REVIEW

Bio-Inspired Algorithms in NLP Techniques: Challenges, Limitations and Its Applications

Huu-Tuong Ho[1], Thi-Thuy-Hoai Nguyen[2], Duong Nguyen Minh Huy[3] and Luong Vuong Nguyen[1,*]

[1]Department of Artificial Intelligence, FPT University, Danang, 550000, Vietnam
[2]Department of Computer Fundamental, FPT University, Danang, 550000, Vietnam
[3]Department of Business, FPT University, Danang, 550000, Vietnam
*Corresponding Author: Luong Vuong Nguyen. Email: vuongnl3@fe.edu.vn

ABSTRACT: Natural Language Processing (NLP) has become essential in text classification, sentiment analysis, machine translation, and speech recognition applications. As these tasks become complex, traditional machine learning and deep learning models encounter challenges with optimization, parameter tuning, and handling large-scale, high-dimensional data. Bio-inspired algorithms, which mimic natural processes, offer robust optimization capabilities that can enhance NLP performance by improving feature selection, optimizing model parameters, and integrating adaptive learning mechanisms. This review explores the state-of-the-art applications of bio-inspired algorithms—such as Genetic Algorithms (GA), Particle Swarm Optimization (PSO), and Ant Colony Optimization (ACO)—across core NLP tasks. We analyze their comparative advantages, discuss their integration with neural network models, and address computational and scalability limitations. Through a synthesis of existing research, this paper highlights the unique strengths and current challenges of bio-inspired approaches in NLP, offering insights into hybrid models and lightweight, resource-efficient adaptations for real-time processing. Finally, we outline future research directions that emphasize the development of scalable, effective bio-inspired methods adaptable to evolving data environments.

KEYWORDS: Natural language processing; bio-inspired; genetic algorithms; ant colony optimization; particle swarm optimization

1 Introduction

Natural Language Processing (NLP) has become the cornerstone of modern artificial intelligence (AI), enabling machines to understand, interpret, and generate human language [1]. NLP endeavors to bridge the divide between human communication and computer comprehension. By developing sophisticated models and algorithms specifically designed to analyze and interpret natural language data [2] effectively, contemporary researchers are not only refining these tools for practical applications—such as creating spoken dialogue systems, speech-to-speech translation engines, and mining social media for health or financial insights [3]–but are also utilizing them to gauge sentiment and emotion toward various products and services; consequently, NLP has become an integral part of our daily lives, being widely employed in numerous computer applications and mobile devices [4]. The rapid proliferation of NLP applications—ranging from chatbots and sentiment analysis (SA) to machine translation (MT) and speech recognition—has significantly impacted diverse domains, including healthcare, finance, and education [5]. NLP generally encompasses tasks requiring machines to process and understand human language, including

text classification, SA, MT, and speech recognition, each with unique challenges [1]. Text classification categorizes text into predefined classes, such as spam detection or document labeling, relying on feature selection techniques to manage high-dimensional data while retaining relevance. SA extracts subjective information like opinions and emotions from text, facing challenges in interpreting nuanced and ambiguous language. MT demands comprehension of semantic and syntactic structures while addressing polysemy and cultural context. Speech recognition converts spoken language into text, requiring robust feature extraction and optimization to manage accents, noise, and speech variations. Despite advances in deep learning (DL) models like Transformers, which power systems such as BERT and GPT, these models require significant computational resources, large labeled datasets, and careful optimization to balance complexity and generalization. However, as NLP tasks grow in complexity and scale, the challenges of optimizing these systems have intensified. These challenges include high-dimensional feature spaces, extensive model parameters, and the need for adaptability in dynamic environments. To address these issues, researchers increasingly turn to bio-inspired algorithms, which mimic natural processes and evolutionary principles, as a promising alternative to traditional optimization techniques.

Optimization algorithms have progressively incorporated inspirations from nature, physics, and human behavior since the introduction of Differential Evolution (DE) in 1995 [6], a population-based method using mutation, crossover, and selection for non-linear problems, followed by the Ant Colony System (ACS) in 1997 [7], which utilizes ant-inspired pheromone trails for combinatorial optimization, and the Artificial Bee Colony (ABC) algorithm in 2005 [8], emulating honey bee foraging to balance exploration and exploitation in continuous optimization, with Grey Wolf Optimization (GWO) launching in 2014 [9], drawing from grey wolf social structures and hunting strategies to tackle complex problems with minimal parameters, while the Equilibrium Optimizer (EO) [10], introduced in 2020, applies physics-based control volume mass balance models to achieve good convergence in both continuous and discrete optimization tasks, alongside the Student Psychology Based Optimization (SPBO) [11], also from 2020, which simulates students' academic efforts as a unique approach with potential in engineering applications, and most recently, the African Vultures Optimization Algorithm (AVOA) [12], developed in 2021, mimicking the foraging and navigation behaviors of African vultures, further enriching the landscape of nature-inspired optimization techniques. Furthermore, the Virus Colony Search (VCS) algorithm, introduced in 2016 [13], simulates virus diffusion and infection strategies, integrating Gaussian random walks, CMA-ES, and immune response mechanisms to efficiently solve unconstrained benchmark functions and constrained engineering problems, often outperforming established methods like GA, PSO, and DE. More recently, the Artificial Circulatory System Algorithm (ACSA) [14], presented in 2025, models human circulatory regulation of blood pressure and volume, excelling on 16 classic and 12 CEC 2022 benchmark functions, surpassing 42 of 64 metaheuristic algorithms statistically, though it requires problem-specific tuning like other metaheuristics. Among this vast array of algorithms, GA, PSO, and ACO were chosen because of their historical significance, extensive application in the NLP literature, and complementary strengths. These algorithms have successfully solved complex optimization problems across various fields [15,16].

Bio-inspired algorithms are computational techniques that draw inspiration from biological phenomena such as evolution, swarm intelligence, and neural processes [17]. Their adaptability, robustness, and ability to navigate vast search spaces make them well-suited for NLP tasks, which often involve large datasets and intricate decision-making processes. Bio-inspired algorithms offer a promising solution to overcome some limitations of traditional and SA-based NLP approaches by leveraging natural processes to address critical challenges [18]. These algorithms optimize complex, high-dimensional parameter spaces, making them ideal for hyperparameter tuning in NLP models. They are particularly effective in feature selection for tasks like text classification and SA, identifying the most relevant features to reduce computational

overhead and enhance performance. Their adaptability to dynamic data environments ensures suitability for evolving datasets and tasks. At the same time, their integration with neural networks enables the optimization of architectures, weights, and learning parameters, improving accuracy and efficiency. Notable examples include GA for feature selection in text classification to improve accuracy, PSO for enhancing MT system parameters, and ACO for aligning sequences in MT and optimizing resource allocation in speech recognition. A specific application of these principles is demonstrated in a 2006 study by Alba et al. [19], published in Information Processing Letters, which explores the use of metaheuristic approaches, specifically a GA, a CHC algorithm, and simulated annealing (SA), for the NLP task of part-of-speech tagging, comparing their performance against the traditional Viterbi algorithm. The study evaluates these methods using integer and binary encodings and both sequential and parallel implementations on the Brown and Susanne corpora, finding that the GA with integer coding generally outperforms CHC and SA, achieving competitive accuracy (around 97%) with Viterbi, especially in scenarios with poorer statistics or non-Markov models, while parallelism enhances both accuracy and efficiency. The results highlight the potential of evolutionary algorithms as flexible, high-quality alternatives for tagging, capable of handling complex contexts beyond traditional methods, with future research planned to explore additional genetic operators and metaheuristics. Researchers have recently explored hybrid models integrating bio-inspired algorithms with traditional SA and SA techniques. These hybrid approaches leverage the strengths of both paradigms: the optimization capabilities of bio-inspired algorithms and the predictive power of neural networks. For instance, neuroevolution—a process where neural network architectures are optimized using evolutionary algorithms—has been applied to NLP tasks like SA and speech recognition [20—22]. This approach improves model performance and reduces the need for manual architecture tuning. Hybrid models also address the computational challenges associated with large-scale NLP tasks. By incorporating bio-inspired optimization methods, these models can achieve better parameter configurations with fewer iterations, reducing training time and resource consumption.

While bio-inspired algorithms offer significant advantages, their application in NLP is not without challenges. These algorithms can be computationally expensive, especially when dealing with large datasets and complex models. Additionally, their performance depends on parameter tuning, which can be time-consuming and require expertise. However, these challenges also present opportunities for innovation. Researchers are actively exploring lightweight bio-inspired algorithms that reduce computational complexity while retaining optimization capabilities. Moreover, advancements in parallel computing and distributed systems enable efficient scaling of these algorithms for large-scale NLP tasks. Another exciting avenue of research is the development of domain-specific bio-inspired techniques tailored to NLP. For example, algorithms inspired by linguistic or cognitive processes could provide novel solutions for tasks like SA and MT. Integrating bio-inspired algorithms with pre-trained language models like BERT and GPT also holds immense potential, combining the contextual understanding of these models with the optimization prowess of bio-inspired approaches.

This review aims to comprehensively analyze the potential of bio-inspired algorithms for enhancing NLP tasks. Specifically, it seeks to:

- Explore state-of-the-art applications of bio-inspired algorithms in key NLP tasks such as text classification, SA, MT, and speech recognition.
- Analyze bio-inspired approaches' comparative advantages and limitations relative to traditional and SA methods.
- Highlight recent advances in hybrid models and their impact on NLP performance.
- Identify key challenges and propose future research directions to address these challenges and unlock the full potential of bio-inspired algorithms in NLP.

By synthesizing existing research and offering actionable insights, this review aims to contribute to the growing body of knowledge on bio-inspired algorithms and their transformative impact on NLP, paving the way for next-generation systems that are both efficient and adaptable.

2 Overview of Bio-Inspired Algorithms

Bio-inspired algorithms are computational techniques that draw inspiration from natural processes such as evolution, swarm intelligence, and biological neural mechanisms. These algorithms are broadly classified into three major categories: Evolutionary Algorithms (EA), Swarm Intelligence (SI), and Neural-based Algorithms (NA). Evolutionary algorithms, such as GA and DE, are inspired by natural selection and genetic mutation. Swarm Intelligence methods, like PSO and ACO, model the collective behavior of groups of simple agents. Neural-based algorithms take inspiration from biological neural networks, such as Artificial Neural Networks (ANN) and Spiking Neural Networks (SNN). These categories collectively provide powerful techniques for solving complex optimization problems in NLP.

Table 1 presents the application domains of bio-inspired optimization algorithms in NLP, categorizing their utilization across different tasks. These algorithms have been extensively used for solving complex optimization problems. This section provides an overview of key bio-inspired algorithms, their mathematical formulations, and their relevance to NLP tasks. In specific functions of NLP, bio-inspired algorithms offer powerful tools for addressing the optimization challenges inherent in NLP tasks. By leveraging natural principles, these methods provide adaptable, efficient, and robust solutions that complement traditional machine-learning approaches.

Table 1: Application domains of bio-inspired optimization algorithms in NLP

Application domain	ACO	PSO	GA
Text classification	Mesleh and Kanaan (2008)—ACO for Arabic text classification [23]. Aghdam et al. (2008)—ACO-optimized SVM [24]. Nema and Sharma (2015)—ACO-kNN hybrid [25]. Aggarwal and Chhabra (2017)—Rule generation via ACO [26]. Hossain et al. (2022)—ACO for deep features [27]. Gite et al. (2023)—Multi-label ACO model [28]. Tripathy et al. (2023) —Semantic features using ACO [29]. Durga et al. (2024)—ACO with fuzzy logic [22]. Alsaduni et al. (2024) —ACO-LDA hybrid [30].	Kurniawati and Pardede (2018)—PSO for feature term optimization [21]. Kumar and Pathak (2021)—PSO-SVM hybrid [31]. Mandal et al. (2018)—Feature selection via PSO [32]. Machova et al. (2020)—PSO combined with GA [33].	Song and Park (2009)—GA-based biomedical feature selection [34,35]. Oezel (2011)—GA using tagged-terms as features [36]. Uguz (2011)—GA-SVM for Turkish datasets [37]. Chen et al. (2013)—GA for term-weighting [38]. Uysal and Gunal (2014) —Dimensionality reduction using GA [39]. Bidi and Elberrichi (2016)—GA-enhanced naive Bayes [40]. Ghareb et al. (2016)—Lexical pruning using GA [41].

