteria is met. Thus, the algorithms suggested in this work are the sequentially combined M-SSA and the interleaved combined M-SSA. For the next section, we will refer to both as Sequentially Combined M-SSA (SM-SSA) and Interleaved Combined M-SSA (IM-SSA), respectively. The test of both methods will be conducted against SSA, i.e., the wrapper-based feature selection and the CBSM local based clustering).

The specific exploration strategy used within the module can be explained in terms of how it balances the trade-off between exploration (searching broadly across the solution space) and exploitation (focusing on refining promising solutions) [34].

D. Memetic Optimization in Feature Selection

The memetic method used for feature selection is based on the idea of combining two types of feature selection methods and treating them as one method. The different studies showed that this technique has more accuracy in terms of the retrieved results. Usually, the methods used for this purpose can be classified into wrapper and filter methods. The first indicates the methods using the iterative methods such as the genetic algorithms while the filter methods are inline methods which are based on only filtering the set of features once without iterating. In Fig. 6, the two models used for the feature selection are GA and MA-based models. In this work, we first tested the GA base shown in Fig. 7.

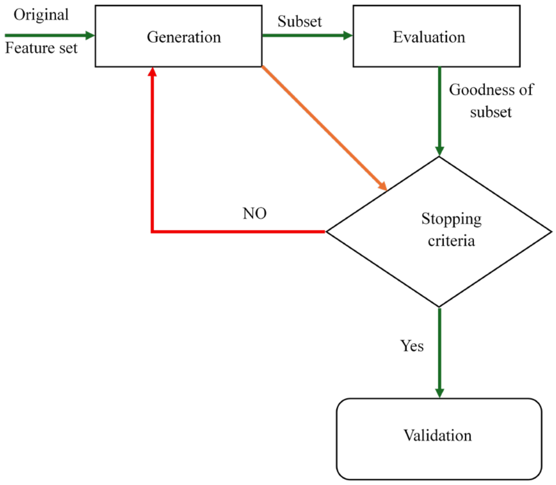


Fig. 6. Feature selection using GA.

E. Pre-processing

The text in its original format is not ready for the further processes of the feature selection. Thus, the following steps are applied to transform the text into a structural format.

1. Text Collection: Sources (e.g., datasets, Application Programming Interface (APIs) and inclusion/exclusion criteria for documents.
2. Text Cleaning: Concrete steps like lowercasing, removing URLs/special characters, punctuation, and non-alphanumeric symbols using regex or libraries like "NLTK".
3. Tokenization & Normalization: Tokenization method (e.g., word vs. subword), lemmatization/stemming (e.g., "spaCy" or "PorterStemmer"), and stop word removal (specifying the list used, e.g., NLTK's 179-word list).
4. Vector Space Modeling: Explicit mention of techniques (e.g., Term Frequency–Inverse Document Frequency (TF-IDF), Bag of Words (BoW) and tools (e.g., "scikit-learn"), including hyperparameters like n-gram range or max document frequency.
5. Dimensionality Analysis: Quantitative comparison of vector space size (e.g., vocab size, sparsity) before/after preprocessing steps like stop word removal or stemming.
6. Similarity Calculation: Metrics (e.g., cosine similarity) and justification for their use.

The revised section will explicitly link preprocessing choices (e.g., stop word lists, tokenization) to their impact on downstream tasks (e.g., vector space sparsity, similarity results), addressing methodological transparency.

