Towards Effective COVID-19 Sentiment Analysis Using Bio-Inspired Feature Optimization

Mohamed Debbab, Moustafa Maaskri, Mohamed Goismi, Abdelkader Boudaoud, and Djamel Seghier

Original scientific article

***Abstract*—Analyzing sentiments in social media content related to COVID-19 presents major challenges, especially given the sheer volume of data and the complexity of text features. These factors often reduce the effectiveness of classification models, making it crucial to apply smart feature selection to boost both accuracy and efficiency. This paper develops a feature selection framework that strategically integrates statistical filtering methods with evolutionary computation techniques, specifically incorporating Genetic Algorithms (GA) and Particle Swarm Optimization (PSO) for optimal feature space reduction—to address these challenges. By focusing on selecting the most meaningful features, our method reduces unnecessary complexity while retaining the information that matters most. Tests carried out on a dataset of COVID-19 tweets show that this approach improves classification accuracy by around 7% compared to standard feature selection methods. These results highlight how combining statistical filtering with bio-inspired optimization can play an important role in improving sentiment analysis, especially during critical situations like the COVID-19 pandemic.**

***Index Terms*—Sentiment Analysis, Covid-19, Wrapper Selection, Feature Selection, GA, PSO, Machine Learning.**

I. Introduction

THe critical role of sentiment analysis in extracting insights from social media data has been magnified during global health emergencies, with the COVID-19 pandemic serving as a prime case study for understanding large-scale public sentiment dynamics. The vast and noisy nature of social media data, especially during crises, makes feature selection a critical task for improving both the accuracy and computational efficiency of classification models. For example, Alqurashi [1] highlighted the growing interest in applying machine learning techniques to sentiment analysis on Twitter, emphasizing challenges specific to large datasets. Research specific to COVID-19 sentiment analysis has also expanded. Moustafa et al. [2] conducted a comprehensive study applying machine learning and deep learning models for multi-class sentiment analysis of COVID-19 tweets, revealing the importance of optimizing data representations to improve accuracy and generalization. These studies underscore the need for effective feature selection to mitigate dimensionality issues, particularly when working with large-scale text data. Among advanced feature selection strategies, Nature-inspired metaheuristics, particularly Genetic Algorithms (GA) and Particle Swarm Optimization (PSO), have demonstrated significant potential in computational optimization tasks. These algorithms leverage principles from evolutionary computation and swarm intelligence to search for optimal feature subsets that enhance classification performance. Hayatin et al. [4] demonstrated the potential of PSO for optimizing feature sets in Twitter-based sentiment analysis, showing notable accuracy improvements. Similarly, Sangam and Shinde [6] used GA for feature selection in opinion mining, highlighting its effectiveness in improving classification accuracy for social media data. Hybrid approaches that combine deep learning models with evolutionary feature selection have also emerged. Pookduang et al. [7] compared transformer-based models like RoBERTa to traditional classifiers in sentiment analysis, emphasizing the advantage of using more advanced techniques to handle complex textual data. Furthermore, Kishore et al. [8] presented a sparse vector-based embedding strategy to manage high-dimensional sentiment features effectively. Despite these advancements, limited work has specifically addressed the use of hybrid feature selection combining bio-inspired algorithms like GA and PSO with deep learning models for COVID-19 sentiment analysis. This study aims to fill that gap by systematically evaluating the effectiveness of GA and PSO in improving classification accuracy on COVID-19-related tweet datasets.

The main contributions of this paper are:

- A wrapper-based feature selection framework for COVID-19 tweet sentiment classification that leverages PSO and GA on top of TF-IDF features.
- Comprehensive evaluation across SVM, KNN (k=3), and MLP using library defaults (no tuning); the same hyperparameters were used with and without PSO/GA feature selection (only the feature subset changes), within a clearly defined preprocessing/balancing pipeline and stratified 10-fold cross-validation on the Coronavirus Tweet NLP dataset; test results are reported once with no additional adjustments.
- Substantial accuracy gains over no feature selection, e.g., MLP: 52.7% → 94.21% (PSO) / 94.48% (GA); consistent boosts for SVM and KNN as well, confirming the value of evolutionary wrappers.

Manuscript received October 24, 2025; revised November 6, 2025. Date of publication December 19, 2025. Date of current version December 19, 2025.
M. Debbab and M. Goismi are with the Department of Science and Technology, Faculty of Applied Sciences, Ibn Khaldoun University, Tiaret, Algeria (e-mails: mohamed.debbab,mohamed.goismi@univ-tiaret.dz).
M. Maaskri and D. Seghier are with the Department of Computer Science, Faculty of Mathematics and Computer Science, Ibn Khaldoun University, Tiaret, Algeria (e-mails: moustafa.maaskri, djamel.seghier@univ-tiaret.dz).
A. Boudaoud is with the Department of Computer Science, Faculty of Mathematics and Computer Science, University Center of El Bayadh, Elbayadh, Algeria, (e-mails: aek.boudaoud@cu-elbayadh.dz).
Digital Object Identifier (DOI): 10.24138/jcomss-2025-0181

- Quantified dimensionality reduction of roughly 50% with both PSO and GA, improving efficiency while maintaining or improving performance.
- Convergence and configuration transparency, including PSO/GA parameter settings and convergence plots to support reproducibility and insight into optimization behavior.

The remainder of this paper is organized as follows. Section II reviews related work. Section III details the materials and methods, including the dataset, preprocessing pipeline, TF–IDF feature construction, and GA/PSO wrapper-based feature selection with the classification setup. Section IV reports and discusses the experimental results. Section V presents the external validation. Section VI concludes the paper and outlines directions for future research.