(Continued)

Table 1 (continued)

Application domain	ACO	PSO	GA
Sentiment analysis	Kwon et al. (2014)—Feature optimization with ACO [42]. Ahmad et al. (2019)—ACO with deep sentiment models [43]. Hamdi et al. (2022)—Hybridized ACO-SVM [44]. Kavitha and Lalitha (2024)—ACO-based rule extraction [45].	Kumar Gupta et al. (2015)—PSO-SVM for sentiment polarity [46]. Machova et al. (2020)—Swarm hybrid with GA [33]. Yue et al. (2023)—PSO-enhanced deep learning [47]. Asri et al. (2023)—PSO in ensemble learning [48].	Katarya and Yadav (2018)—GA-optimized random forest [49]. Keshavarz and Abadeh (2017)—GA-based SVM [50]. Iqbal et al. (2019)—Sentiment classification using GA [51]. Li and Liang (2020)—Hybrid GA with NB [52]. Ernawati et al. (2020) – Feature tuning via GA [53]. Junior et al. (2021)—GA for multilingual sentiment [54]. Dangi et al. (2022)—GA-selected semantic features [55]. Huang et al. (2023)—Hybrid GA+CNN [56]. Ferreira et al. (2015)—Social media analysis using GA [57].
Machine translation	Tambouratzis (2009)—Phrase alignment via ACO [58]. Alsaeedan et al. (2017)—ACO for syntax trees [59]. Habibulla et al. (2023)—Hybrid ACO-RNN [60]. Zhang et al. (2023)—ACO for word sense disambiguation [61]. Shi and Tao (2024)—Multi-objective ACO [62].	Farzi and Faili (2015)—Phrase extraction using PSO [63]. Montes-Olguin et al. (2018)—Bilingual term alignment [64]. Li (2022)—PSO-enhanced encoder-decoder [65]. Zhao and Zheng (2023)—Neural translation boosted by PSO [66]. Chartcharnchai et al. (2024)—Statistical MT with PSO [67].	Echizenya et al. (1996)—Early GA-based MT [68]. Sofianopoulos et al. (2010)—GA for grammar optimization [69]. Singh et al. (2018)—Phrase table optimization [70]. Ganapathy (2020)—GA-based attention tuning [71]. Yu (2022)—Evolutionary fine-tuning in NMT [72].
Speech recognition	Nemati and Basiri (2011)—Phoneme recognition using ACO [73]. Rashno et al. (2015)—ACO for MFCC feature selection [74]. Jalili et al. (2016)—Speech pattern matching [75].	Najkar et al. (2010)—PSO for HMM optimization [76]. Chan et al. (2012)—Swarm learning of acoustic models [77]. Selvaraj et al. (2014)—PSO for noise-robust features [78]. Menditatta et al. (2016)—PSO in hybrid neural models [20]. Batista et al. (2016, 2019)—PSO in prosodic modeling [79,80]. Yogesh et al. (2017)—PSO in ANN training [81]. Mesleh and Kanaan (2008)— Dual-domain optimization [23].	Kwong and Chau (1996, 1997)—GA for speech parameterization [82,83]. Lin et al. (2000)—GA-HMM hybrid for speech [84]. Zamalloa et al. (2006)—Syllable segmentation using GA [85]. Pan et al. (2007, 2010)—Robust GA-based speech models [86,87]. Sedaaghi et al. (2007)—GA-enhanced GMM [88]. Priyadarshani et al. (2012)—Voice feature optimization [89]. Ibrahim et al. (2020)—GA for smart device voice control [90].

2.1 Genetic Algorithms

This section explores GAs, emphasizing their theoretical foundations, practical applications, and technological advancements. Reference [91] examines the theoretical underpinnings of GAs, detailing

genetic operators and their diverse applications across various domains. In [92], the authors offer a historical and contemporary review of GAs, highlighting their evolution, current trends, and potential future developments. While reference [93] underscores the application of GAs in ancestral genome inference, showcasing their effectiveness in solving complex biological problems, the study in [94] places GAs within the broader context of artificial intelligence, providing practical insights into their implementation for problem-solving. These studies illustrate the versatility of GAs and their critical role in addressing challenges across computational, biological, and artificial intelligence domains.

In recap, GAs mimic natural selection and evolution principles, operating on a population of potential solutions. They utilize genetic operators such as selection, crossover, and mutation to evolve improved solutions iteratively. GAs operate through an iterative process. The algorithm begins by initializing a population of potential solutions called chromosomes. Each individual's fitness is evaluated using a predefined fitness function. Based on their fitness, individuals are selected for reproduction. Crossover and mutation operators are then applied to generate offspring, ensuring genetic diversity. The old population is replaced with the new one, and this process is repeated until a stopping criterion, such as a maximum number of iterations or a satisfactory fitness level, is met.

Mathematically, the fitness function $f(x)$ measures the quality of a solution x. The selection probability for an individual i is calculated as:

$$P_i = \frac{f(x_i)}{\sum_{j=1}^{N} f(x_j)}, \quad i = 1, \ldots, N. \tag{1}$$

where N is the population size. Crossover combines two parents, P_1 and P_2, to produce offspring O using:

$$O = \alpha P_1 + (1 - \alpha) P_2, \quad \alpha \in [0,1]. \tag{2}$$

Mutation introduces small random changes to offspring, expressed as:

$$O' = O + \delta, \quad \delta \sim \mathcal{N}(0, \sigma^2). \tag{3}$$

The selection probability (Eq. (1)) uses Roulette Wheel Selection, where P_i reflects an individual's fitness proportion [91]. Crossover (Eq. (2)) employs single-point crossover with α blending parents, and mutation (Eq. (3)) introduces Gaussian noise $\delta \sim N(0, \sigma^2)$ to enhance diversity. This iterative process facilitates the evolution of optimal solutions over successive generations.

2.2 Particle Swarm Optimzation

The collective behavior of swarms, such as flocks of birds or schools of fish, inspires PSO [95]. It optimizes a problem by iteratively improving candidate solutions based on the movements of particles in a search space. The following are key equations.

- *Velocity Update:* Each particle updates its velocity based on its personal best position (p_i) and the global best position (g):

$$v_i(t + 1) = \omega v_i(t) + c_1 r_1 (p_i - x_i(t)) + c_2 r_2 (g - x_i(t)) \tag{4}$$

- *Position Update:* The position of the particle is updated using its velocity:

$$x_i(t + 1) = x_i(t) + v_i(t + 1) \tag{5}$$

Here, $omega$ is the inertia weight, c_1 and c_2 are acceleration coefficients, and r_1 and r_2 are random numbers in $[0, 1]$.

In natural language processing (NLP), Particle Swarm Optimization (PSO) has been applied to hyperparameter tuning in NLP models and the optimization of word embeddings, leveraging its ability to efficiently explore high-dimensional spaces. The velocity update equation (Eq. (4)) incorporates inertia (ω), personal best positions (p_i), and the global best position (g), with c_1, c_2 serving as acceleration coefficients [95]. Meanwhile, the position update equation (Eq. (5)) iteratively adjusts particle positions.

2.3 Ant Colony Optimization

Ants' foraging behavior inspires ACO. It is particularly effective for solving combinatorial optimization problems [96]. The formulations are described as follows.

- *Pheromone Update:* The pheromone trail is updated based on the quality of the solutions found:

$$\tau_{ij} \leftarrow (1 - \rho)\tau_{ij} + \Delta\tau_{ij} \tag{6}$$

where ρ is the evaporation rate and $\Delta\tau_{ij}$ is the pheromone deposited by ants.

- *Transition Probability:* The probability of moving from node i to node j is given by:

$$P_{ij} = \frac{(\tau_{ij})^\alpha (\eta_{ij})^\beta}{\sum_{k \in N_i} (\tau_{ik})^\alpha (\eta_{ik})^\beta} \tag{7}$$

where η_{ij} is the heuristic information (e.g., inverse of distance), and α and β control the influence of pheromone and heuristic information, respectively.

ACO has been applied to NLP tasks to optimize sequence alignment in MT and resource allocation in speech recognition systems. Pheromone updates (Eq. (6)) decay with rate ρ and increment via $\Delta\tau_{ij}$ based on solution quality [96]. Transition probability (Eq. (7)) weighs pheromone (τ_{ij}) and heuristic (η_{ij}) factors with exponents α and β.

3 Applications of Bio-Inspired Algorithms in NLP Tasks

3.1 Text Classification

Text classification is the automated process of categorizing natural language text into predefined semantic categories. In simpler terms, it involves assigning a specific class label, such as "Tech," to a given document from a range of possible labels like "Business," "Sports," or "Entertainment." Because manually creating classification rules is costly and inefficient, most modern methods automatically use SA techniques to learn text classifiers from examples. However, traditional SA methods rely on feature engineering, whereas bio-inspired algorithms optimize feature selection and model parameters.

Let $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$ represent the dataset, where x_i is a text instance and $y_i \in C$ is its corresponding label from the set of categories C. The objective is to find a function $f : X \to C$ that maximizes classification accuracy. By applying the bio-inspired approach, we can optimize the selection of features F by encoding them as chromosomes and evolving the population to maximize the fitness function, $f(F)$ (e.g., GA) or optimize hyperparameters of SA models like SVMs or neural networks (e.g., PSO). Particularly, the fitness function $f(F)$ is defined as follows:

$$F = \frac{1}{|D|} \sum_{i=1}^{n} \delta(f(x_i), y_i) \tag{8}$$

where δ is the Kronecker delta function, indicating whether the prediction matches the true label.

Following these concepts, many studies have addressed the text classification tasks from simple to complex. In 2007, AntMiner+ was an advanced ant-based classification technique that improved previous versions by utilizing an ant system, creating a defined environment for better exploration and handling of multiclass problems [97]. It automates the adjustment of algorithmic parameters, enhancing its efficiency compared to earlier iterations. Extensive benchmark tests show that AntMiner+ achieves superior accuracy and is competitive with renowned classifiers like C4.5, RIPPER, and SVM, making it a valuable tool in healthcare and finance.

One year later, reference [23] studied a new feature subset selection method for Arabic text classification, using ACO combined with Chi-square statistics to improve Support Vector Machine (SVM) performance. Their work addresses the growing need for practical tools to manage the increasing volume of Arabic digital information. The study highlights the importance of the document pre-processing phase, particularly Feature Subset Selection (FSS), which reduces dataset dimensionality, speeds up SVM training, and enhances classifier performance. The proposed method was tested against six other FSS techniques using a custom Arabic document corpus, showing superior precision, recall, and F1 score results. A study in [24] proposed a novel feature selection algorithm based on ACO. This algorithm is particularly significant because it does not require prior knowledge of the features, which makes it more accessible for various applications, especially with unknown or new datasets. Comparative analyses show that this new method outperforms traditional techniques, such as Information Gain and the CHI algorithm, to achieve higher accuracy with fewer features in text categorization tasks, mainly when tested on the Reuters dataset.

Nema and Sharma present a novel approach to multi-label text categorization that utilizes feature optimization through ACO and relevance clustering techniques [25]. Their method effectively addresses the challenges of feature selection for multi-label data, where each instance can belong to multiple categories. By beginning with feature extraction and employing ACO, classification accuracy is enhanced while dimensionality is reduced. Empirical testing against traditional methods, including fuzzy systems, demonstrated significant improvements in precision, recall, and F1 Score on complex datasets like Webkb and RCV1, highlighting the robustness and effectiveness of their approach in practical applications. While traditional methods like bags-of-words (BOW) and parts-of-speech (POS) fail to capture contextual nuances, the study in [98] focuses on a refined BOW approach that sorts and stems lexicon sets for more meaningful classifications. This study employs a train-test experimental model with a TF-IDF vector approach, where ants iteratively determine the best classification paths, mirroring their natural foraging behavior. The authors suggest that further refinement in feature selection could improve ACO's effectiveness in text classification.

In [26], the authors explore SA techniques for analyzing the sentiments of tweets through Naive Bayes and SVM classifiers enhanced by ACO. It focuses on differentiating subjective (opinion-based) from objective (fact-based) expressions, addressing the challenges posed by tweets' brief and informal nature. The study employs a dataset of 200,000 tweets, preprocessing the data to optimize features and improve sentiment classification into positive or negative categories. It employs unigram and bigram features while excluding emoticons to reduce bias and discusses the technical aspects of data preprocessing, including tokenization and TF/IDF. Overall, the research demonstrates a systematic approach to SA that leverages ACO for more efficient classification. AntMiner algorithms, a subset of ACO, focus on rule-based classification, which is vital in areas like medicine and finance, where interpretability is essential. The modular framework of ACO allows for methodological refinements. This review highlights real-world applications of AntMiner, showcasing its broad usability in fields prioritizing interpretative classifiers [27]. It covers key ACO strategies and components, such as rule construction, transition probability functions, and pheromone updating strategies, all crucial for the efficiency of AntMiner algorithms. The review also addresses challenges like handling continuous attributes and classifying unseen data, emphasizing the complexity of optimizing these algorithms.