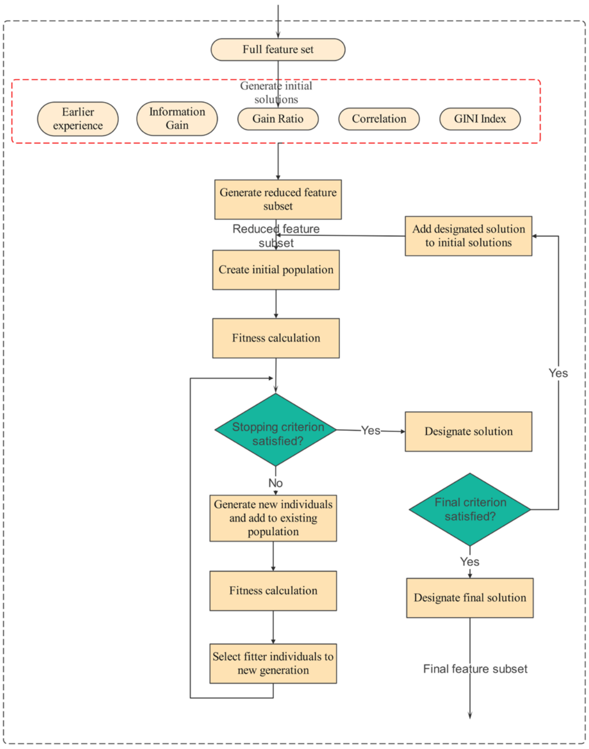

Full feature set
Generate initial solutions
Earlier experience
Information Gain
Gain Ratio
Correlation
GINI Index
Generate reduced feature subset
Reduced feature subset
Add designated solution to initial solutions
Create initial population
Fitness calculation
Yes
Stopping criterion satisfied?
Yes
Designate solution
No
Generate new individuals and add to existing population
Final criterion satisfied?
Fitness calculation
Yes
Designate final solution
Select fitter individuals to new generation
Final feature subset

Fig. 7. Feature selection using MA.

V. Dataset

Open Source Arabic Corpora (OSAC) is sourced from several British Broadcasting Corporation (BBC) and Cable News Network (CNN) Arabic websites. The corpus consists of 5843 text documents divided into six groups, each containing 300 to 2000 raw documents, as indicated in Table I. The corpus is partitioned into two parts to construct the training and testing data for the classification system. It is important to observe that this corpus has already been preprocessed, so the preprocessing stage is omitted in this study.

TABLE I. OSAC ARABIC DATASET

Class	Number of documents
Social Economy	3299
Social Law	944
Sport	717
Religions-Islam-General	210
Science-Applied Health	373
Pure science-Astronomy	300

These problems arise from multiple factors:

1. An inconsistency problem frequently arises, particularly during transitions between various encoding systems (e.g., Unicode versus CP-1256). Inaccurate encoding may result in improper text presentation, compromising the dataset's integrity.
2. Arabic has a complicated morphology, with affixes and clitics that change the shape of words. This could make processing harder, causing noise or mistakes in automated parsing or feature extraction.
3. The Arabic language has synonymy and ambiguity due to various meanings for a single word, and elevated synonymy can lead to inconsistencies, mainly if not managed appropriately during preprocessing or annotation.
4. The Arabic language has vowel omission, which means Short vowels are frequently missed in written Arabic, resulting in ambiguity and noise in data unless appropriate preprocessing or discretization is used.

These elements jointly influence the dataset's quality, generating noise and inconsistencies that must be meticulously addressed using preprocessing approaches.

VI. TEST RESULTS

This section details the experimental results we acquired after evaluating the suggested strategy in several settings.

A. *Local, Global and Hybrid Methods Comparison*

This section details the classification test results obtained by comparing the proposed technique's two components: the SSO as the wrapper global search and the Symmetrical Uncertainty (SU). Table II displays the test results of applying three feature selection methods to the OSAC dataset. Three classifiers utilized are Naïve Bayesian, J48, and SVM.