II. Related Work

Research on COVID-19 Twitter sentiment has evolved along three intersecting axes: (i) representations for short, noisy text, (ii) modeling paradigms (classical ML, ensembles, and transformers), and (iii) feature selection/optimization to control dimensionality and cost.

Text representations. Early pipelines relied on sparse lexical features—bag-of-words, character and word n-grams, and TF–IDF vectors—which remain competitive under careful feature engineering and regularization [26], [18]. Dense semantic encoders (word2vec/GloVe), and later contextual transformers (BERT family), improved robustness to lexical variation and context [21], [9]. For COVID-19 tweets in particular, CT-BERT, pre-trained on tens of millions of in-domain messages, consistently outperformed generic BERT on pandemic-related classification and analysis tasks [27].

Classical vs. transformer baselines. On five-class COVID-19 Twitter sentiment, Maaskri et al. compared TF–IDF plus ensemble learners (voting, bagging, stacking) with a BERT baseline; the best macro F1 was ≈ 0.74 for BERT and ≈ 0.65 for the strongest ensemble, establishing a widely cited reference point for multi-class tweet sentiment during the pandemic [2]. Hybrid neural designs tailored to regional content also proved effective: Sitaula and Shahi showed that syntactic + semantic feature channels coupled with a CNN boost robustness on Nepali tweets, a useful indicator for multilingual settings with noisy orthography and sparse resources [16]. Other hybrids fused transformer embeddings with statistical features (e.g., hierarchical TF–IDF + BERT) to balance accuracy and runtime [5].

Feature selection and metaheuristics. Because TF–IDF and n-gram spaces are extremely high-dimensional, feature selection (FS) is crucial for stability, interpretability, and speed. Filter methods (e.g., Chi-Square, mutual information) are scalable but model-agnostic; embedded methods (L1/LASSO, elastic net) incorporate sparsity during training; wrapper methods explicitly search for a performant subset with a model in the loop [26], [14]. Bio-inspired search—Genetic Algorithms (GA) [22] and Particle Swarm Optimization (PSO) [23]—has become popular in text mining because it directly optimizes task metrics over discrete feature subsets while controlling redundancy. In sentiment analysis, PSO-based and GA-based wrappers routinely enhance classical learners and can even serve as hyper-parameter optimizers for neural models, improving accuracy/efficiency trade-offs on short social text [20], [15], [12]. Recent nature-inspired variants (e.g., teaching–learning-based optimization, cuckoo search) report similar gains by retaining compact, discriminative subsets that generalize well [11], [10].

COVID-19-specific pipelines. Within the pandemic literature, hybrid bio-inspired FS has been particularly effective. Goismi et al. combined Chi-Square filtering with GA/PSO (and other heuristics such as HHO/WOA)[30], showing substantial dimensionality reduction without sacrificing accuracy on COVID-19 tweets—evidence that careful wrapper design can make classical models competitive with heavy neural baselines when features are curated. Related comparisons confirm the trend: Moustafa et al. contrasted ensembles (voting, bagging, stacking) against BERT and observed the expected transformer advantage but also highlighted that robust feature engineering narrows the gap at much lower cost [2]. Complementary studies explored transformer-plus-FS hybrids (e.g., H-TFIDF + BERT embeddings) that improved accuracy and runtime—an attractive compromise for production pipelines [32].

Practical considerations. COVID-19 tweet corpora are imbalanced and noisy; thus resampling (e.g., SMOTE) and stratified evaluation are standard [13]. Efficiency matters for streaming or large-scale monitoring, where wrapper FS can shrink inference latency and memory footprint. Finally, interpretability (e.g., LIME/SHAP) benefits from compact feature sets, making post-hoc explanations more faithful and easier to audit in public-health contexts [3].

Positioning of our work. Building on these insights, we adopt wrapper-based GA/PSO selection over TF–IDF for five-class COVID-19 tweet sentiment. Our results align with prior evidence that metaheuristic FS yields sizable gains for classical models; they further show that with a well-designed wrapper and stable evaluation, classical learners (especially MLP/SVM) can match or surpass heavier baselines at a fraction of the cost.

III. Materials and Methods

The figure 1 illustrates a framework for sentiment analysis, tailored to process social media content related to COVID-19. The workflow starts with collecting a dataset of COVID-19 tweets, followed by a crucial data preprocessing stage. During this step, unnecessary elements such as usernames, URLs, punctuation, numbers, special characters, stopwords, and short words are removed. The text is then tokenized, stemmed, and converted to lowercase to ensure consistency and reduce noise.

Following data preprocessing, textual features are vectorized using TF-IDF (Term Frequency-Inverse Document Frequency) transformation, converting unstructured text into structured numerical representations compatible with machine learning models. To further enhance the model's efficiency, bio-inspired wrapper-based feature selection algorithms—specifically Particle Swarm Optimization (PSO) and Genetic Algorithm (GA)—are applied to select the most relevant features.

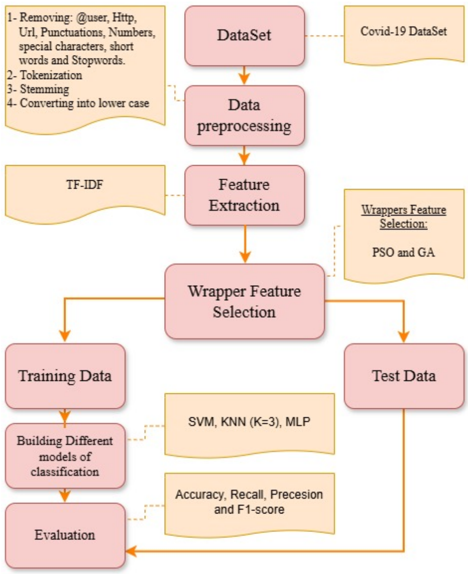

Fig. 1. Proposed approach.