The research by Gite et al. investigates the application of ACO to enhance feature selection and extraction in text categorization tasks, underscoring the importance of feature engineering for improving SA model performance [28]. The study tested ACO on numerical and text-based datasets, specifically on stroke prediction and hate speech detection. Using models such as Logistic Regression, K-Nearest Neighbors, Stochastic Gradient Descent, and Random Forest, the researchers found a notable accuracy improvement, with the Random Forest model achieving a 10.07% increase using TF-IDF feature extraction. However, they also highlighted challenges with textual data, such as higher computational costs and complexity, leading to an accuracy gap of about 18.43% compared to numerical data. The study presents a hybrid approach to SA using ACO and Artificial Neural Networks (ANN) to analyze linguistic data from social media. Given the abundance of user-generated reviews, there is a critical need for effective tools to interpret public sentiment accurately. This method involves ACO identifying optimal features from the data, which are then classified by ANN, enhancing classification accuracy. Utilizing the aclIMDB dataset, which contains a balanced mix of positive and negative reviews, the technique demonstrates its effectiveness through performance metrics like precision, recall, and overall accuracy, achieving an impressive accuracy rate of 97.5% [29].

Another study focuses on classifying drug reviews through aspect-based sentiment analysis (ABSA) using a hybrid model that integrates the RoBERTa model, a Bidirectional Long Short-Term Memory (Bi-LSTM) network, and an ACO for feature selection as shown in Fig. 1. This approach allows for a more detailed analysis of sentiments related to specific drug attributes, addressing the shortcomings of traditional SA in the healthcare context. By efficiently processing text data and identifying optimal features, the model achieved impressive results on datasets like druglib.com and drugs.com, with an accuracy of 96.78% and an F1 score of 98.29%, showcasing its effectiveness in harnessing user-generated content to support healthcare decision-making [22]. A new approach for sentiment classification in large drug review datasets using an Ant Colony-based Deep Belief Neural (AC-DBN) framework is presented in [30]. This framework combines NLP with advanced SA techniques to address the inefficiencies of traditional methods. The AC-DBN model implements an optimization algorithm to minimize errors and enhance learning by processing a comprehensive dataset that includes various sentiments. The results show marked improvements in accuracy and performance metrics such as precision and recall compared to conventional models like Weighted Word Representation and SA.
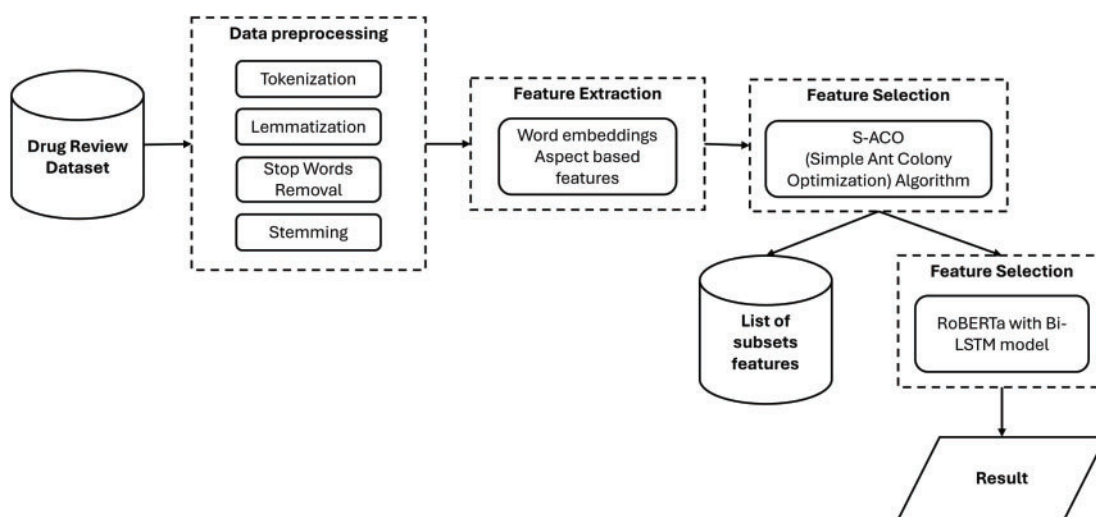
Figure 1: Workflow of ABSA for drug reviews using hybrid model

A study by Song and Park presents a GA based on Latent Semantic Indexing (GAL) to enhance text clustering, addressing the challenges posed by the high dimensionality typical of textual data. Unlike traditional GA, which relies on the inefficient Vector Space Model, GAL effectively reduces dimensionality and captures essential semantic relationships among documents [34]. This approach uses variable string lengths for chromosome encoding, allowing dynamic cluster adjustments. Their extensive tests utilizing the Reuters-21578 text collection show that GAL outperforms traditional GAs' clustering accuracy and computational efficiency, mainly when LSI significantly reduces dimensionality. In that same year, the authors identified that traditional text clustering techniques often fail to capture more profound conceptual similarities, as they rely heavily on identical terminology. To address this limitation, they developed a modified GA that combines thesaurus-based and corpus-based ontologies, utilizing WordNet's structured taxonomy and a novel transformed Latent Semantic Indexing (LSI) model [35]. This approach enhances semantic understanding, enabling the model to identify associations between documents without shared terms. Experiments on the Reuters-21578 text corpus reveal significant improvements in clustering precision, recall, and convergence compared to traditional methods like k-means. Integrating ontology-based measures with GAs represents a notable advancement in unsupervised text clustering, with potential applications in organizing large text datasets and opportunities for future refinement in specialized domains.

Uğuz and Harun introduced a two-stage feature selection and extraction model that uses information gain (IG), principal component analysis (PCA), and GA to improve classification performance. In the first stage, terms are ranked by relevance through IG, followed by PCA and GA to reduce dimensionality among these prioritized terms. The efficacy of this model is evaluated using two datasets, Reuters-21,578 and Classic3, with the k-nearest neighbor (KNN) and C4.5 decision tree classifiers. Results indicate a significant enhancement in classification metrics such as precision, recall, and F-measure compared to using all features [37]. Additionally, the study highlights preprocessing techniques like stopword removal and stemming and emphasizes that the hybridized approach reduces computational complexity while improving overall classifier performance.

A study developed an automatic web page classification system using a GA (GA) that combined HTML tags and terms as features. It demonstrated a 95% accuracy, particularly when training sets had more negative documents. The GA-based classifier outperformed Naïve Bayes and KNN, showing adaptability to large feature sets for comprehensive tasks. The research concluded that using tagged terms effectively manages diverse web document features, with suggestions for future work on varying document class proportions and applying the method to larger datasets [36]. Another study in [38], a new method that enhances text categorization by combining chaos optimization and GAs. CGFSO improves classification accuracy while reducing the feature set and iteratively assessing subsets based on precision and recall. It outperforms traditional methods of GA and SVM in the Reuters document set, effectively minimizing computational complexity. Future work could explore additional hybrid optimization techniques for broader data categorization applications.

On the other hand, the method in [39] enhances document representation through a two-stage process involving feature selection using advanced filter-based techniques, followed by a transformation stage that modifies Latent Semantic Indexing (LSI) with a GA. By utilizing singular vectors linked to large and small singular values, GALSF captures more discriminative elements than traditional LSI methods, focusing solely on the largest values. Experiments on benchmark datasets demonstrate that GALSF consistently outperforms standalone LSI and traditional feature selection methods, effectively identifying underlying patterns and variations in the text. A recent study explored a GA approach for selecting optimal feature subsets to improve classifier performance while minimizing subset sizes. This method was evaluated using three classifiers—Naive Bayes, KNN, and SVM—on two datasets, the 20 Newsgroups and Reuters-21578 [40].

The results showed significant classification accuracy and efficiency enhancements when incorporating GA-based feature selection. Various text representation methods, such as bag-of-words and N-grams, were also analyzed for their effect on classifier performance, revealing substantial improvements across different datasets and classifiers. Additionally, the study provided an overview of GA concepts relevant to feature selection, including selection, crossover, and mutation processes, highlighting their role in achieving more accurate classification outcomes.

Another study presents hybrid feature selection approaches using an Enhanced GA (EGA) for text categorization, focusing on high-dimensional feature spaces to improve performance. By refining traditional crossover and mutation functions of GAs based on feature importance and classifier performance, EGA targets feature selection more effectively than random methods. The research combines EGA with established filter feature selection techniques to enhance dimensionality reduction and classifier efficiency. Thorough testing on three Arabic text datasets reveals significant improvements in dimensionality reduction, execution time, and categorization performance compared to traditional methods [41].

Reference [21] discusses a hybrid method that combines Information Gain and PSO for feature selection in SVM-based SA. This approach addresses the challenges posed by the informal language and high attribute-to-document-length ratio found in social media data, such as tweets. By refining attribute selection through PSO and Information Gain before classification with SVM, the researchers achieved an accuracy of 94.80% and an Area Under the Curve (AUC) of 0.98 in experiments focused on the West Java Governor election sentiment, demonstrating significant improvements in handling high-dimensional data typical in NLP tasks.

The study in [31] highlights the main role of SA in businesses for extracting insights from online customer feedback, enabling a better understanding of customer preferences and brand perception to inform data-driven decisions. Through this methodology, PSO effectively processes significant features in unstructured data, thus improving operational efficiency and accuracy. The proposed model outperformed various algorithms, including GA and ACO, in feature handling and execution time. A research initiative by Mandal et al. explores the integration of PSO with SA for text summarization, focusing primarily on extractive methods. The authors highlight the limitations of past studies in summarization techniques and propose a novel model that efficiently selects meaningful text clusters by evaluating each particle's sentiment score derived from SentiWordNet. Their methodology combines PSO with GA, ensuring the generated summaries are comprehensive, non-redundant, and concise by adhering to constraints like coverage and diversity. Testing on various datasets revealed significant recall and F-measure metrics improvements under ROUGE standards, marking advancements in automated text summarization research [32].

Yue et al. introduce an advanced SA model that combines BiLSTM, TextCNN, and Self-Attention mechanisms and is optimized through an Improved PSO (IPSO) method [47]. This model effectively tackles short text SA by incorporating a Generative Adversarial Network (GAN) to generate modified texts, enhancing its robustness against data irregularities. The IPSO technique, inspired by bird foraging behavior, improves hyperparameter optimization by dynamically adjusting learning factors and inertia weights, leading to superior performance across four datasets compared to traditional methods. The study details a comprehensive methodology from noise injection via GANs to meticulous feature extraction and performance testing, showcasing the successful integration of advanced neural networks with effective optimization strategies.

3.2 Sentiment Analysis

SA aims to determine textual data's sentiment polarity (positive, negative, or neutral). The problem formulation is described as follow: "Given a text $T = \{w_1, w_2, \ldots, w_m\}$, the objective is to determine the sentiment label $s \in \{Positive, Negative, Neutral\}$". In a sentiment classification task, ACO optimized word

embedding dimensions by balancing accuracy and computational efficiency, achieving improved results compared to grid search methods. It means using bio-inspired methods to select optimal feature sets for SA.

To follow up on the problems of SA, the researchers explore how sentiment-based lexicons or embeddings can be optimized using algorithms like ACO and GA. In 2014, Kwon et al. examined SA, particularly using tools such as SentiWordNet, which helps to understand emotions in different languages [42]. They also highlighted using ACO algorithms, which improve problem-solving, particularly in network routing and data management. Ahmad et al. present a novel text feature selection method that combines ACO with the k-nearest neighbor (KNN) algorithm to improve sentiment classification accuracy in high-dimensional feature spaces. ACO effectively guides the feature selection process by efficiently navigating the complex feature space, while KNN evaluates the selected subsets based on mean squared error performance. The study creates features from customer review data and processes them through the ACO-KNN algorithm, resulting in a relevant and minimized set of features for classification. The ACO-KNN approach outperforms other hybrid methods, such as GA combined with information gain (IG-GA) and IG with rough set attribute reduction (IG-RSAR), particularly in tasks involving customer reviews of electronic products, as evidenced by improved precision, recall, and F-score metrics [43].

Hamdi introduces a novel algorithm called Affirmative ACO-Based SVM (AACOSVM) aimed at enhancing the sentiment classification of customer reviews from Amazon [44]. As SA is essential for businesses to gauge customer perceptions in the digital age, traditional algorithms often struggle with accuracy and performance on large datasets. Hamdi's AACOSVM leverages the natural foraging behavior of ants to optimize the selection of SVM parameters, employing key components of state transition and state updates to minimize classification errors. Results indicate that AACOSVM significantly outperforms traditional classifiers like EBC and EFAN regarding classification accuracy. Reference [45] aims to enhance SA by implementing an ACO algorithm for feature selection in social media data. Recognizing the challenges posed by social media text's complex and high-dimensional nature, the researchers advocate for ACO's collective intelligence approach, which mimics ant foraging behavior to efficiently evaluate and select relevant features. This innovative method addresses the limitations of traditional feature selection techniques, leading to significant improvements in classification metrics such as accuracy, precision, and recall. The findings highlight ACO's effectiveness in refining SA by selecting more predictive features, ultimately improving overall performance.