TABLE II. TEST RESULTS OF EVALUATING THE PROPOSED METHOD WITH THE SU (LOCAL SEARCH) AND THE MSSO (GLOBAL SEARCH)

Collected data	classification	precision	recall	F-measure
Memetic Salp Swarm Optimization-based Subset Selection Using Uncertainty (M-SSOSU)	Naïve Bayesian	0.96	0.99	0.99
	J48	0.95	0.9	0.9
	SVM	0.92	0.93	0.96
	Naïve Bayesian	0.92	0.92	0.91
	J48	0.92	0.92	0.92
	SVM	0.99	0.93	0.9
SSO	Naïve Bayesian	0.9	0.91	0.9
	J48	0.91	0.9	0.92
	SVM	0.92	0.91	0.91
Interleaved-MSSO	Naïve Bayesian	0.93	0.9	0.9
Improved Memetic Salp Swarm Optimization (I-MSSO)	J48	0.92	0.92	0.88
	SVM	0.92	0.9	0.91
Sequential Memetic Salp Swarm Optimization (M-SSO)	Naïve Bayesian	0.99	0.98	0.99
Simplified Memetic Salp Swarm Optimization (S-MSSO)	J48	0.95	0.95	0.92
	SVM	0.93	0.95	0.95
Sequential M-SSO	Naïve Bayesian	0.98	0.95	0.94
(S-MSSO)	J48	0.92	0.92	0.91
	SVM	0.9	0.92	0.95

B. *Comparison with Other State-of-the-Art Methods*

This section presents the results of a thorough comparison between the suggested strategy and accepted approaches commonly recognized as the best in the area. Our assessment is based on using the OSAC dataset, a common benchmark dataset used in related research. One methodology being examined is the speed, accuracy, and feature relevance (BSO-CHI-SVM) technique [35], which is well-known for its effectiveness in optimizing feature representation via iterative swarm-based algorithms. On the other hand, Marie-Sainte *et al.* [36] outlines a technique for developing an Arabic Text Categorizer system (ATC-FA) that combines Support Vector Machine (SVM) classifiers with the evolutionary algorithmic Firefly optimization method. By contrasting these methodologies' performance measures with the suggested methodology, more will be known about the strengths and weaknesses of each strategy. The classification accuracy, computational efficiency, scalability, and robustness will be used as performance metrics. This comparative research highlights the possibility for innovation and improvement in text categorization algorithms and provides insights into the relative merits of the strategies under investigation.

C. *Baseline Methods Comparison*

Other feature selection approaches, including the well-known rankers, were used when utilizing attribute evaluation algorithms. Their application's functionality is likewise contingent on the assessment technique employed. The method calculates the coefficients used to determine the ranking. Decrease the value to be eliminated from the database. Our investigation employed the following approach: A suitable classifier was chosen and used to classify data using various subsets of characteristics. They are defined as the following: Initially, the characteristic with the lowest rank is eliminated. Subsequently, the two least-ranking qualities are eliminated in succession. This subgroup was chosen based on achieving maximum classification accuracy. To identify the best selection method (Table III), classification is conducted using the optimum attribute selected subsets for each algorithm. The indications from the categorization are being contrasted. This algorithm was chosen based on the indicators True Positive (TP) rate, precision, and F-measure, which have the highest values. Various methods may be utilized to identify the most suitable subset of features for each algorithm. The most often employed method is to establish a threshold value for the coefficient and any characteristics that possess it. Tables III–V compare MSSO with six feature selection methods (One_R, TF-IDF, Information Gain Correlation Attribute Eval, Classifier Attribute Eval, One_R and Cfs_Sub_set_Eval) implemented in WEKA. Precision (P), Recall (R), and F-measure (F) are metrics used to evaluate the categorization accuracy (Table V).