The dataset is then split into training and testing sets, and several classification algorithms, including Support Vector Machines (SVM), K-Nearest Neighbors (K=3), and Multi-Layer Perceptron (MLP), are used to build predictive models. Finally, the system evaluates these models using standard metrics such as Accuracy, Precision, Recall, and F1-score. This framework demonstrates the value of integrating bio-inspired feature selection to improve classification accuracy, particularly when dealing with the high-dimensional, noisy text typical of social media data during the COVID-19 pandemic.

A. Particle Swarm Optimization (PSO)

Particle Swarm Optimization (PSO) is a nature-inspired metaheuristic algorithm that simulates the social behavior observed in flocks of birds and schools of fish. Originally introduced by Kennedy and Eberhart in 1995. [23], PSO has since become a widely adopted method for solving both continuous and discrete optimization problems, particularly in the domain of machine learning feature selection. The algorithm operates by maintaining a population of candidate solutions, referred to as particles, which navigate the search space by iteratively adjusting their positions and velocities. These adjustments are guided by each particle's own best-known position as well as the best position discovered by the entire swarm. When applied to feature selection, each particle represents a potential subset of features, with each dimension corresponding to a specific feature within the dataset. The objective of PSO in this context is to identify an optimal feature subset that maximizes the performance of the classification model (see Fig. 2).

Mathematical formulation:

$$v_i(t+1) = w \cdot v_i(t) + c_1 r_1 \left(p_{\text{best},i} - x_i(t)\right) + c_2 r_2 \left(g_{\text{best}} - x_i(t)\right) \tag{1}$$

where:

- $v_i(t+1)$ is the velocity of particle i at time $t+1$,
- $x_i(t)$ is the position of particle i at time t,
- $p_{\text{best},i}$ and g_{best} are the personal best and global best positions,
- w is the inertia weight,
- c_1, c_2 are cognitive and social coefficients,
- r_1, r_2 are random numbers.

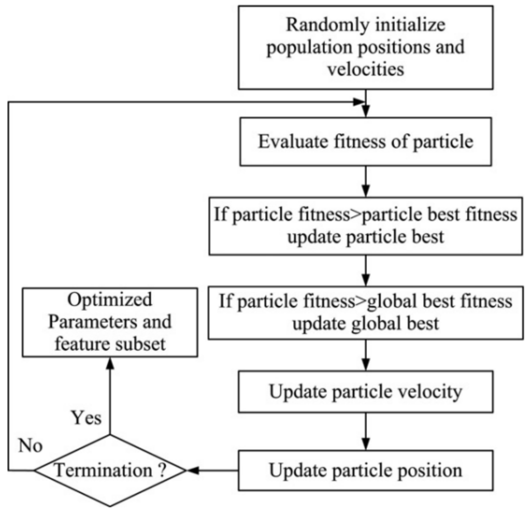

Fig. 2. The flowchart of PSO algorithm.

B. Genetic Algorithm (GA)

Genetic Algorithms (GAs) belong to a family of optimization techniques inspired by the principles of natural selection and genetic evolution. Initially introduced by Holland in 1975 [34], GAs emulate the evolutionary process by evolving a population of candidate solutions over successive generations. This evolution is driven by key genetic operators, namely selection, crossover, and mutation. In the context of feature selection, each individual in the population—referred to as a chromosome—encodes a possible subset of features, and its fitness is typically assessed based on classification accuracy or other performance criteria. The evolutionary cycle involves selecting the fittest individuals, combining them through crossover, and occasionally applying mutations to explore new areas of the search space. Due to their robustness and adaptability, GAs have proven particularly effective for combinatorial optimization tasks such as feature selection, where the objective is to identify an optimal subset of features that enhances the overall performance of the learning model (see Fig. 3).

- Crossover: Combining two parent solutions to create offspring.
- Mutation: Randomly altering an offspring to maintain diversity.
- Fitness: Evaluating the quality of a solution (e.g., classification accuracy).

C. DataSet

This study utilized English-language tweets concerning the coronavirus, published on Twitter between January 1st and

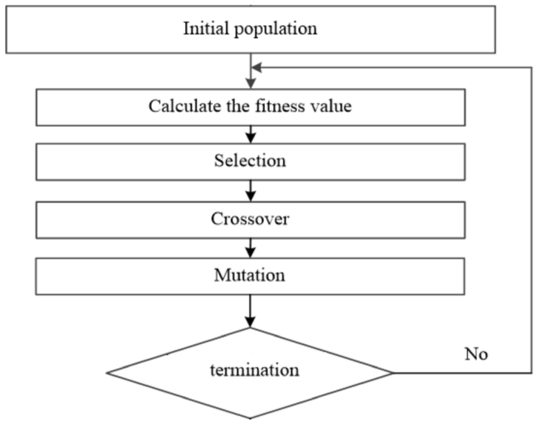

Fig. 3. The flowchart of GA algorithm.

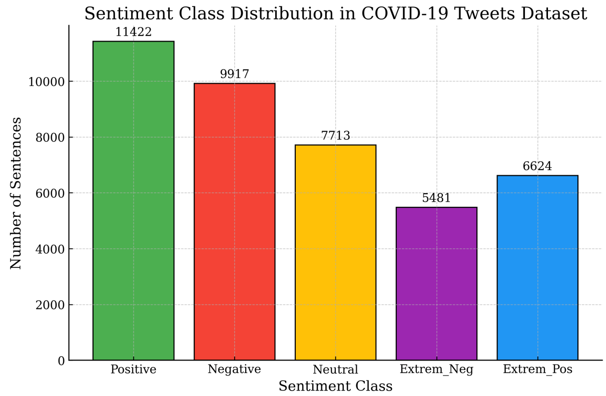

Fig. 4. Covid-19 Imbalanced DataSet Class distribution.