In [46], the author presents an effective method for extracting sentiments from online reviews, leveraging PSO to optimize feature selection for aspect term extraction and sentiment classification using a Conditional Random Field (CRF) model. Tested on the SemEval-2014 dataset in the restaurant and laptop review domains, the approach achieved significant reductions in feature sets—from 67 to 41 and from 83 to 35—while maintaining or enhancing classification accuracy. This demonstrates PSO's capability to reduce computational complexity and improve performance compared to previous state-of-the-art systems. In 2020, a study explores advancements in SA, mainly through automating lexicon labeling using PSO and Bare-bones PSO (BBPSO) variation. By comparing automated labels to those produced by humans, the experiments found that both PSO variants outperformed human labeling in classifying sentiments, with BBPSO being the most effective [33]. To address the shortcomings of lexicon-based methods in cases where text lacks relevant terms, a hybrid model incorporating a Naive Bayes classifier was introduced, which improved sentiment classification by considering unclassified posts. Additionally, the paper highlighted topic identification to enhance sentiment classification and examined strategies for managing intensifiers and negations to refine sentiment assessment based on contextual cues. By eliminating irrelevant or redundant features, feature selection simplifies models, reduces computational costs, and enhances performance, addressing the challenges of high dimensionality.

The study in [48] uses PSO—an optimization algorithm that mimics swarming behavior—to identify key features, demonstrating superior performance compared to GA and ACO across various datasets. Experiments show that PSO achieves higher precision, recall, F-score, and accuracy, attributed to its ability to efficiently focus on essential features for building robust SA models. GAs are increasingly important in SA, particularly for enhancing business operations through social media. As a core part of NLP, SA extracts opinions from text data, which aids in refining marketing strategies, product features, and customer service. Conducted at document, sentence, and aspect levels, aspect-level analysis provides detailed feedback by focusing on specific product features. GA, which mimics natural selection to optimize solutions over generations, demonstrates effectiveness in various SA tasks, as highlighted in a structured comparative analysis. This showcases their adaptability and ability to integrate with other computational methods for improved accuracy [49]. In [50], the authors introduced ALGA (Adaptive Lexicon Learning using GA), a novel Twitter SA approach that addresses the platform's dynamic language and slang. By combining corpus-based and lexicon-based techniques, ALGA creates adaptive lexicons for classifying sentiments as positive or negative, optimizing the identification of effective lexicons. It incorporates all words, including stop words, to understand tweet sentiment context better. Tests on six diverse datasets revealed that ALGA achieves over 80% accuracy and F-measure, demonstrating its adaptability to cultural changes and superior performance compared to existing methods.

In addition, Farkhund et al. introduce a hybrid SA framework that integrates a GA-based feature reduction technique to improve scalability and accuracy in handling large datasets [51]. By reducing the feature set size by up to 42% without compromising accuracy, this framework effectively addresses the limitations of both lexicon-based and machine-learning methods. Evaluations using precision, recall, and F-measure metrics demonstrated the framework's robustness across various datasets, including online reviews, tweets, and geopolitical data, showing significant improvements in precision compared to traditional dimensionality reduction methods like PCA and LSA. The GA-enhanced model successfully managed larger data sets while maintaining high accuracy. A model developed by researchers from Shandong University enhances word embeddings for SA by integrating an improved GA and a sentiment lexicon [52]. While traditional word embeddings effectively capture semantic relationships, they often overlook crucial sentiment information, leading to misclassifications. The proposed refinement model addresses this issue by optimizing word vectors to reflect sentiment context better, as demonstrated through experimental analysis on datasets like IMDB and the Stanford Sentiment Treebank. The results indicate that the refined embeddings significantly improve sentiment classification accuracy when used with various neural network architectures compared to existing methods.

Ernawati et al. aimed to analyze and classify user reviews of digital learning applications into positive and negative sentiments, which are crucial for assessing trust in these platforms. The study used two feature selection methods, GA and PSO, in conjunction with the Naive Bayes classification algorithm [53]. Using a dataset of two hundred reviews evenly split between sentiments, the authors optimized various parameters to enhance classification accuracy. The results demonstrated that the Naïve Bayes algorithm combined with PSO outperformed the GA method, achieving an accuracy of 98.00% compared to 95.50%, while also being more efficient in feature reduction. SA has gained importance with the rise of social media, leading to research in polarity detection—classifying text as positive, negative, or neutral. The complexities of language, such as sarcasm and slang, make traditional classifier development resource-intensive. A genetic programming approach has been proposed to automate the selection and combination of essential lexicons and textual features, demonstrating competitive performance compared to established SA methods across various datasets, including sarcastic text. GP's ability to adjust lexicon importance and manage neutral sentiment thresholds highlights its effectiveness. The research concludes that GP not only streamlines the SA

process but also holds potential for future enhancements, including integrating more lexicons and expanded text analysis capabilities [54].

Dangi et al. focus on enhancing SA in social media by combining Convolutional Neural Networks (CNNs) with a GA. As manual analysis of the vast data available on social media is impractical, the study aims to automate sentiment classification based on facial expressions, categorizing them as "Happy" or "Sad" [55]. Using a facial expression dataset from Kaggle, the researchers implemented various preprocessing steps and adjusted the CNN architecture. Key improvements involved mutating the GA to optimize hyperparameters, leading to a notable classification accuracy of 96.984%. One study in [56] explores an advanced approach to predicting stock market fluctuations by integrating social media SA, GA, and DL techniques, specifically the LSTM model as illustrated in Fig. 2. It aims to improve stock price prediction accuracy by incorporating online comments and traditional quantitative indicators. The study uses a hybrid GA to identify relevant technical indicators and applies the Taguchi method to optimize hyperparameters for the LSTM model. Additionally, Grey Relational Analysis is employed to determine which social media sentiment variables are most correlated with stock prices. The findings confirm that including social media sentiment significantly enhances the predictive reliability of the LSTM models. The study explores the challenges of using SVMs for SA in imbalanced datasets, where certain emotions are overrepresented while others are underrepresented [57]. It introduces a GA that intelligently oversamples minority and undersamples majority classes to create a more balanced dataset, thereby improving classifier performance. Tested on Brazilian Portuguese texts, the results revealed significant gains in accuracy and F1 scores when identifying six emotions, highlighting that data balancing is especially crucial in multiclass scenarios. The paper concludes with suggestions for future research on applying this approach to various datasets and exploring different classification algorithms for analyzing imbalanced sentiment.
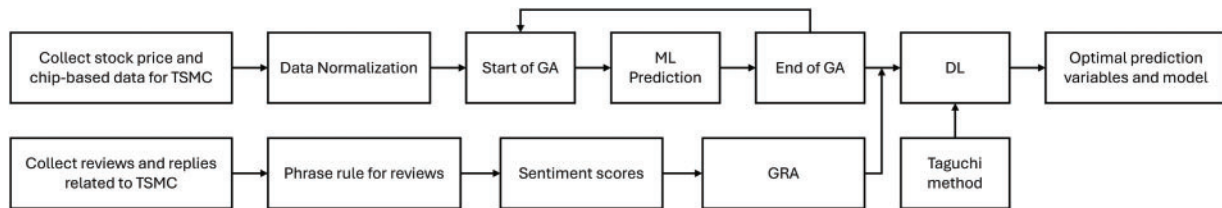
Figure 2: Framework for predicting stock market fluctuations

3.3 Machine Translation

MT involves converting text from one language to another. The challenge lies in handling contextual meaning and linguistic nuances. Given a sentence $S = \{w_1, w_2, \ldots, w_m\}$ in the source language, the aim is to generate a translated sentence $T = \{v_1, v_2, \ldots, v_n\}$ in the target language. To solve this problem formulation with the bio-inspired approach, we first define the equation:

$$P(T|S) = \prod_{i=1}^{n} P(v_i|v_{<i}, S) \tag{9}$$

Here, $P(v_i|v_{<i}, S)$ is the probability of translating the ith word based on the context and the source sentence. Then, ACO was applied to align phrases in a statistical MT model, improving BLEU scores on low-resource language pairs. Generally, the application of ACO and GA for sequence alignment and translation phrase optimization. Integrating bio-inspired algorithms with neural MT to optimize parameters, handle large search spaces, and improve translation accuracy. The application of bio-inspired in this is

divided into two tasks: i) optimize the alignment of source and target phrases by treating alignment as a path-finding problem and ii) enhance neural machine translation (NMT) by optimizing learning rates and architecture parameters.

In 2009, Tambouratzis's research applied Ant Colony Optimization (ACO) to enhance the automated segmentation of Ancient Greek words into stems and endings, addressing the complexity of stemming in this highly inflectional language. Traditionally reliant on manual intervention by expert linguists, the study modified the Ant System—originally designed for optimization problems like the traveling salesman problem—to fit the needs of word segmentation. Fig. 3 illustrates the processing steps involved in the automated morphological processor (AMP) variant for handling Ancient Greek. Experimental results indicated that ACO effectively tuned system parameters, often outperforming manually optimized configurations and demonstrating its robustness in improving segmentation quality. This advancement reduces the need for expert input, streamlines the analysis of historical languages, and opens up opportunities for scaling and adapting the technology to other complex languages [58]. A new approach combining hybrid genetic and ant colony optimization algorithms has been developed to address the Word Sense Disambiguation (WSD) problem. Reference [59] introduces two algorithms, MMASWSD and ACSWSD, which enhance WSD's contextual handling by incorporating semantic relatedness through extended glosses. A self-adaptive genetic algorithm (SAGA) is integrated to adjust the parameters of these algorithms dynamically. Empirical tests showed that the GMMASWSD variant significantly outperformed its counterpart GACSWSD and other existing algorithms on fine-grained corpora. However, it fell short of top performance on coarse-grained data, indicating potential for further optimization. These hybrid algorithms demonstrate significant improvements in WSD tasks for fine-grained data. However, further refinement is needed for coarse-grained scenarios, with future research suggesting exploring additional integrations and adaptations to broaden their applicability.
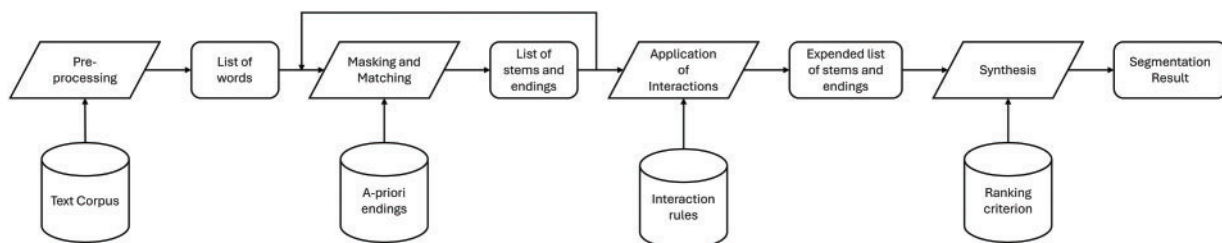
Figure 3: AMP system for ancient Greek

Habibulla et al. introduce a novel English-Chinese machine translation approach utilizing the Ant Colony Optimization (ACO) algorithm, showcased at the 2023 NMITCON [60]. The system improves speech recognition capabilities by integrating advanced neural network enhancements and addresses translation accuracy gaps. Designed to be mobile-friendly, it functions effectively on devices like tablets and smartphones while achieving a remarkable 95.7% accuracy rate across various speech scenarios. The findings, supported by rigorous evaluations against other advanced systems, signify a substantial advancement in translation technology, inviting further exploration to enhance its application on more complex datasets and expand its algorithmic integrations. This research aims to bridge educational and communicative gaps caused by language barriers, fostering a more linguistically inclusive global community.

Recent research by Zhang et al. has advanced Japanese-Chinese machine translation by integrating enhanced ant colony algorithms with grey entropy for precise word segmentation, effectively addressing

the complexities of Chinese text processing [61]. The study presents a multi-strategy translation architecture incorporating advanced computational techniques such as lexical, syntactic, and semantic analyses, alongside a semantic classification system that improves verb categorization and translation accuracy. Using multiple translation engines, the research demonstrates enhanced translation output compared to single-engine capabilities. This innovative approach showcases the blend of computational intelligence and linguistic strategies, laying a robust foundation for future advancements in machine translation and promising greater efficiency and accuracy in cross-language communication. The newly designed English machine translation system launched in 2024 utilizes innovative techniques, including the ant colony algorithm, to enhance translation quality and reliability [62]. The system efficiently tackles complex translation tasks with integrated hardware, such as the EMMY simulator and the FXI32 static translator, while achieving higher BLEU scores that reflect closer alignment with human translation standards. Its advanced software framework supports effective feature extraction and model building, contributing to improved natural language processing. Overall, this system represents a significant advancement in machine-assisted translation, providing solutions for quality, efficiency, cost reduction, and greater accessibility to high-quality translation services.