TABLE III. TESTING THE PROPOSED METHOD WITH OTHER STATE-OF-THE-ART METHODS

Class	FAFS			BSO-CHI-SVM			MSSO		
	Pr	**Recall**	**F-score**	**Pe**	**Recall**	**F-score**	**Pr**	**Recall**	**F-score**
Social Economy	1	0.96	0.96	0.99	0.99	1	0.99	0.99	1
Social Law	1	0.93	0.96	1	1	1	1	1	1
Sport	1	1	1	0.99	1	0.99	0.99	1	0.99
Religions-Islam-General	0.9	0.83	0.86	0.96	0.96	0.92	1	1	0.92
Science -Applied Health	0.9	0.96	0.94	0.99	0.99	0.98	0.99	0.99	1
Pure Science-Astronomy	1	1	1	0.99	0.99	0.98	0.99	0.99	1
Weighted Avg.	1	0.95	0.95	0.99	0.99	0.99	0.99	1	1

FAFs: Firefly Algorithm-based Feature Selection; Pr: Probability of feature selection; Pe: Elitism probability

TABLE IV. BASELINE METHODS COMPARISON

Class	Correlation_AttributeEval			Classifier_Attribute Eval			Cfs_SubsetEval			MSSO		
	P	**R**	**F**	**P**	**R**	**F**	**P**	**R**	**F**	**P**	**R**	**F**
Social Economy	0.96	0.98	0.99	0.96	0.98	0.99	0.98	0.99	0.99	0.99	1	1
Social Law	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	1	1	1	1
Sport	0.94	0.95	0.91	0.94	0.95	0.91	0.96	0.97	0.95	1	1	1
Religions-Islam-General	0.9	0.9	0.92	0.9	0.9	0.92	1	1	1	0.99	0.99	0.97
Science-Applied Health	1	1	1	1	1	1	1	0.99	0.99	1	1	0.99
Pure Science-Astronomy	1	1	1	1	1	1	1	1	1	1	0.99	0.99
Weighted Avg.	0.98	0.98	0.98	0.98	0.98	0.98	0.99	0.99	0.99	1	1	1

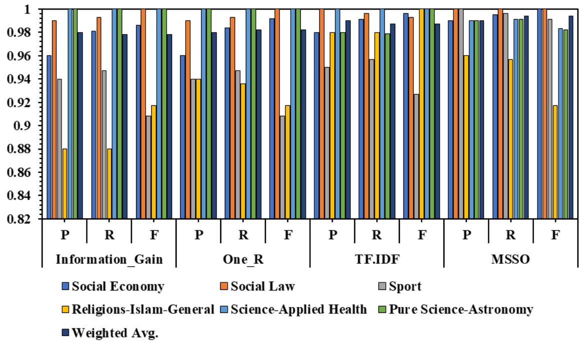
TABLE V. SECOND GROUP OF BASELINE METHODS COMPARISON

Class	Information_Gain			One_R			TF.IDF			MSSO		
	P	**R**	**F**	**P**	**R**	**F**	**P**	**R**	**F**	**P**	**R**	**F**
Social Economy	0.96	0.981	0.986	0.96	0.984	0.992	0.98	0.99	0.9	0.99	0.995	1
Social Law	0.99	0.993	1	0.99	0.993	1	1	0.996	0.99	1	1	1
Sport	0.94	0.947	0.908	0.94	0.947	0.908	0.95	0.957	0.927	1	0.996	0.991
Religions-Islam-General	0.88	0.88	0.917	0.94	0.936	0.917	0.98	0.98	1	0.96	0.957	0.917
Science-Applied Health	1	1	1	1	1	1	1	1	1	0.99	0.991	0.983
Pure Science-Astronomy	1	1	1	1	1	1	0.98	0.979	1	0.99	0.991	0.982
Weighted Avg.	0.98	0.978	0.978	0.98	0.982	0.982	0.99	0.987	0.987	0.99	0.994	0.994

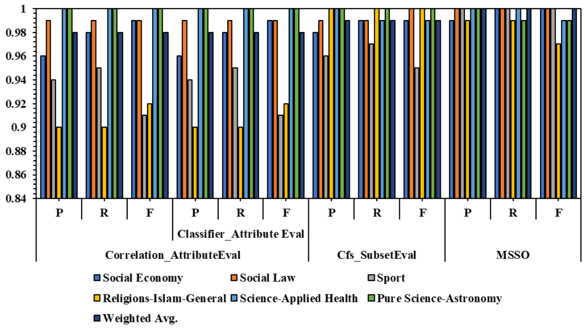
Based on the outcomes obtained from the conducted experiments, it is possible that the evaluation was performed using the Multi-Strategy Symmetrical Uncertainty-based Optimization (MSSO) approach. The evaluation metrics employed include Precision, recall, and F-score, which collectively offer a comprehensive view of the model's classification performance by capturing its accuracy and robustness in identifying relevant features [37].