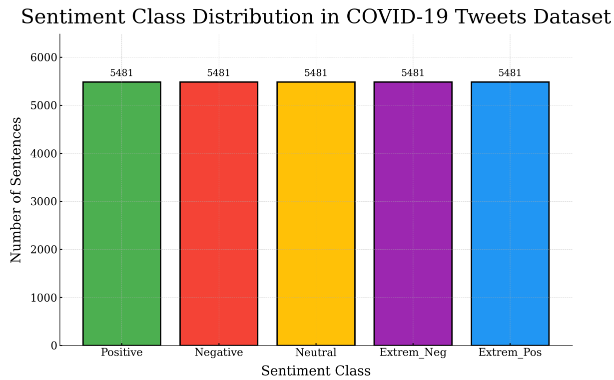

Fig. 5. Covid-19 balanced DataSet Class distribution.

December 31st, 2020. The dataset includes tweets gathered from several countries across the globe during the pandemic timeline and is publicly available in [17]. To compile the dataset, a set of predefined and commonly used keywords associated with COVID-19, such as "COVID-19", "coronavirus", "lockdown", "isolation", "quarantine", "pandemic", and "ncov-2020" was employed as filtering criteria.

The dataset consists of a training set comprising 41,157 tweets and a testing set of 3,798 tweets. Figure 4 illustrates the distribution of the unbalanced data across the five sentiment classes, highlighting the class imbalance inherent in the raw dataset. To address this imbalance, a downsampling technique was applied to the majority classes [33]. The resulting balanced distribution after downsampling is depicted in Figure 5, which ensures an equal number of samples for each sentiment class, facilitating more robust and unbiased model training.

a) Downsampling strategy: To address class imbalance in the Coronavirus Tweet NLP data, we performed class-wise *random downsampling* of the four majority classes to match the minority class size. From the raw split (Fig. 4: Positive $= 11{,}422$; Negative $= 9{,}917$; Neutral $= 7{,}713$; Ext_Neg $= 5{,}481$; Ext_Pos $= 6{,}624$), each class was downsampled to $5{,}481$ instances (the minority count), yielding a balanced training set of 27,405 tweets (Fig. 5). Downsampling is applied *only* within the training folds during cross-validation to prevent test-set leakage; the held-out test set preserves the original label proportions.

D. Data preprocessing

Before raw textual data can be effectively used in machine learning models, it must undergo a preprocessing stage to ensure it's clean, consistent, and suitable for analysis. In this system, preprocessing is handled using Natural Language Processing (NLP) techniques [19].

The process begins by converting all text to lowercase to standardize the input. Common stop words—words that carry little meaningful information are then removed using a predefined list from the Python NLTK library. Additionally, contractions are expanded using a custom-built function to maintain clarity in the text.

To further enhance text quality, a spell-checking step is included to minimize the risk of errors due to typos. The pipeline also removes special characters, URLs, and HTML tags to eliminate noise. Once the text is cleaned, it undergoes tokenization, normalization, and lemmatization [24]. These steps are essential in NLP, where techniques like stemming, tokenization, and normalization help structure the raw text for effective classification.

1) Tokenization is a key process in Natural Language Processing (NLP), which consists of segmenting textual data into discrete elements known as tokens. In this framework, each word is typically considered a separate token, facilitating subsequent analysis and computational processing [25].
2) Stemming refers to the technique of reducing words to their base or root form, under the assumption that various morphological variants of a word convey a similar semantic meaning. The resulting stem doesn't necessarily have to be a valid dictionary word. However, all variations of a word should ideally be reduced to the same stem after the process. When applying stemming, it's important to consider two key factors to ensure its effectiveness [17]:
    a) It is generally assumed that the different morphological forms of a word share the same basic meaning, so they should all be reduced to a common stem.

b) It is important to ensure that words with distinct meanings are not mistakenly treated as the same.

These two principles are typically adequate, provided that the generated stems remain effective for the intended text mining and natural language processing applications. In most cases, stemming is used primarily to improve recall by grouping related word forms together. However, its impact tends to be less significant in languages with simpler morphological structures compared to those with more complex morphology.

3) Normalization refers to the process of transforming text into a consistent and standardized form, making it easier to process and analyze.
Sometimes, people use words in unconventional ways to express their ideas [28]. To ensure clarity, such content needs to be standardized, and any spelling mistakes should be corrected.

E. Experimental Environment and Dataset Splitting

All experiments were performed on a machine featuring a 13th-generation Intel Core i5 processor clocked at 2.5 GHz, equipped with 16 GB of RAM, and running the Windows 11 operating system. The experimental evaluations utilized the Coronavirus Tweet NLP Dataset, as described in the Dataset section. The dataset was randomly sampled and shuffled, then divided into training and testing subsets using an 80:20 class distribution ratio. To enhance the robustness and generalizability of the classification models, a 10-fold cross-validation approach was employed during the training process, where the training data was further split into 10 subsets, iteratively using one fold for validation and the remaining folds for model training.

a) Reproducibility statement: We use the Kaggle *Coronavirus tweets NLP—text classification* dataset (see References) with the preprocessing and TF–IDF pipeline described in this paper. Data are split once into 80/20 train/test, with *stratified* 10-fold CV on the training portion. We report a single evaluation on the fixed test set without additional tuning. All PSO/GA runs and model training use fixed random seeds. For replication, we provide the exact GA/PSO settings (Tables I–II), convergence curves (Figs. 4–5), and environment details (CPU/RAM/OS) along with the library defaults used for SVM/KNN/MLP.

b) GA/PSO hyperparameter justification: We chose stable, widely used defaults for wrapper-based feature search. For PSO, we set the inertia and acceleration terms to $w = 0.7298$ and $c_1 = c_2 = 1.496$, with 100 particles and 100 iterations. For GA, we used a population of 100, mutation rate $= 0.2$, crossover rate $=$ 0.9, and 100 generations. These "safe" settings strike a practical balance between exploration and exploitation in high-dimensional binary spaces, and brief pilot sweeps (10–20% variation around each value) did not change conclusions in a meaningful way. During search, the evaluator is a Random Forest to keep the objective fast and robust to noise; after selection, we recompute metrics with the target classifier (SVM, KNN, or MLP) on the chosen subset to avoid overfitting to a single learner's quirks (see Tables I–II, Figs. 6–7).