Early in 1996, Echizen-ya et al. developed an MT method that combines inductive learning and GAs to enhance the accuracy of English to Japanese translations while requiring fewer data inputs [68]. This innovative approach generates and refines translation rules through an iterative feedback system involving user input and validation. By mimicking natural selection, GAs continuously evolve these rules for improved effectiveness, introducing variability through techniques like crossover and mutation. Experiments have shown that this method significantly enhances translation accuracy compared to traditional rule-based and example-based systems, paving the way for more adaptive language processing tools that can learn similarly to human language acquisition. A study in [69] investigates the optimization of parameters in a hybrid MT (MT) system that relies on monolingual corpora rather than bilingual text corpora. The research uses GAs, specifically the SPEA2 algorithm, to enhance translation quality by optimizing parameters based on MT evaluation metrics like BLEU and METEOR. The findings reveal significant improvements in translation output compared to a baseline system with manually defined parameters. Multiple experiments demonstrated the importance of diverse training sets and refined databases for higher translation quality. Overall, the research highlights the effectiveness of evolutionary computation techniques for optimizing hybrid MT systems, particularly in scenarios where parallel corpora are unavailable, suggesting a promising direction for future automated tuning processes [69].

Reference [71] focuses on using GAs to optimize hyperparameters in neural networks for MT. This study addresses the challenges of fine-tuning hyperparameters as traditional methods like grid search become impractical with the increasing complexity of neural networks. Inspired by Darwin's theory of natural selection, Ganapathy's approach evaluates various network configurations based on their translation accuracy, measured by the BLEU score. The GA was tested against a baseline random search method in several trials, demonstrating that it consistently required fewer adjustments to reach or exceed the target BLEU score, thus validating its efficiency in hyperparameter optimization. In 2022, Yuxiu Yu introduced the GA-MLP-NN model, which enhances Chinese-English translation quality by integrating multilayer perceptron methods, GAs, and advanced neural network strategies. This sophisticated system effectively manages and refines a parallel corpus, demonstrating superior performance in identifying high-quality translation data compared to traditional systems [72]. The research achieves significant advancements in automated translation processing by employing innovative filtering methods and a uniquely designed neural network. Although the model outperforms single-layered approaches, the study emphasizes the necessity for ongoing refinement, particularly for domain-specific translations, to enhance adaptability and precision

in diverse text types. Singh et al. introduced a GA-based MT System (GA-MTS) specifically for translating Sanskrit to Hindi, addressing the linguistic complexities existing translation systems often struggle with, such as semantics, word order, and idiomatic expressions [70]. The GA-MTS utilizes advanced techniques like tokenization, morphological analysis, and parsing to maintain linguistic integrity, improving translation quality. The system ensures adherence to both languages' grammar, syntax, and semantics by employing a GA for efficient translation mapping. The initial results demonstrate high accuracy in sample translations, indicating the system's potential as a valuable tool for overcoming language barriers in India's diverse cultural landscape. Future work will aim to broaden the corpus and include complex sentence structures to enhance reliability and effectiveness.

Reference [63] investigates a re-ranking system for Statistical Machine Translation (SMT) that employs a Quantum-behaved PSO (QPSO) algorithm to improve translation quality by addressing issues of ungrammatical and non-fluent outputs. The QPSO-enhanced re-ranker utilizes advanced non-syntactical features from language and translation models, categorized into language model features, translation model features, N-best list features, and others like anchor matcher and Part-of-Speech (POS) tag features. Through extensive testing on two different language pairs and various genres, the study demonstrates the QPSO algorithm's superiority over traditional methods in terms of translation accuracy, measured by BLEU scores and computational efficiency. The researchers also analyze the contributions of different feature sets, providing valuable insights into their impact on translation quality across diverse datasets. In 2018, a study explored PSO in MT, specifically on English and Spanish translations, which proposed a novel approach that combines evolutionary algorithms with cosine similarity to enhance translation efficiency and accuracy [64]. By adapting PSO to optimize its velocity attribute, the methodology simplifies complex cost functions and reduces translation ambiguity by measuring semantic similarities between source and target texts. This approach prioritizes sentence-level translations over intricate word-based methods, aiming to streamline the translation process and eliminate the need for extensive linguistic analysis. Another study [65] develops a computer-aided English-Chinese translation system utilizing a PSO algorithm as its core strategy, enhancing the efficiency and effectiveness of the translation process. This system addresses the growing demand for efficient translation methods that cannot be met by humans or basic machines alone. While Computer-Assisted Translation (CAT) tools assist translators by reducing redundancy and improving memory capabilities, challenges remain regarding file format adaptability and user experience. The study compares PSO to other evolutionary algorithms, emphasizing its ease of application and cost-effectiveness advantages, particularly for complex optimization problems. Overall, the research envisions a future where automated translation systems, guided by human oversight, enable translators to focus on intricate translation elements rather than procedural tasks.

Besides, Zhao and Zheng, in the study [66], focus on improving machine English translation through an enhanced method utilizing Improved PSO. It addresses common challenges in traditional translation, such as incorrect word choices and context interpretation issues. By integrating PSO with a neural MT model, the study aims to improve translation accuracy using a probabilistic fuzzy search approach, which enhances semantic retrieval and selection. Experiments conducted in Matlab demonstrate significant improvements in translation accuracy and reliability, effectively reducing errors and better managing semantic distributions. With the same purpose, in [67], the authors introduce a novel hyperparameter optimization method called set-based Categorical PSO (C-PSO) for low-resource Transformer-based NMT models. This approach diverges from traditional hyperparameter tuning by using particles as probability distributions, enhancing translation quality on German and Thai datasets. The vanilla Transformer architecture's performance heavily relies on hyperparameters, and conventional methods like grid and random search often prove inadequate in complex spaces. The C-PSO demonstrates better search performance and convergence, effectively improving

BLEU scores while closing the performance gap between high-resource and low-resource datasets. This indicates that high-quality translations can be achieved even with limited data.

In 2024, Macedo et al. explored the use of neuroevolution to enhance user interfaces and improve human-machine interaction. Using ANN optimized through evolutionary algorithms, the study demonstrated the potential to create interfaces that translate user input into precise commands, addressing challenges associated with traditional rigid HCI systems [99]. The authors compared neuroevolutionary methods with conventional backpropagation techniques, finding that neuroevolution excelled at handling complex, high-dimensional data, particularly in scenarios involving camera input for head movements and facial expressions. This approach promises to create more intuitive and adaptable interfaces, including improving accessibility for users with disabilities and advancing technology integration into daily life.

3.4 Speech Recognition

Speech recognition translates spoken language into text, requiring optimization of acoustic and language models. The first attempts at speech recognition aimed to develop machines that mimic human speech rather than recognize it. This was due to the lack of knowledge about the acoustic resonance tubes that approximate the human vocal tract. The formulation of this problem is stated as follows: given an audio signal A, decode it into the most probable word sequence W^*:

$$W^* = \arg\max_W P(W|A) \tag{10}$$

where $P(W|A)$ is:

$$P(W|A) \propto P(A|W)P(W) \tag{11}$$

where $P(A|W)$ is the acoustic model, and $P(W)$ is the language model. Following the bio-inspired approach, we first optimize the parameters of Hidden Markov Models (HMMs) or DL models by applying PSO. Then, select features from Mel Frequency Cepstral Coefficients (MFCCs) for improved recognition accuracy with GA.

In 2011, Nemati et al. explored the enhancement of automatic speaker verification (ASV) systems through ACO for feature selection, focusing on a text-independent approach without constraints on speech samples. The research aims to reduce system complexity and verification time by minimizing the feature space's dimensionality, eliminating redundant features while maintaining predictive accuracy [73]. A comparative analysis with GA demonstrated that ACO outperformed GA by significantly decreasing the number of features and accelerating the verification process. By integrating the optimized features into a Gaussian Mixture Model Universal Background Model (GMM-UBM), the study achieved over 80% reduction in feature dimensions, resulting in improved ASV performance, as measured by equal error rate (EER) and detection cost function (DCF), with ACO showing better accuracy and lower false recognition rates. In a notable study, Abdolreza Rashno and his team presented a new method to improve the accuracy and efficiency of automatic speaker verification (ASV) systems. They addressed the challenges of high-dimensional feature vectors, often including irrelevant data, by introducing a feature selection approach that combines ACO and SVM with feature relief weights [74]. This method achieved a 64% reduction in feature dimensionality and a low Equal Error Rate (EER) of 1.745% using the polynomial kernel of SVM. Compared to traditional GAs, their approach demonstrated superior performance with consistently lower EER and fewer feature selections. At the same time, the ACO framework helped accelerate the optimization process by pre-determining feature weights. The study in [75] presents advanced speech recognition methods

that enhance signal clarity and reduce noise. It integrates a fuzzy logic system with an ACO algorithm to classify and process speech signals efficiently. Initially, the fuzzy system minimizes data dimensionality, while the ACO algorithm organizes signals into clusters, separating noise for removal. The study details how the ACO mimics ant behavior in optimizing paths and clustering based on pheromone levels. The fuzzy model addresses variability in speech signals using Gaussian functions. Results indicate that the combined approach outperforms existing techniques, achieving higher recognition rates, improved noise management, and greater computational efficiency.

To enhance speech recognition systems using HMMs, the author in [83] implemented Parallel GAs (PGA). Traditional optimization methods for HMM parameters often led to local maxima, but the introduction of GA allows for global searches, improving the training process. By parallelizing computations through an island model, where multiple GAs work across different nodes of a multiprocessor system, the study achieved an 18% increase in recognition rates compared to non-parallel methods. The paper details the encoding mechanisms and selection processes critical to GA performance, demonstrating that PGA yields faster and higher-quality solutions in speech recognition than conventional methods. Another work, Lin et al., addresses the challenges of speech recognition in noisy environments, highlighting how complex speech waveforms complicate detection [84]. The Two-Dimensional Cepstrum (TDC) method struggles with background noise; thus, they introduce the Modified TDC (M_TDC), which uses a temporal filter to reduce interference.

Additionally, GAs optimize the selection of robust TDC coefficients from the M_TDC matrix, improving adaptability to noise. Experimental results show that the GA-based M_TDC outperforms traditional TDC and various HMMs in noisy settings while maintaining efficiency in clean conditions, making it a promising approach for practical applications. Focusing on HMM and ANN, reference [82] explores effective architecture for chip-embedded speech recognition. It critiques traditional gradient descent training for ANNs, which can lead to local optima, and introduces a GA that enhances performance by identifying solutions closer to the global optimum. Key pre-processing stages such as speech sampling and feature extraction using Mel-Frequency Cepstral Coefficients (MFCC) are thoroughly examined, highlighting their impact on recognition rates. Results indicate that GA-trained ANNs outperform those trained with gradient descent in terms of speed and accuracy on FPGA chips. The paper also emphasizes using Integer Fast Fourier Transform (FFT) to reduce time and power consumption. Ultimately, it underscores GAs' potential to improve speech recognition systems' efficiency for real-world applications in embedded consumer electronics. Lan et al. examine the effectiveness of using GA to train Backpropagation Neural Networks (BPNN) for improved speech recognition efficiency. It begins by reviewing current speech recognition methods, including Dynamic Time Warping (DTW) and HHMs, emphasizing the advantages of ANN for lower-power computing environments [100]. The study aims to address the local optimum problem associated with traditional Sequential Decision Making (SDM) by incorporating GA, which enhances the exploration of optimal weight configurations. Through experiments, the authors demonstrated that integrating GA with BPNN training led to increased recognition rates. The paper also highlights the importance of preprocessing techniques, such as point detection and Mel-Frequency Cepstral Coefficients (MFCC) extraction, in enhancing overall performance in speech recognition tasks.

Reference [85] aims to enhance speaker recognition systems by using Mel-Frequency Cepstral Coefficients (MFCC) and GAs to optimize feature weighting and selection. It identifies and assigns weights to a comprehensive 38-dimensional feature set, significantly reducing error rates by up to 25%. By selecting the most valuable features, the approach creates a more streamlined model that minimizes dimensionality without sacrificing accuracy, enhancing processing efficiency on low-resource devices. The research demonstrates that combining feature weighting with selection effectively optimizes speaker recognition

performance, with future work suggested to refine these techniques further for more constrained systems. In other studies, Pan et al. investigate maximizing the training of ANN for Mandarin digit speech recognition by using GAs instead of traditional training methods, particularly highlighting the shortcomings of the Steepest Descent Method (SDM), which struggles with local optima [86]. The proposed GA enhances evolutionary steps through strategic chromosome crossover, facilitating the discovery of global solutions. Utilizing a Back-Propagation Neural Network (BPNN) with a multi-layer perceptron architecture, the study emphasizes the importance of preprocessing tasks like point detection and feature extraction via Mel-Frequency Cepstral Coefficients (MFCC) to ensure high-quality input data. Training results indicate that the improved GA significantly boosts learning performance, achieving speech recognition accuracy of up to 95%. They continue introducing an Enhanced GA to train ANNs to improve non-specific speaker Mandarin digit speech recognition [87]. The study highlights the limitations of traditional training methods, particularly the Steepest Descent Method (SDM). It outlines a comprehensive speech recognition process involving pre-processing steps like point detection and feature extraction using Mel-frequency Cepstral Coefficients (MFCC). The Improved GA is specifically designed to overcome the shortcomings of SDM by conducting a more thorough search for optimal weight configurations. Empirical results show that the Improved GA significantly outperforms both the Traditional GA and SDM, leading to higher speech recognition accuracy.