MSSO consistently demonstrates superior performance across all three metrics compared to the other two approaches under consideration. This superiority can be attributed to the inherent capability of MSSO to enhance the effectiveness of localized search, allowing it to focus more intelligently on promising regions of the feature space. By integrating symmetrical uncertainty into its search mechanism, MSSO can prioritize feature subsets that are both highly relevant and minimally redundant, thus improving the quality of the selected features. The strength of MSSO lies in its multi-strategy optimization process, which combines exploration and exploitation techniques to avoid local optima while refining the search for high-quality solutions. This balanced strategy increases the likelihood of identifying optimal feature subsets and accelerates the search process's convergence.

These advantages are clearly reflected in the experimental results. Fig. 8 visually presents the performance comparisons, corroborating the data summarized in Tables I and V. The figure illustrates the consistent improvement achieved by MSSO, providing further empirical support for its effectiveness and reliability in feature selection tasks. Such outcomes affirm the potential of MSSO as a robust and efficient method for high-dimensional data analysis in real-world applications.



(a)



(b)

Fig. 8. Baseline methods comparison for two groups: (a) First group (b) Second group.

D. Parameters Incorporated in the Salp Swarm Algorithm (SSA)

The performance of the Memetic Salp Swarm Algorithm (MSSA) computationally is heavily based on parameters native to the SSA component, such as the values (c_1), (c_2), and (c_3), population size, and number of maximum iterations (L). The (c_1) term, ($c_1 = 2e^{-(4l/L)2}$), dynamically balances exploration and exploitation by controlling the step size of the motion of the leader salp to the food (Eq. (2)). The higher (c_1) in early iterations supports exploration, allowing the swarm to traverse a range of regions of the search space. (c_1) decreases with ongoing iterations, directing focus towards exploitation for refinement of solutions. If (c_1) decays too rapidly, the algorithm has the potential to prematurely converge on poorer feature sets; conversely, slow decay potentially becomes unnecessarily long and increases computational expense. The parameters (c_2) and (c_3), random numbers in [0, 1], introduce randomness to avoid local optima but must be adjusted so that too much randomness does not destabilize convergence. Population size regulates candidate solution diversity larger populations favor exploration but add to computational cost, particularly in high-dimensional Arabic text corpora. Small populations risk lack of diversity and stagnation. The maximum iteration number (L) regulates the algorithm running time; insufficient iterations can truncate convergence, while excessive iterations are wasteful. Experimental results (Table III) highlight that SSA alone achieved lower F-scores (e.g., 0.91 for SVM) compared to MSSA (0.95 for SVM), highlighting the importance of balancing these parameters in order to optimize both accuracy and efficiency.

E. Symmetrical Uncertainty (SU) Integration and Threshold Settings

The incorporation of Symmetrical Uncertainty (SU) as a filter-based local search process has parameters that have a direct influence on feature relevance estimation and redundancy elimination. SU measures mutual information between features and the target class, and the greater the value, the greater the relevance. The choice of threshold for selecting features according to SU rankings controls subset size and quality. An SU cutoff with high value (stringent threshold) can filter out features that are slightly useful, potentially decreasing the classification rate, as evident from Table III when individual SU had lower F-scores (e.g., 0.91 for Naïve Bayes) than the hybrid MSSA (0.99). An SU cutoff with low value (suspicious threshold) retains duplicate features, increasing dimension and computational demands. The use of SU in a memetic framework in the paper ensures that SU-guided local search improves global SSA outputs by eliminating redundancy while preserving significant features. For instance, in Section IV, SU's role in clustering high-SU features around centroids ensures that the algorithm prefers informative features like morphological roots in Arabic text, overcoming problems like omission of vowels and synonymy. However, imprecise SU thresholds can disrupt this balance—over-pruning removes linguistically nuanced features, while under-pruning introduces noise. The success of the hybrid solution (Table IV: MSSA's 1.0 F-score on "Social Law" vs. BSO-CHI-SVM's 1.0) depends on calibrated integration of SU, demonstrating that parameterizing SU thresholds based on dataset characteristics (e.g., class skewness) is key to robustness.