TABLE I
PSO CONFIGURATION HYPER-PARAMETERS USED IN THE EXPERIMENTS

Nbr of Particles	Nbr of iteration	Inertia weight W	Accel factor C1	Accel factor C2	Eval
100	100	0.7298	1.496	1.496	Random Forest

TABLE II
GA CONFIGURATION HYPER-PARAMETERS USED IN THE EXPERIMENTS

Nbr of Chrom	Nbr of iter	Mut rate	Cross rate	Evaluator
100	100	0.2	0.9	Random Forest

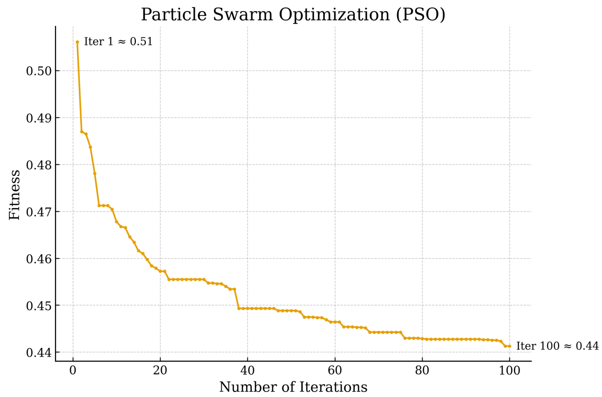

Fig. 6. PSO Convergence

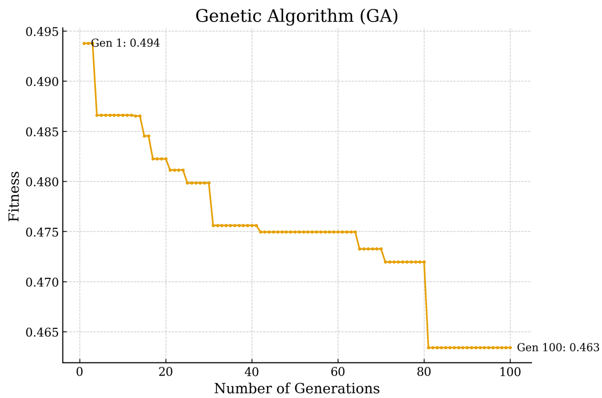

Fig. 7. GA Convergence

IV. Results and Discussions

A. Models performances

TABLE III
PERFORMANCES OF MODELS WITHOUT FEATURE SELECTION

	Accuracy %	Precision %	Recall %	F1-score %
KNN =3	22,17	59,22	22,27	11,59
SVM	57,22	54,99	57,01	55,55
MLP	52,79	53,24	52,76	52,84

The results obtained from this study clearly demonstrate the substantial impact that bio-inspired optimization methods can

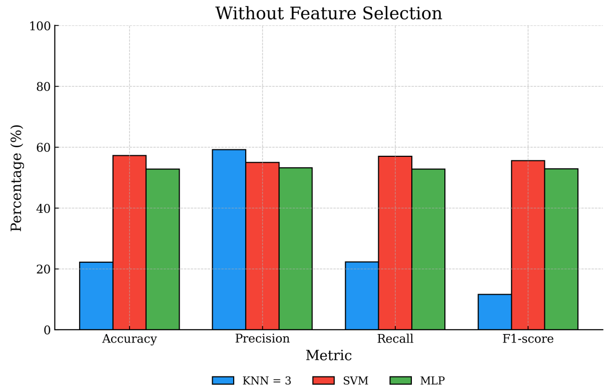

Fig. 8. Performances of models without Feature Selection

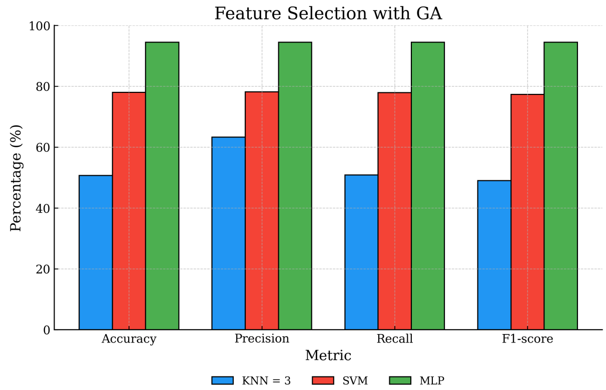

Fig. 10. Performances of models with Feature Selection by GA

TABLE IV
PERFORMANCES OF MODELS WITH FEATURE SELECTION BY PSO

	Accuracy %	Precision %	Recall %	F1-score %
KNN =3	47,91	62,94	48,09	45,46
SVM	77,55	77,75	77,50	76,84
MLP	94,21	94,22	94,21	94,21