Sedaaghi and colleagues, in [88], present an approach that utilizes adaptive GAs to enhance speech-emotion recognition systems. By applying discrete emotional theories and leveraging SA, they aim to filter out ineffective features and utterances through a two-stage process guided by a Bayes classifier. This adaptive GA adjusts mutation and crossover probabilities to maintain diversity within the population, preventing stagnation in local optima and allowing for more effective solution exploration. Their application of these methodologies to the Danish Emotional Speech database showed improved accuracy in emotion recognition compared to traditional methods. The study highlights the potential of adaptive GAs to optimize feature selection and system performance, paving the way for future advancements in handling complex multi-feature datasets and enhancing human-computer interactions. In [89], the authors aimed to alleviate the need for user-provided speech samples while achieving high accuracy in both speaker-dependent and independent contexts. It explored three approaches to Automatic Speech Recognition (ASR)–acoustic phonetic, pattern recognition, and artificial intelligence—emphasizing GA's role in optimizing recognition processes through crossover, mutation, and selection. The results showed a nearly 96% word recognition rate, highlighting GA's effectiveness in adapting to various speakers. With the same purpose, reference [90] presents a study focused on enhancing the efficiency and accuracy of Arabic phoneme recognition systems. Utilizing a GA combined with feedforward neural networks, the study optimizes distinctive phonetic features through significant feature size reduction, achieving a decrease of approximately 50% in vector dimensions while maintaining 90% recognition accuracy. The researchers utilized the King Abdulaziz City for Science and Technology Arabic Phonetic Database, applying various configurations of GAs and extensive preprocessing of features like spectrograms and MFCCs. The results, validated using Wilcoxon signed-rank tests, indicate a robust approach that streamlines processing efforts and improves system scalability without compromising performance. This work marks a pivotal step toward developing more efficient and accurate speech recognition systems while paving the way for future research.

Najkar et al. present an innovative approach to enhancing HMM-based speech recognition systems through PSO, replacing the traditional Viterbi algorithm. This method involves the generation and iterative optimization of segmentation vectors, utilizing both standard and global strategies to improve likelihood measures. Experiments on isolated word recognition and phone classification demonstrate that the PSO approach retains accuracy while reducing computational complexity. The advantages of PSO over GAs and the Baum-Welch algorithm are highlighted, particularly its superior global search capabilities and

efficient convergence. This methodology shows promise for future applications in continuous speech recognition, especially in scenarios with limited computational resources [76]. Speech control mechanisms are increasingly used in industrial operations to enhance efficiency, but they often struggle with performance degradation due to environmental noise. A recent paper presents an innovative approach using Intelligent PSO (IPSO) to optimize beamformers that improve speech recognition systems in commercial settings [77]. Traditional methods can distort signals during noise filtration, while conventional optimization algorithms fail to address the discontinuous nature of recognition accuracy landscapes. IPSO effectively manages these discontinuities by adjusting velocities and maintaining particle diversity, resulting in superior recognition rates compared to other algorithms and traditional methods. This makes IPSO a promising tool for enhancing speech recognition in noisy environments and advancing industrial automation systems.

In [78], the authors present an innovative speech recognition approach by integrating Improved PSO (IPSO) with HMM. The methodology involves denoising speech signals with a median filter and extracting pitch, peak spectrum, and Mel-Frequency Cepstral Coefficients (MFCC) features. A key aspect of the research is using a GA-based vector quantization technique that optimally clusters these features to enhance the training set, resulting in a highly accurate speech recognition system with a success rate of 97.14%. This approach effectively addresses common challenges in speech recognition, such as signal noise and feature extraction limitations, positioning it as a superior alternative to traditional methods in acoustic signal processing. Reference [20] enhanced ASR systems through a hybrid algorithm that combines Artificial Bee Colony (ABC) and PSO. This approach focuses on optimal feature selection from speech data to train an SVM classifier, significantly increasing recognition accuracy and performance. The ASR system involves three key stages: preprocessing with a Wiener filter to reduce noise, extracting eight statistical and acoustic features, and utilizing the ABC-PSO algorithm for feature selection. Testing on MATLAB with recorded speech datasets reveals that this method drastically improves recognition accuracy compared to systems without feature selection. The study underscores the potential for hybrid algorithms to enhance ASR systems, making them applicable for real-world scenarios requiring reliable speech recognition. Another research explores the integration of SVM and PSO to enhance ASR for recognizing digits in Brazilian Portuguese. SVM generates an optimal hyperplane for efficient pattern categorization, while PSO mimics natural social behaviors to optimize data representation. The study employs preprocessing techniques, such as mel-cepstral coefficients and Discrete Cosine Transform (DCT), to convert speech signals into a two-dimensional matrix for input into the PSO algorithm before training the SVM classifier. The findings indicate that this hybrid approach significantly improves processing times and classification accuracy compared to SVM alone, with kernel functions like Radial Basis Function (RBF) enhancing performance. This combination represents a notable advancement in ASR, offering the potential for more efficient systems capable of handling diverse languages and dialects [79].

In [81], a novel tool called PSOBBO combines PSO and Biogeography-Based Optimization (BBO) to improve feature selection efficiency in emotion recognition systems. Fig. 4 illustrates the speech emotion recognition pipeline from the study. This tool is significant for enhancing human-machine interactions in various fields, such as health diagnostics and customer service, particularly addressing the challenges posed by the variability of emotional expression among speakers. The algorithm was rigorously tested against established databases, demonstrating superior performance over standalone PSO and BBO regarding classification accuracy and reduced feature dimensionality. Another hybrid model, in [80], which also employs Discrete Cosine Transform (DCT), achieves a remarkable 99% success rate in command recognition for Brazilian Portuguese while increasing processing speed by 25 times compared to non-optimized systems. By being speaker-independent, the ASR can operate effectively in real-world scenarios without requiring the same voice for training and testing. The study highlights the system's efficiency compared to traditional

models like Gaussian Mixture Models (GMM). It discusses its potential for implementation on various data processing boards, paving the way for practical applications in voice-activated technologies across diverse environments. Reference [23] explores an innovative method for enhancing text classification systems, specifically for Arabic texts, by integrating ACO with Chi-square statistics. This approach aims to improve the performance of SVMs by addressing the challenges posed by the sparsity of the Arabic language and reducing the dimensionality of training datasets through effective Feature Subset Selection (FSS). The study demonstrates that this method outperforms six contemporary FSS techniques in precision, recall, and F1 measures when applied to a specially curated corpus of Arabic documents. Ultimately, the combination of ACO and Chi-square statistics not only aids in managing large amounts of information but also optimizes SVM performance, paving the way for future research into its applicability across various datasets and classification tasks.
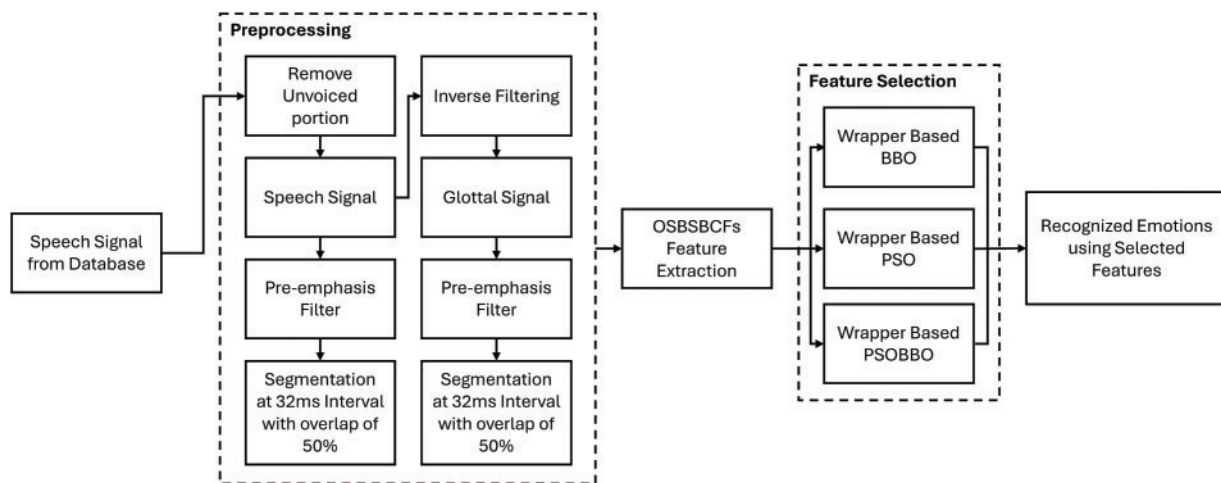
Figure 4: Block diagram of hybrid PSO-BBO for speech emotion recognition

4 Limitations and Future Directions

Despite their success in NLP tasks, bio-inspired algorithms face several challenges and limitations.

- High Computational Cost: Bio-inspired algorithms often require numerous iterations to converge to an optimal solution, leading to high computational demands [27,92]. For instance, genetic algorithms evaluate a population's fitness over multiple generations, which can be time-intensive for large datasets. For example, applying PSO to large corpora like PubMed abstracts in document clustering tasks results in prolonged computation times due to iterative fitness evaluations. This limits real-time applications, such as conversational AI, where low-latency responses are essential.
- Scalability Issues: Scaling bio-inspired algorithms to handle massive datasets or complex models is challenging. The stochastic nature of these algorithms can lead to inefficiencies when applied to high-dimensional data. Furthermore, the increasing size of modern NLP datasets exacerbates this challenge [17,101]. The specific task, such as training a DL model for language modeling with hyperparameters optimized via GA, showed degraded performance when scaling to datasets with millions of sentences. Such scalability issues can hinder their use in industrial-scale applications like search engines or recommendation systems.
- Parameter Sensitivity: The performance of bio-inspired algorithms heavily depends on their parameter settings (e.g., population size, mutation rate, and crossover rate in GA). Improper tuning can lead

to suboptimal results [102]. Finding the optimal parameter settings is often non-trivial and requires domain expertise.

- Lack of Deterministic Guarantees: Unlike traditional optimization methods, bio-inspired algorithms do not guarantee convergence to a global optimum. Their stochastic nature may result in different outcomes for the same problem. This lack of consistency can make them less reliable for critical applications [101]. For example, applying ACO to optimize dependency parsing trees yielded inconsistent results across multiple runs. This unpredictability can impact reproducibility in scientific research and production deployments.
- Integration Complexity: Integrating bio-inspired algorithms with existing NLP pipelines can be complex, especially when combining them with deep learning frameworks. Adapting these algorithms for specific tasks and frameworks often requires custom engineering [92,103]. Consider the scenarios incorporating GA for feature selection in transformer-based NLP systems that require significant customization to handle tensor computations. Such complexity can act as a barrier to adoption in fast-paced industrial environments.
- Resource and Energy Consumption: Bio-inspired algorithms can be resource-intensive, consuming significant computational power and energy [104]. This has environmental and cost implications, particularly when applied to large-scale NLP tasks. Running a GA for hyperparameter tuning on transformer models required extensive GPU hours, leading to high operational costs and carbon footprint. Such concerns are becoming increasingly important in sustainable AI research.

Several promising research directions have been proposed to overcome the challenges and unlock the full potential of bio-inspired algorithms in NLP. Hybrid approaches, which combine bio-inspired algorithms with traditional optimization methods or deep learning frameworks, offer a pathway to address computational and scalability issues. For example, hybrid models that utilize PSO for hyperparameter tuning alongside neural networks for feature extraction have shown potential for improved performance. Additionally, leveraging ensemble techniques can mitigate the shortcomings of individual methods by combining their strengths. Algorithmic enhancements are another vital area of focus. Developing advanced versions of bio-inspired algorithms, such as adaptive GA or multi-swarm PSO, can significantly improve their efficiency and convergence rates. Innovations like self-adaptive parameters or dynamically changing fitness functions can make these algorithms more robust and adaptive to various tasks.

Furthermore, parallel and distributed computing frameworks can alleviate computational challenges. We can effectively manage large-scale NLP tasks by implementing bio-inspired algorithms on cloud platforms or utilizing GPU and TPU acceleration while reducing processing times. Tailoring bio-inspired algorithms to specific NLP domains is another promising direction. Domain-specific customizations, such as incorporating linguistic constraints into the fitness function, can enhance performance in tasks like syntactic parsing or machine translation. Explainability frameworks for bio-inspired algorithms are also essential to increase trust and usability. Techniques such as surrogate modeling, visualization of evolutionary paths, and sensitivity analysis can improve the interpretability of these solutions, making them more suitable for critical applications.

Finally, energy-efficient designs for bio-inspired algorithms can reduce their environmental impact and operational costs. Methods like early stopping, sparsity induction, and efficient population sampling offer opportunities to make these algorithms more sustainable. By addressing these areas, researchers can enhance the applicability and effectiveness of bio-inspired algorithms in NLP, paving the way for broader adoption and success in real-world scenarios.