F. Hybridization Strategies and Memetic Balancing

The hybridization approach sequential global-first, sequential global-last, or interleaved—has an impact on MSSA's ability to balance exploration and exploitation. Sequential global-first (Section IV) gives more importance to SSA's global search in order to traverse extensive regions of the feature space before applying SU-based local refinement. This approach is more effective for complex search landscapes but has the potential to bring about convergence delay, as evident from Table III's sequential global-first MSSA achieving higher F-scores (0.99 for Naïve Bayes) at the cost of longer runtime. Conversely, sequential global-last (Section IV) executes SU initially in order to pursue locally promising features with faster early convergence but risks leaving the algorithm in suboptimal areas should local search become too restrictive. Interleaved hybridization (Section IV) interleaves a mixture of global and local searches with dynamic switching among them to better adapt to the search space. This strategy achieved competitive performance (e.g., 0.91 F-score for SVM in Table III) by diversity via global exploration from time to time and local solution optimization. Memetic balance between SSA and SU is also influenced by how often local search is used.

Excessive iterations of local search can disperse the coherence of the swarm, and underapplication may fail to eliminate redundancies. For example, the better performance of sequential global-first MSSA on the OSAC dataset (Table IV: 1.0 weighted F-score) results from its orderly exploration-exploitation sequence, modified according to the high dimension of Arabic text. Thus, choice of hybridization strategy and memetic optimization should be dataset-dependent, solving problems such as class imbalance (addressed via SMOTE) and morphological richness, for optimal accuracy and efficiency in computation.

VII. Conclusions

Text data is usually associated with high-dimensional feature spaces. Optimization approaches can be used efficiently with Feature Selection (FS) and are increasingly adopted. Our research uses the Salp Swarm Algorithm (SSA) as a wrapper feature selection method. Like other optimization methods, SSA struggles to maintain population variety and avoid local optima. To address these issues, this work introduces the Memetic Salp Swarm Algorithm (MSSA), an upgraded variant of SSA. Combining SSA with Symmetrical Uncertainty creates the Memetic Salp Swarm Algorithm (MSSA). The suggested method reduces computing complexity and improves classification accuracy to find the best Arabic text characteristics. MSSA effectively finds the optimal subsets of attributes by integrating salp behaviour with memetic algorithms. Integrating Symmetrical Uncertainty ranking enhances selection by precisely assessing feature value to the classification aim. The suggested solution outperformed current methods in classification accuracy and feature subset size on Arabic text datasets. The test results showed that the proposed method outperformed the other state-of-the-art methods regarding the evaluation criteria utilized.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Mohammed Ghassan Abdulkareem: Writing—Original Draft; Alhasan Amer Ibrahim: Writing—Review & Editing and Ibrahem Amer Hammed: experimental demonstration and data collection, G. Abdulkareem-Alsultan: Writing—Original Draft, Writing—Review & Editing, Supervision. All authors had approved the final version.

Acknowledgment

The authors wish to thank Prof. Dr. Alsultan from PutraCAT for helping us to revise and edit this work.

References

[1] A. Jović, K. Brkić, and N. Bogunović, "A review of feature selection methods with applications," in *Proc. 2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, 2015, pp. 1200–1205.