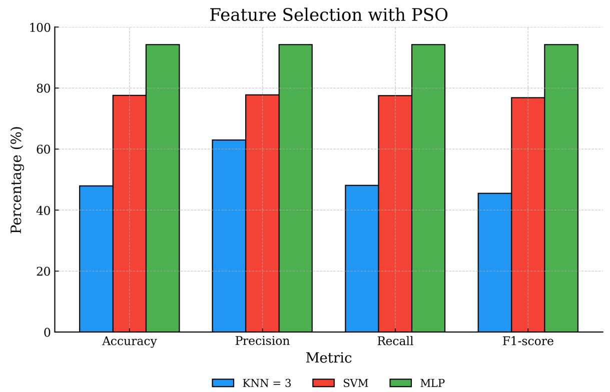

Fig. 9. Performances of models with Feature Selection by PSO

TABLE V
PERFORMANCES OF MODELS WITH FEATURE SELECTION BY GA

	Accuracy %	Precision %	Recall %	F1-score %
KNN =3	50,73	63,29	50,87	49,01
SVM	77,97	78,19	77,93	77,37
MLP	94,48	94,48	94,47	94,48

have on sentiment analysis tasks, particularly when applied to high-dimensional textual data such as COVID-19-related tweets. Without feature selection, all models—especially KNN—struggled to achieve acceptable performance levels. For instance, the KNN classifier yielded an accuracy of only 22.17% and a particularly low F1-score of 11.59%, indicating its inability to manage the noise and sparsity typical of raw textual data. Even the more robust MLP model only achieved 52.79% accuracy in the same setting. However, once Particle Swarm Optimization (PSO) was applied, a marked improvement across all classifiers was observed. Notably, MLP accuracy increased to 94.21%, while SVM and KNN saw boosts to 77.55% and 47.91% respectively. These gains suggest that PSO was effective in selecting a more informative and compact feature subset, thus reducing redundancy and enhancing classification quality. When Genetic Algorithms (GA) were employed, the improvements were even more pronounced in some cases. MLP performance peaked at 94.48% accuracy, with equally high precision, recall, and F1-score—demonstrating the algorithm's ability to find more globally optimal feature sets through evolutionary operations. SVM and KNN also experienced additional performance gains with GA, although marginal compared to PSO. Overall, the comparison between PSO and GA indicates that while both methods are effective, GA offers a slight advantage in fine-tuning feature subsets, albeit with increased computational complexity. Importantly, the consistent superiority of MLP across all conditions confirms the advantage of deep learning-based models in capturing sentiment patterns, especially when guided by an optimized feature space. These results reinforce the argument that evolutionary algorithms, when used as standalone optimization tools, are powerful allies in navigating the complexities of sentiment-rich, high-dimensional data, particularly during fast-evolving public health events like the COVID-19 pandemic.

B. Reduction rate

An important aspect of this study lies in the significant reduction of the feature space achieved through the use of bio-inspired optimization algorithms. Specifically, Particle Swarm Optimization (PSO) resulted in a feature reduction rate of 50.05%, while Genetic Algorithms (GA) achieved a slightly higher reduction of 50.32%. These results indicate that both methods were able to effectively eliminate approximately half of the original features without compromising—and indeed while improving—the performance of the classifiers. This substantial dimensionality reduction contributes not only to increased classification accuracy, as demonstrated in previous sections, but also to improved computational efficiency during model training and inference. Moreover, reducing the feature set by such a degree helps mitigate overfitting and enhances the generalizability of the models, particularly when dealing with large, noisy, and redundant datasets such as COVID-19-related tweets. The near-equivalence of the reduction rates between PSO and GA also highlights the consistency of these algorithms in identifying compact, high-quality feature subsets, reinforcing their applicability in high-dimensional natural language processing tasks.

V. Validation

Compared with Maaskri *et al.* [2], who evaluated five–class COVID-19 tweet sentiment using classical ensembles and a BERT baseline (macro $F_1 \approx 0.74$), our findings show that *feature selection is decisive*: starting from comparable single–model baselines without selection (e.g., SVM ≈ 0.56 in our setup), applying PSO or GA feature selection yields large, consistent gains for classical learners—most notably MLP ($F_1 \approx 94\%$) and SVM ($F_1 \approx 77\%$). These results surpass both the ensemble ML scores ($F_1 \approx 0.65$) and the reported BERT baseline, indicating that targeted search–based selection on high–dimensional tweet features can recover a compact, discriminative subset that classical models exploit effectively. While exact parity of splits and preprocessing prevents a one–to–one replication, the ranking agreement (KNN $\ll$ SVM $<$ MLP $\ll$ MLP + FS) across studies supports the external validity of our conclusions.

VI. Conclusion

In conclusion, this study presents a robust and efficient framework for sentiment analysis of COVID-19-related social media content, relying solely on bio-inspired optimization techniques for feature selection. By integrating Genetic Algorithms and Particle Swarm Optimization with traditional machine learning classifiers, the proposed approach effectively reduces feature dimensionality while significantly enhancing classification performance. The experimental results, particularly with the MLP model, demonstrate that optimizing the feature space without the need for conventional statistical filtering leads to substantial gains in accuracy, precision, recall, and F1-score. These findings underscore the critical role of intelligent feature selection in dealing with noisy, large-scale textual data, as often encountered in real-time crisis communication scenarios. Moreover, the comparative evaluation between GA and PSO highlights their respective strengths—GA offering slightly higher predictive accuracy, and PSO delivering faster convergence. This work not only contributes a practical method for improving sentiment classification under urgent and dynamic conditions but also opens avenues for future exploration, such as combining these optimization techniques with contextual language models or applying them to multilingual datasets. Ultimately, the study affirms that evolutionary computing remains a promising and adaptable strategy for enhancing natural language processing tasks in complex, real-world applications.