5 Conclusion

Inspired by natural processes and biological systems, bio-inspired algorithms offer a robust and adaptable approach to solving complex problems in NLP. This review explored the transformative potential of bio-inspired algorithms—specifically GA, PSO, and ACO—in enhancing NLP tasks such as text classification, sentiment analysis, machine translation, and speech recognition. Our objectives were to analyze their applications, assess comparative advantages, and identify limitations. Key findings reveal that these algorithms excel in optimization tasks, with GA improving feature selection (e.g., 95% accuracy in [36]), and PSO enhancing parameter tuning (e.g., 97.14% recognition rate in [78]) However, bio-inspired algorithms face notable challenges alongside their potential, including high computational cost, scalability issues, parameter sensitivity, and integration complexity with existing NLP systems. Despite these limitations, their stochastic and flexible nature makes them a compelling choice for tackling optimization challenges in NLP, especially when combined with other techniques. While bio-inspired algorithms are not without challenges, their contributions to NLP are undeniable. With advancements in algorithmic development and interdisciplinary research, these methods are poised to play a transformative role in the evolution of NLP technologies, addressing the complexities of human language with a blend of innovation and computational efficiency. Future research in hybrid approaches, algorithmic enhancements, and sustainability can further enhance their impact. With continued advancements, bio-inspired algorithms are poised to contribute significantly to NLP technologies' evolution.

Author Contributions: The authors confirm contribution to the paper as follows: study conception and design: Luong Vuong Nguyen, Huu-Tuong Ho; data collection: Huu-Tuong Ho, Luong Vuong Nguyen, Duong Nguyen Minh Huy, Thi-Thuy-Hoai Nguyen; analysis and interpretation of results: Luong Vuong Nguyen, Huu-Tuong Ho; draft manuscript preparation: Huu-Tuong Ho, Duong Nguyen Minh Huy, Thi-Thuy-Hoai Nguyen. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: The data and materials utilized in this review originate from publicly available databases and previously published studies, with proper citations included throughout the text. References to these sources can be found in the bibliography.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

Nomenclature
NLP   Natural Language Processing
ML   Machine Learning
DL   Deep Learning
PSO   Particle Swarm Optimization
GA   Genetic Algorithms
ACO   Ant Colony Optimization
ABC   Artificial Bee Colony
SA   Sentiment Analysis
SVM   Support Vector Machine

References

1. Chowdhary K, Chowdhary K. Natural language processing. In: Fundamentals of artificial intelligence. New Delhi, India: Springer; 2020. p. 603–49.

2. Mishra BK, Kumar R. Natural language processing in artificial intelligence. Boca Raton, FL, USA: CRC Press; 2020.

3. Hirschberg J, Manning CD. Advances in natural language processing. Science. 2015;349(6245):261–6. doi:10.1126/science.aaa8685.

4. Fanni SC, Febi M, Aghakhanyan G, Neri E. Natural language processing. In: Introduction to artificial intelligence. Wiesbaden, Germany: Springer; 2023. p. 87–99.

5. Khurana D, Koli A, Khatter K, Singh S. Natural language processing: state of the art, current trends and challenges. Multimed Tools Appl. 2023;82(3):3713–44. doi:10.1007/s11042-022-13428-4.

6. Storn R, Price K. Differential evolution-a simple and efficient heuristic for global optimization over continuous spaces. J Global Optimiz. 1997;11(4):341–59. doi:10.1023/A:1008202821328.

7. Dorigo M, Gambardella LM. Ant colony system: a cooperative learning approach to the traveling salesman problem. IEEE Trans Evol Comput. 1997;1(1):53–66. doi:10.1109/4235.585892.

8. Karaboga D, Basturk B. A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. J Global Optimiz. 2007;39(3):459–71. doi:10.1007/s10898-007-9149-x.

9. Mirjalili S, Mirjalili SM, Lewis A. Grey wolf optimizer. Adv Eng Softw. 2014;69:46–61. doi:10.1016/j.advengsoft.2013.12.007.

10. Faramarzi A, Heidarinejad M, Stephens B, Mirjalili S. Equilibrium optimizer: a novel optimization algorithm. Knowl Based Syst. 2020;191:105190. doi:10.1016/j.knosys.2019.105190.

11. Das B, Mukherjee V, Das D. Student psychology based optimization algorithm: a new population based optimization algorithm for solving optimization problems. Adv Eng Softw. 2020;146(4):102804. doi:10.1016/j.advengsoft.2020.102804.

12. Abdollahzadeh B, Gharehchopogh FS, Mirjalili S. African vultures optimization algorithm: a new nature-inspired metaheuristic algorithm for global optimization problems. Comput Ind Eng. 2021;158(4):107408. doi:10.1016/j.cie.2021.107408.

13. Li MD, Zhao H, Weng XW, Han T. A novel nature-inspired algorithm for optimization: virus colony search. Adv Eng Softw. 2016;92(3):65–88. doi:10.1016/j.advengsoft.2015.11.004.

14. Özcan N, Utku S, Berber T. Artificial circulation system algorithm: a novel bio-inspired algorithm. Comput Model Eng Sci. 2025;142(1):635–63. doi:10.32604/cmes.2024.055860.

15. Nguyen TH, Nguyen LV, Jung JJ, Agbehadji IE, Frimpong SO, Millham RC. Bio-inspired approaches for smart energy management: state of the art and challenges. Sustainability. 2020;12(20):8495. doi:10.3390/su12208495.

16. Nguyen LV, Nguyen TH, Pham HTN, Vo QT, Duong HT, Nguyen-Thi TA. Bio-inspired clustering: an ensemble method for user-based collaborative filtering. In: 2023 International Conference on Intelligence of Things; 2023 Oct 25–27; Ho Chi Minh City, Vietnam. p. 26–35.

17. Nguyen LV. Swarm intelligence-based multi-robotics: a comprehensive review. AppliedMath. 2024;4(4):1192–210. doi:10.3390/appliedmath4040064.

18. Nguyen LV. Advanced clustering techniques with bio-inspired for collaborative filtering recommendation systems. Vietnam J Comput Sci. 2025;12(1):21–36. doi:10.1142/S2196888824400013.

19. Alba E, Luque G, Araujo L. Natural language tagging with genetic algorithms. Inf Proces Lett. 2006;100(5):173–82. doi:10.1016/j.ipl.2006.07.002.

20. Mendiratta S, Turk N, Bansal D. Automatic speech recognition using optimal selection of features based on hybrid ABC-PSO. In: 2016 International Conference on Inventive Computation Technologies (ICICT). Coimbatore, India; 2016. p. 1–7.

21. Kurniawati I, Pardede HF. Hybrid Method of Information Gain and Particle Swarm Optimization for Selection of Features of SVM-Based Sentiment Analysis. In: 2018 International Conference on Information Technology Systems and Innovation (ICITSI); 2018 Oct 22–26; Padang, Indonesia. p. 1–5.

22. Durga P, Godavarthi D, Kant S, Basa SS. Aspect-based drug review classification through a hybrid model with ant colony optimization using deep learning. Discov Comput. 2024;27(1):19. doi:10.1007/s10791-024-09441-w.

23. Mesleh M, Kanaan G. Support vector machine text classification system: using ant colony optimization based feature subset selection. In: 2008 International Conference on Computer Engineering & Systems; Cairo, Egypt; 2008. p. 143–8. doi:10.1109/ICCES.2008.4772984.

24. Aghdam MH, Ghasem-Aghaee N, Ehsan Basiri M. Application of ant colony optimization for feature selection in text categorization. In: 2008 IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence). Hong Kong, China; 2008. p. 2867–73.

25. Nema P, Sharma V. Multi-label text categorization based on feature optimization using ant colony optimization and relevance clustering technique. In: 2015 International Conference on Computers, Communications, and Systems (ICCCS); 2015 Nov 2–3; Kanyakumari, India. p. 1–5.

26. Aggarwal S, Chhabra B. Sentimental analysis of tweets using ant colony optimizations. In: 2017 International Conference on Intelligent Sustainable Systems (ICISS). Palladam, India; 2017. p. 1219–23.

27. Hossain SKM, Ema SA, Sohn H. Rule-based classification based on ant colony optimization: a comprehensive review. Appl Comput Intell Soft Computi. 2022;2022:1–17. doi:10.1155/2022/2232000.

28. Gite S, Patil S, Dharrao D, Yadav M, Basak S, Rajendran A, et al. Textual feature extraction using ant colony optimization for hate speech classification. Big Data Cogn Comput. 2023;7(1):45. doi:10.3390/bdcc7010045.

29. Tripathy A, De UC, Dash BB, Patra SS, Pattanayak BK, Pandey TN. Sentiment classification of reviews using combination of ant colony optimization and ANN. In: 2023 3rd International Conference on Mobile Networks and Wireless Communications (ICMNWC). Tumkur, India; 2023; p. 1–6.

30. Alsaduni I, Baseer MA, Alluhaidan M, Tazeen N. A novel deep belief neural framework based on ant colony to analyze the sentiment score in drug review dataset. Cham, Switzerland: Springer Nature; 2024. p. 384–99.

31. Kumar S, Pathak SK. A Novel Approach for Sentiment Analysis and Optimization Using PSO. In: 2021 International Conference on Advances in Technology, Management & Education (ICATME). Bhopal, India; 2021. p. 43–7.

32. Mandal S, Singh GK, Pal A. PSO-based text summarization approach using sentiment analysis. Singapore: Springer; 2018. p. 845–54.

33. Machová K, Mikula M, Gao X, Mach M. Lexicon-based sentiment analysis using the particle swarm optimization. Electronics. 2020;9(8):1317. doi:10.3390/electronics9081317.

34. Song W, Park SC. Genetic algorithm for text clustering based on latent semantic indexing. Comput Math Appl. 2009;57(11–12):1901–7. doi:10.1016/j.camwa.2008.10.010.

35. Song W, Li CH, Park SC. Genetic algorithm for text clustering using ontology and evaluating the validity of various semantic similarity measures. Expert Syst Appl. 2009;36(5):9095–104. doi:10.1016/j.eswa.2008.12.046.

36. Özel SA. A Web page classification system based on a genetic algorithm using tagged-terms as features. Expert Syst Appl. 2011;38(4):3407–15. doi:10.1016/j.eswa.2010.08.126.

37. Uğuz H. A two-stage feature selection method for text categorization by using information gain, principal component analysis and genetic algorithm. Knowl-Based Syst. 2011;24(7):1024–32. doi:10.1016/j.knosys.2011.04.014.

38. Chen H, Jiang W, Li C, Li R. A heuristic feature selection approach for text categorization by using chaos optimization and genetic algorithm. Math Problems Eng. 2013;2013:1–6.

39. Uysal AK, Gunal S. Text classification using genetic algorithm oriented latent semantic features. Expert Syst Appl. 2014;41(13):5938–47. doi:10.1016/j.eswa.2014.03.041.

40. Bidi N, Elberrichi Z. Feature selection for text classification using genetic algorithms. In: 2016 8th International Conference on Modelling, Identification and Control (ICMIC); 2016 Nov 15–17; Algiers, Algeria; 2016. p. 806–10.

41. Ghareb AS, Bakar AA, Hamdan AR. Hybrid feature selection based on enhanced genetic algorithm for text categorization. Expert Syst Appl. 2016;49(5):31–47. doi:10.1016/j.eswa.2015.12.004.

42. Kwon K, Kang D, Choi S, Park H, Chung IJ. A study on sentiment trend analysis method using ant colony optimization algorithm and SentiWordNet. In: Proceedings of the Korea Information Processing Society Conference. Seoul, republic of Korea; 2014. p. 948–51.

43. Ahmad SR, Bakar AA, Yaakub MR. Ant colony optimization for text feature selection in sentiment analysis. Intell Data Anal. 2019;23(1):133–58. doi:10.3233/IDA-173740.

44. Hamdi M. Affirmative ant colony optimization based support vector machine for sentiment classification. Electronics. 2022;11(7):1051. doi:10.3390/electronics11071051.

45. Kavitha P, Lalitha SD. ANT colony optimization algorithm for feature selection in sentiment analysis of social media data. ICTACT J Soft Comput. 2024;14(4):3334–9. doi:10.21917/ijsc.2024.0468.

46. Kumar Gupta D, Srikanth Reddy K, Shweta, Ekbal A. PSO-ASent: feature selection using particle swarm optimization for aspect based sentiment analysis. Cham, Switzerland: Springer International Publishing; 2015. p. 220–33.

47. Yue Y, Peng Y, Wang D. Deep learning short text sentiment analysis based on improved particle swarm optimization. Electronics. 2023;12(19):4119. doi:10.3390/electronics12194119.

48. Asri AM, Ahmad SR, Yusop NMM. Feature selection using particle swarm optimization for sentiment analysis of drug reviews. Int J Adv Comput Sci Appl. 2023;14(5):286–95. doi:10.14569/issn.2156-5570.

49. Katarya R, Yadav A. A comparative study of genetic algorithm in sentiment analysis. In: 2018 2nd International Conference on Inventive Systems and Control (ICISC). Coimbatore, India; 2018. p. 136–41.