[2] Z. Xu *et al.*, "Discriminative semi-supervised feature selection via manifold regularization," *IEEE Transactions on Neural networks,* vol. 21, no. 7, pp. 1033–1047, 2010.

[3] N. Hoque, D. K. Bhattacharyya, and J. K. Kalita, "MIFS-ND: A mutual information-based feature selection method," *Expert Systems with Applications,* vol. 41, no. 14, pp. 6371–6385, 2014.

[4] I. H. Witten *et al.*, "Practical machine learning tools and techniques," in *Data Mining*, San Francisco: Morgan Kaufmann Publishers In., 2005, ch. 2, pp. 403–413.

[5] Y. Farida *et al.*, "Comparing support vector machine and naïve bayes methods with a selection of fast correlation based filter features in detecting Parkinson's disease," *Lontar Komputer: Jurnal Ilmiah Teknologi Informasi,* vol. 14, no. 2, pp. 80–90, 2023.

[6] V. V. Bukhtoyarov *et al.*, "A study on a probabilistic method for designing artificial neural networks for the formation of intelligent technology assemblies with high variability," *Electronics,* vol. 12, no. 1, 215, 2023.

[7] X. S. Yang, "Nature-inspired optimization algorithms: Challenges and open problems," *Journal of Computational Science,* vol. 46, 101104, 2020.

[8] P. Z. Lappas and A. N. Yannacopoulos, "A machine learning approach combining expert knowledge with genetic algorithms in feature selection for credit risk assessment," *Applied Soft Computing,* vol. 107, 107391, 2021.

[9] J. J. Grefenstette, "Genetic algorithms and machine learning," in *Proc. Sixth Annual Conf. on Computational Learning Theory*, 1993, pp. 3–4.

[10] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. of ICNN'95-International Conf. on Neural Networks*, 1995, pp. 1942–1948.

[11] K. Price, R. M. Storn, and J. A. Lampinen, *Differential Evolution: A Practical Approach to Global Optimization*, Heidelberg: Springer Berlin, 2006, ch. 1.

[12] O. A. Alomari *et al.*, "Gene selection for microarray data classification based on gray wolf optimizer enhanced with TRIZ-inspired operators," *Knowledge-Based Systems,* vol. 223, 107034, 2021.

[13] L. Zhang *et al.*, "Feature selection using firefly optimization for classification and regression models," *Decision Support Systems,* vol. 106, pp. 64–85, 2018.

[14] M. A. Awadallah *et al.*, "Binary JAYA algorithm with adaptive mutation for feature selection," *Arabian Journal for Science and Engineering,* vol. 45, pp. 10875–10890, 2020.

[15] A. I. Hammouri *et al.*, "An improved dragonfly algorithm for feature selection," *Knowledge-Based Systems,* vol. 203, 106131, 2020.

[16] M. Mafarja, "Binary dragonfly optimization for feature selection using time-varying transfer functions," *Knowledge-Based Systems,* vol. 161, pp. 185–204, 2018.

[17] E. O. Abiodun *et al.*, "A systematic review of emerging feature selection optimization methods for optimal text classification: The present state and prospective opportunities," *Neural Computing and Applications,* vol. 33, pp. 15091–15118, 2021.

[18] W. Seo *et al.*, "Effective memetic algorithm for multilabel feature selection using hybridization-based communication," *Expert Systems with Applications,* vol. 201, 117064, 2022.

[19] P. Agrawal *et al.*, " Metaheuristic algorithms on feature selection: A survey of one decade of research (2009–2019)," *IEEE Access,* vol. 9, pp. 26766–26791, 2021.

[20] A. D. Li, B. Xue, and M. Zhang, "Improved binary particle swarm optimization for feature selection with new initialization and search space reduction strategies," *Applied Soft Computing,* vol. 106, 107302, 2021.

[21] S. J. Berlin and M. John, "Particle swarm optimization with deep learning for human action recognition," *Multimedia Tools and Applications,* vol. 79, pp. 17349–17371, 2020.