References

[1] T. Alqurashi, “Arabic Sentiment Analysis for Twitter Data: A Systematic Literature Review,” *Engineering, Technology & Applied Science Research*, vol. 13, no. 2, pp. 10292–10300, Apr. 2023, doi:https://doi.org/10.48084/etasr.5662

[2] M. Moustafa, S. A. Mokhtar-Mostefaoui, M. Hadj-Meghazi, and M. Goismi, ”Multi-class sentiment analysis of COVID-19 tweets by machine learning and deep learning approaches,” *Computación y Sistemas*, vol. 28, no. 2, pp. 507–516, 2024, doi:https://doi.org/10.13053/CyS-28-2-4568

[3] M. T. Ribeiro, S. Singh, and C. Guestrin, “Why Should I Trust You? Explaining the Predictions of Any Classifier,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 1135–1144, 2016. doi:https://doi.org/10.1145/2939672.2939778

[4] N. Hayatin, G. I. Marthasari, and L. Nuarini, “Optimization of sentiment analysis for Indonesian presidential election using Naïve Bayes and particle swarm optimization,” *Jurnal Online Informatika*, vol. 5, no. 1, pp. 81–88, 2020. doi:https://doi.org/10.15575/join.v5i1.558

[5] M. Grootendorst, “BERTopic: Neural topic modeling with a class-based TF–IDF procedure,” *arXiv preprint* arXiv:2203.05794, 2022. doi:https://doi.org/10.48550/arXiv.2203.05794

[6] S. Sangam and S. Shinde, “A novel feature selection method based on genetic algorithm for opinion mining of social media reviews,” in *Information, Communication and Computing Technology*, Singapore: Springer, pp. 167–175, 2019. doi:https://doi.org/10.1007/978-981-13-5992-7_15

[7] P. Pookduang, R. Klangbunrueang, and W. Chansanam, “Advancing sentiment analysis: Evaluating RoBERTa against traditional and deep learning models,” *Engineering, Technology & Applied Science Research*, vol. 15, no. 1, pp. 20167–20174, Feb. 2025. doi:https://doi.org/10.48084/etasr.9703

[8] G. R. Kishore, B. S. Harish, and C. K. Roopa, “Unenhanced sparse vector-based embedding method for sentiment analysis,” *Engineering, Technology & Applied Science Research*, vol. 15, no. 2, pp. 21225–21231, Apr. 2025. doi:https://doi.org/10.48084/etasr.10098

[9] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, pp. 4171–4186, 2019. doi:https://doi.org/10.18653/v1/N19-1423.

[10] H. E. Kizilöz, A. Deniz, T. Dökeroglu, and A. Coşar, “Novel multiobjective TLBO algorithms for the feature subset selection problem,” *Neurocomputing*, vol. 306, pp. 94–107, 2018. doi:https://doi.org/10.1016/j.neucom.2018.04.020

[11] M. Allam and M. Nandhini, “Optimal feature selection using binary teaching learning based optimization algorithm,” *Journal of King Saud University – Computer and Information Sciences*, vol. 34, no. 2, pp. 329–341, Feb. 2022. doi:https://doi.org/10.1016/j.jksuci.2018.12.001

[12] P. Rani, J. Shokeen, A. Majithia, A. Agarwal, A. Bhatghare, and J. Malhotra, “Designing an LSTM and Genetic Algorithm-based Sentiment Analysis Model for COVID-19,” in *Proceedings of Data Analytics and Management* (Lecture Notes on Data Engineering and Communications Technologies, vol. 91), Singapore: Springer, pp. 209–216, 2022. doi:https://doi.org/10.1007/978-981-16-6285-0_17

[13] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, “SMOTE: Synthetic minority over-sampling technique,” *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002. doi:https://doi.org/10.1613/jair.953

[14] R. Tibshirani, “Regression shrinkage and selection via the LASSO,” *Journal of the Royal Statistical Society: Series B*, vol. 58, no. 1, pp. 267–288, 1996. doi:https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

[15] Y. Yue, Y. Peng, and D. Wang, “Deep Learning Short Text Sentiment Analysis Based on Improved Particle Swarm Optimization,” *Electronics*, vol. 12, no. 19, article 4119, 2023. doi:https://doi.org/10.3390/electronics12194119

[16] C. Sitaula and T. B. Shahi, “A hybrid feature extraction method for Nepali COVID-19-related tweets,” *Computational Intelligence and Neuroscience*, vol. 2022, article 5681574, Mar. 2022. doi:https://doi.org/10.1155/2022/5681574

[17] A. Miglani, ”Coronavirus tweets NLP - text classification,” 2020. [Online]. Available: https://www.kaggle.com/datasets/datatattle/covid-19-nlp-text-classification. Accessed on: Jun. 13, 2025.

[18] A. Blum and P. Langley, “Selection of relevant features and examples in machine learning,” *Artificial Intelligence*, vol. 97, no. 1–2, pp. 245–271, 1997. doi:https://doi.org/10.1016/S0004-3702(97)00063-5

[19] L. Nemes and A. Kiss, ”Social media sentiment analysis based on COVID-19,” *Journal of Information and Telecommunication*, vol. 5, no. 1, pp. 1–15, 2020, doi: 10.1080/24751839.2020.1790793.

[20] S. Kumar, M. B. Khan, M. H. Abul Hasanat, A. K. J. Saudagar, A. AlTameem, and M. AlKhathami, “Sigmoidal particle swarm optimization for Twitter sentiment analysis,” *Computers,*

Materials & Continua, vol. 74, no. 1, pp. 897–914, 2023. doi:https://doi.org/10.32604/cmc.2023.031867

[21] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient Estimation of Word Representations in Vector Space," in *Proc. ICLR (Workshop Track)*, 2013. doi:https://doi.org/10.48550/arXiv.1301.3781

[22] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Reading, MA: Addison-Wesley, 1989.