50. Keshavarz H, Abadeh MS. ALGA: adaptive lexicon learning using genetic algorithm for sentiment analysis of microblogs. Knowl-Based Syst. 2017;122(4):1–16. doi:10.1016/j.knosys.2017.01.028.

51. Iqbal F, Hashmi JM, Fung BCM, Batool R, Khattak AM, Aleem S, et al. A hybrid framework for sentiment analysis using genetic algorithm based feature reduction. IEEE Access. 2019;7:14637–52. doi:10.1109/ACCESS.2019.2892852.

52. Li J, Liang Y. Refining word embeddings based on improved genetic algorithm for sentiment analysis. In: 2020 IEEE 9th Joint International Information Technology and Artificial Intelligence Conference (ITAIC); 2020 Dec 11–13; Chongqing, China. p. 213–6.

53. Ernawati S, Wati R, Nuris N, Marita LS, Yulia ER. Comparison of naïve bayes algorithm with genetic algorithm and particle swarm optimization as feature selection for sentiment analysis review of digital learning application. J Phys: Conf Ser. 2020;1641(1):012040. doi:10.1088/1742-6596/1641/1/012040.

54. Junior AB, da Silva NFF, Rosa TC, Junior CGC. Sentiment analysis with genetic programming. Inf Sci. 2021;562(3):116–35. doi:10.1016/j.ins.2021.01.025.

55. Dangi D, Bhagat A, Kumar Dixit D. Emerging applications of artificial intelligence, machine learning and data science. Comput Mater Contin. 2022;70(3):5399–419. doi:10.32604/cmc.2022.020431.

56. Huang JY, Tung CL, Lin WZ. Using social network sentiment analysis and genetic algorithm to improve the stock prediction accuracy of the deep learning-based approach. Int J Comput Intell Syst. 2023;16(1):93. doi:10.1007/s44196-023-00276-9.

57. Ferreira L, Dosciatti M, Nievola J, Paraiso E. Using a genetic algorithm approach to study the impact of imbalanced corpora in sentiment analysis. In: Proceedings of the Twenty-Eighth International Florida Artificial Intelligence Research Society Conference; 2015 May 18–20; Hollywood, FL, USA. p. 163–8.

58. Tambouratzis G. Using an ant colony metaheuristic to optimize automatic word segmentation for ancient Greek. IEEE Trans Evol Comput. 2009;13(4):742–53. doi:10.1109/TEVC.2009.2014363.

59. Alsaeedan W, Menai MEB, Al-Ahmadi S. A hybrid genetic-ant colony optimization algorithm for the word sense disambiguation problem. Inf Sci. 2017;417(4):20–38. doi:10.1016/j.ins.2017.07.002.

60. Habibulla I, Akhat A, Mijit S, Husiyin M. Intelligent recognition system for English-Chinese translation based on ant colony optimization algorithm. In: 2023 International Conference on Network, Multimedia and Information Technology (NMITCON); 2023 Sep 1–2; Bengaluru, India; 2023. p. 1–6.

61. Zhang W, Lian H, Zhang X. Research on Japanese-Chinese statistical machine translation based on improved ant colony algorithm. In: 2023 International Conference on Mechatronics, IoT and Industrial Informatics (ICMIII); 2023 Jun 9–11; Melbourne, VIC, Australia. p. 428–31.

62. Shi J, Tao L. Design of english machine translation system based on ant colony algorithm. Singapore: Springer Nature; 2024. p. 401–10.

63. Farzi S, Faili H. A swarm-inspired re-ranker system for statistical machine translation. Comput Speech Lang. 2015;29(1):45–62. doi:10.1016/j.csl.2014.07.002.

64. Montes Olguín JA, Mizera-Pietraszko J, Rodriguez Jorge R, Martínez García EA. Particle swarm optimization as a new measure of machine translation efficiency. Cham, Switzerland: Springer International Publishing; 2018. p. 161–70.

65. Li J. Development and research of computer-aided english-chinese translation system based on particle swarm optimization programming. Sci Program. 2022;2022(7):1–12. doi:10.1155/2022/4254932.

66. Zhao L, Zheng Y. Research on correcting the accuracy of english translation based on improved particle swarm optimization algorithm. In: 2023 IEEE International Conference on Sensors, Electronics and Computer Engineering (ICSECE); 2023 Aug 18–20; Jinzhou, China. p. 1504–8.

67. Chartcharnchai P, Jewajinda Y, Praditwong K. A categorical particle swarm optimization for hyperparameter optimization in low-resource transformer-based machine translation. In: 2024 28th International Computer Science and Engineering Conference (ICSEC); 2024 Nov 6–8; Khon Kaen, Thailand. 2024. p. 1–6.

68. Echizen-ya H, Araki K, Momouchi Y, Tochinai K. Machine translation method using inductive learning with genetic algorithms. In: Proceedings of the 16th conference on Computational linguistics; 1996 Aug 5–9; Copenhagen, Denmark. p. 1020–3.

69. Sofianopoulos S, Tambouratzis G. Multi-objective optimisation of real-valued parameters of a hybrid MT system using genetic algorithms. Pattern Recognit Lett. 2010;31(12):1672–82. doi:10.1016/j.patrec.2010.05.018.

70. Singh M, Kumar R, Chana I. GA-based machine translation system for Sanskrit to Hindi language. Singapore: Springer; 2018. p. 419–27.

71. Ganapathy K. A study of genetic algorithms for hyperparameter optimization of neural networks in machine translation. arXiv: 2009.08928. 2020.

72. Yu Y. Application of the neural network based on the multilayer perceptron genetic algorithm in chinese-english two-way translation. J Sensors. 2022;2022(1):1–12. doi:10.1155/2022/1510868.

73. Nemati S, Basiri ME. Text-independent speaker verification using ant colony optimization-based selected features. Exp Syst Appl. 2011;38(1):620–30. doi:10.1016/j.eswa.2010.07.011.

74. Rashno A, Ahadi SM, Kelarestaghi M. Text-independent speaker verification with ant colony optimization feature selection and support vector machine. In: 2015 2nd International Conference on Pattern Recognition and Image Analysis (IPRIA). Rasht, Iran; 2015. p. 1–5.

75. Jalili F, Jafari Barani M. Speech recognition using combined fuzzy and ant colony algorithm. Int J Electr Comput Eng. 2016;6(5):2205.

76. Najkar N, Razzazi F, Sameti H. A novel approach to HMM-based speech recognition systems using particle swarm optimization. Math Comput Model. 2010;52(11–12):1910–20. doi:10.1016/j.mcm.2010.03.041.

77. Chan KY, Yiu CKF, Dillon TS, Nordholm S, Ling SH. Enhancement of speech recognitions for control automation using an intelligent particle swarm optimization. IEEE Trans Ind Inform. 2012;8(4):869–79. doi:10.1109/TII.2012.2187910.

78. Selvaraj L, Ganesan B. Enhancing speech recognition using improved particle swarm optimization based hidden markov model. Sci World J. 2014;2014(2):1–10. doi:10.1155/2014/270576.

79. Batista GC, Santos Silva WL, Menezes AG. Automatic speech recognition using support vector machine and particle swarm optimization. In: 2016 IEEE Symposium Series on Computational Intelligence (SSCI). Athens, Greece; 2016. p. 1–6.

80. Batista GC, Silva WLS, de Oliveira DL, Saotome O. Automatic speech patterns recognition of commands using SVM and PSO. Multimed Tools Appl. 2019;78(22):31709–31. doi:10.1007/s11042-019-07956-9.

81. Yogesh CK, Hariharan M, Ngadiran R, Adom AH, Yaacob S, Berkai C, et al. A new hybrid PSO assisted biogeography-based optimization for emotion and stress recognition from speech signal. Exp Syst Appl. 2017;69:149–58. doi:10.1016/j.eswa.2016.10.035.

82. Kwong S, Chau CW, Halang WA. Genetic algorithm for optimizing the nonlinear time alignment of automatic speech recognition systems. IEEE Trans Ind Electron. 1996;43(5):559–66. doi:10.1109/41.538613.

83. Kwong S, Chau CW. Analysis of parallel genetic algorithms on HMM based speech recognition system. IEEE Trans Consumer Electron. 1997;43(4):1229–33. doi:10.1109/30.642391.

84. Lin CT, Nein HW, Hwu JY. GA-based noisy speech recognition using two-dimensional cepstrum. IEEE Trans Speech Audio Process. 2000;8(6):664–75. doi:10.1109/89.876300.

85. Zamalloa M, Bordel G, Rodriguez L, Penagarikano M. Feature selection based on genetic algorithms for speaker recognition. In: 2006 IEEE Odyssey—The Speaker and Language Recognition Workshop. San Juan, PR, USA; 2006. p. 1–8.

86. Pan ST, Wu CH, Lai CC. The application of improved genetic algorithm on the training of neural network for speech recognition. In: Second International Conference on Innovative Computing, Informatio and Control (ICICIC 2007); 2007 Sep 5–7; Kumamoto, Japan. p. 168–8.

87. Pan ST, Chen CF, Zeng JH. Speech recognition via Hidden Markov Model and neural network trained by genetic algorithm. In: 2010 International Conference on Machine Learning and Cybernetics; 2010 Jul 11–14; Qingdao, China. p. 2950–5.

88. Sedaaghi MH, Kotropoulos C, Ververidis D. Using adaptive genetic algorithms to improve speech emotion recognition. In: 2007 IEEE 9th Workshop on Multimedia Signal Processing; 2007 Oct 1–3; Crete, Greece. p. 461–4.

89. Priyadarshani PGN, Dias NGJ, Punchihewa A. Genetic algorithm approach for sinhala speech recognition. In: 2012 IEEE 55th International Midwest Symposium on Circuits and Systems (MWSCAS); 2012 Aug 5–8; Boise, ID, USA. p. 896–9.

90. Ibrahim AB, Seddiq YM, Meftah AH, Alghamdi M, Selouani SA, Qamhan MA, et al. Optimizing arabic speech distinctive phonetic features and phoneme recognition using genetic algorithm. IEEE Access. 2020;8:200395–411. doi:10.1109/ACCESS.2020.3034762.

91. Alhijawi B, Awajan A. Genetic algorithms: theory, genetic operators, solutions, and applications. Evol Intell. 2024;17(3):1245–56. doi:10.1007/s12065-023-00822-6.

92. Katoch S, Chauhan SS, Kumar V. A review on genetic algorithm: past, present, and future. Multimed Tools Appl. 2021;80(5):8091–126. doi:10.1007/s11042-020-10139-6.

93. Gao N, Yang N, Tang J. Ancestral genome inference using a genetic algorithm approach. PLoS One. 2013;8(5):e62156. doi:10.1371/journal.pone.0062156.

94. Weng W. Artificial intelligence–genetic algorithm. In: A Beginner's guide to informatics and artificial intelligence: a short course towards practical problem solving. Cham, Switzerland: Springer; 2024. p. 39–49.

95. Wang D, Tan D, Liu L. Particle swarm optimization algorithm: an overview. Soft Comput. 2018;22(2):387–408. doi:10.1007/s00500-016-2474-6.

96. Dorigo M, Birattari M, Stutzle T. Ant colony optimization. IEEE Comput Intell Mag. 2006;1(4):28–39. doi:10.1109/MCI.2006.329691.

97. Martens D, De Backer M, Haesen R, Vanthienen J, Snoeck M, Baesens B. Classification with ant colony optimization. IEEE Trans Evol Comput. 2007;11(5):651–65. doi:10.1109/TEVC.2006.890229.

98. Wajeed MA, Adilakshmi T. Adopting ant colony optimization for supervised text classification. In: 2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI). Jaipur, India; 2016. p. 2562–6.

99. Macedo J, Gidey HK, Rebuli KB, Machado P. Evolving user interfaces: a neuroevolution approach for natural human-machine interaction. Cham, Switzerland: Springer Nature; 2024. p. 246–64.

100. Lan ML, Pan ST, Lai CC. Using genetic algorithm to improve the performance of speech recognition based on artificial neural network. In: First International Conference on Innovative Computing, Information and Control—Volume I (ICICIC'06). Beijing, China; 2006. p. 527–30. doi:10.1109/ICICIC.2006.372

101. Gill SS, Buyya R. Bio-inspired algorithms for big data analytics: a survey, taxonomy, and open challenges. In: Big data analytics for intelligent healthcare management. Amsterdam, Netherlands: Elsevier; 2019. p. 1–17.

102. Yadav A, Vishwakarma DK. A comparative study on bio-inspired algorithms for sentiment analysis. Cluster Comput. 2020;23(4):2969–89. doi:10.1007/s10586-020-03062-w.

103. Pham TH, Raahemi B. Bio-inspired feature selection algorithms with their applications: a systematic literature review. IEEE Access. 2023;11:43733–58. doi:10.1109/ACCESS.2023.3272556.

104. Siddique N, Adeli H. Nature inspired computing: an overview and some future directions. Cogn Comput. 2015;7(6):706–14. doi:10.1007/s12559-015-9370-8.