[22] P. Samal and D. Panigrahy, "Simultaneous feeder reconfiguration, DSTATCOM allocation, and sizing using seagull optimization algorithm in unbalanced radial distribution systems," *Soft Computing,* vol. 28, pp. 6403–6421, 2024.

[23] S. N. Chegini, A. Bagheri, and F. Najafi, "PSOSCALF: A new hybrid PSO based on sine cosine algorithm and levy flight for solving optimization problems," *Applied Soft Computing,* vol. 73, pp. 697–726, 2018.

[24] S. Jain and R. Dharavath, "Memetic salp swarm optimization algorithm based feature selection approach for crop disease detection system," *Journal of Ambient Intelligence and Humanized Computing,* vol. 14, pp. 1817–1835, 2021.

[25] Z. Tao *et al.*, "GA-SVM based feature selection and parameter optimization in hospitalization expense modeling," *Applied Soft Computing,* vol. 75, pp. 323–332, 2019.

[26] Y. Liang and L. Wang, "Applying genetic algorithm and ant colony optimization algorithm into marine investigation path planning model," *Soft Computing,* vol. 24, pp. 8199–8210, 2019.

[27] S. Hosseini and B. M. H. Zade, "New hybrid method for attack detection using combination of evolutionary algorithms, SVM, and ANN," *Computer Networks,* vol. 173, 107168, 2020.

[28] R. Shuler, "A theoretical treatment of memetic traits using gene-meme, meme-meme and population equilibrium," Preprints.org, January 2021. https://doi.org/10.20944/preprints202012.0689.v2

[29] J. Li *et al.*, "A novel and efficient salp swarm algorithm for large-scale QoS-aware service composition selection," *Computing,* vol. 104, pp. 2031–2051, 2022.

[30] S. Kassaymeh *et al.*, "Optimizing beyond boundaries: Empowering the salp swarm algorithm for global optimization and defective software module classification," *Neural Computing and Applications,* vol. 36, pp. 18727–18759, 2024.

[31] T. Wang *et al.*, "Research on stochastic feature selection optimization algorithm fusing symmetric uncertainty," in *Proc. 2024 IEEE 9th International Conf. on Data Science in Cyberspace (DSC)*, 2024, pp. 549–554.

[32] M. Z. Ali *et al.*, "Advances and challenges in feature selection methods: A comprehensive review," *J. Artif. Intell. Metaheuristics*, vol. 7, no. 1, pp. 67–77, 2024.

[33] R. Priyadarshi and R. R. Kumar, "Evolution of swarm intelligence: A systematic review of particle swarm and ant colony optimization approaches in modern research," *Archives of Computational Methods in Engineering* pp. 1–42, 2025. https://dx.doi.org/10.1007/s11831-025-10247-2

[34] M. Moradi *et al.*, "Adaptive network approach to exploration–exploitation trade-off in reinforcement learning," *Chaos: An Interdisciplinary Journal of Nonlinear Science,* vol. 34, 123120, 2024.

[35] R. Belkebir and A. Guessoum, "A hybrid BSO-Chi2-SVM approach to Arabic text categorization," in *Proc. 2013 ACS International Conf. on Computer Systems and Applications (AICCSA)*, 2013, pp. 1–7.

[36] S. L. Marie-Sainte and N. Alalyani, "Firefly algorithm based feature selection for Arabic text classification," *Journal of King Saud University-Computer and Information Sciences,* vol. 32, no. 3, pp. 320–328, 2020.

[37] R. Diallo, C. Edalo, and O. O. Awe, "Machine Learning Evaluation of Imbalanced Health Data: A Comparative Analysis of Balanced Accuracy, MCC, and F1 Score," in *Practical Statistical Learning and Data Science Methods: Case Studies from LISA 2020 Global Network, USA*, Cham: Springer, 2024, ch. 1, pp. 283–312.

Copyright © 2025 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (CC BY 4.0).