[23] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of the IEEE International Conference on Neural Networks (ICNN'95)*, Perth, Australia, 1995, pp. 1942–1948. doi:https://doi.org/10.1109/ICNN.1995.488968

[24] F. M. Javed-Mehedi-Shamrat, Z. Tasnim, P. Ghosh, A. Majumder, and M. Z. Hasan, "Personalization of job circular announcement to applicants using decision tree classification algorithm," in *Proc. IEEE Int. Conf. Innovation Technology (INOCON)*, Piscataway, NJ, USA: IEEE, 2020, pp. 1–5, doi: 10.1109/INOCON50539.2020.9298253.

[25] H. Ghulam, F. Zeng, W. Li, and Y. Xiao, "Deep learning-based sentiment analysis for Roman Urdu text," *Procedia Computer Science*, vol. 147, pp. 131–135, 2019, doi: 10.1016/j.procs.2019.01.202.

[26] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *Journal of Machine Learning Research*, vol. 3, pp. 1157–1182, 2003.

[27] M. Müller, M. Salathé, and P. E. Kummervold, "COVID-Twitter-BERT: A natural language processing model to analyse COVID-19 content on Twitter," *Frontiers in Artificial Intelligence*, vol. 6, article 1023281, 2023. doi:https://doi.org/10.3389/frai.2023.1023281

[28] R. P. Mehta, M. A. Sanghvi, D. K. Shah, and A. Singh, "Sentiment analysis of tweets using supervised learning algorithms," in *First International Conference on Sustainable Technologies for Computational Intelligence*, (AISC, vol. 1045) Singapore: Springer, pp. 323–338, 2020. doi:https://doi.org/10.1007/978-981-15-0029-9_26

[29] T. B. Shahi, C. Sitaula, and N. Paudel, "A hybrid feature extraction method for Nepali COVID-19-related tweets classification," *Computational Intelligence and Neuroscience*, vol. 2022, no. 1, Art. 5681574, Mar. 2022. doi:https://doi.org/10.1155/2022/5681574

[30] M. Goismi, R. M. Hamou, A. Toumouh, and M. Moustafa, "Enhancing Twitter sentiment classification with a hybrid bio-inspired feature selection approach," *Journal of Information Systems Engineering & Management*, vol. 10, no. 53s, pp. 74–89, 2025, doi: 10.52783/jisem.v10i53s.10840.

[31] S. Kumar, M. B. Khan, M. H. Abul Hasanat, A. K. Jilani Saudagar, A. AlTameem, and M. AlKhathami, "Sigmoidal particle swarm optimization for Twitter sentiment analysis," *Computers, Materials & Continua*, vol. 74, no. 1, pp. 897–914, 2023, doi: 10.32604/cmc.2023.031867.

[32] M. A. Syed, E. Arsevska, M. Roche, and M. Teisseire, "Feature selection for sentiment classification of COVID-19 tweets: H-TFIDF featuring BERT," in *Proc. 15th Int. Joint Conf. Biomedical Engineering Systems and Technologies (BIOSTEC 2022) – HEALTHINF*, SciTePress, 2022, pp. 648–656, doi: 10.5220/0010887800003123.

[33] G. E. A. P. Batista, R. C. Prati, and M. C. Monard, "A study of the behavior of several methods for balancing machine learning training data," *ACM SIGKDD Explorations Newsletter*, vol. 6, no. 1, pp. 20–29, 2004. doi:https://doi.org/10.1145/1007730.1007735

[34] J. H. Holland, *"Adaptation in Natural and Artificial Systems,"* Ann Arbor, MI, USA: University of Michigan Press, 1975.

Mohamed Debbab received his engineering degree in Computer Science from the University of Sciences and Technology of Oran (USTO), Algeria, in 2007, and his M.S. degree in Computer Science from the University of Sciences and Technology Houari Boumediene (USTHB), Algiers, Algeria, in 2010. He obtained his Ph.D in Computer Science from the University of Sidi Bel Abbès (UDL) in June 2025. He is currently working as an Associate Professor in the Department of Science and Technology at Ibn Khaldoun University of Tiaret. His research interests include Natural Language Processing (NLP), machine learning, misinformation detection, and sentiment analysis.

Moustafa Maaskri is currently working as an Associate Professor in Electrical Engineering Department at Ibn Khaldoun University of Tiaret, Algeria. He obtained his PhD in Computer Science from the University of Sidi Bel Abbès in 2022 and his Engineering degree in Computer Science from Ibn Khaldoun University, Algeria in 2006. His research interests include Natural Language Processing (NLP), machine learning, misinformation detection and sentiment analysis.

Mohamed Goismi is currently working as an Associate Professor in the Department of Science and Technology at Ibn khaldoun University of Tiaret, Algeria. He obtained his PhD in Computer Science from the University of Sidi bel abbes in 2025 and his Engineering degree in Computer Science from Ibn khaldoun University, Algeria, in 2006. His research interests are focused on Natural Language Processing (NLP) and include machine learning, Misinformation detection and sentiment analysis.

Abdelkader Boudaoud is currently working as an Associate Professor in Computer Science Department at University Center of El Bayadh, Algeria. He obtained his PhD in Computer Science from the University of Sidi bel abbes in 2025 and his Engineering degree in Computer Science from USTO University, Algeria, in 2007. His research interests lie primarily in the fields of cryptography and steganography, with a particular emphasis on secure communication, data confidentiality, and advanced techniques for information hiding and protection.

Djamel Seghier is a first-year PhD student in Computer Science at Ibn Khaldoun University in Tiaret, Algeria. He commenced his doctoral studies in 2025, focusing on Natural Language Processing (NLP), machine learning, misinformation detection and sentiment analysis.