SSRG International Journal of Electrical and Electronics Engineering
ISSN: 2348-8379/ https://doi.org/10.14445/23488379/IJEEE-V12I6P115

Volume 12 Issue 6, 180-186, June 2025

Original Article

Integrating Swarm Intelligence with Deep Learning for Enhanced Social Media Sentiment Analysis

Parminder Singh[1], Saurabh Dhyani[2]

[1,2]Uttaranchal School of Computing Sciences, Uttaranchal University, Dehradun, Uttarakhand, India

[1]Corresponding Author : digiageinfo@gmail.com

Received: 11 April 2025 Revised: 13 May 2025 Accepted: 12 June 2025 Published: 30 June 2025

Abstract *- In our digital world, it is vital to know the opinion of the masses on social network sites. The study has presented a new framework to merge the advantages of deep learning and swarm intelligence, PSAM (Particle Swarm-Accelerated Model). It combines Long Short-Term Memory (LSTM) to perform a good sentiment classification and Particle Swarm Optimization (PSO) to perform good feature selection and hyperparameter optimization. The model has been used to label the YouTube reviews of the film Housefull 5 by classifying these sentiments as positive, negative or neutral with an impressive accuracy of 95.2 percent. The sentiment analysis pipeline starts with the extraction of comments through the YouTube API, and their pre-processing consists of removing punctuation, removing low-frequency words, normalising colloquial vocabulary, and emoji analysis. PSO has been essential in ensuring that the relevance of subsets has been determined and that feature reduction enhances the performance of the LSTM model tremendously. Comparing it with the conventional algorithms like Naïve Bayes, SVM, CNN and Random Forest, PSAM gives desirable results in all the major criteria of the build, like the accuracy and F1-score. The hybrid approach is very competent in sentiment analysis and has great potential to be extended to real-time systems and cross-platform social media mining.*

Keywords *- PSAM model, Deep Learning, Sentiment classification, Hybrid optimization, Social media mining, LSTM-PSO framework.*

1. Introduction

There has been a monumental growth in the user-generated content on websites such as Facebook, Twitter and YouTube, which has caused an increase in the application of sophisticated sentiment analysis methods. Sentiment classification can be more difficult because these platforms tend to display the users' emotionally lively expressions. As a subfield of Natural Language Processing (NLP), sentiment analysis is useful in identifying and deriving the feelings, opinions, and attitudes in text data, giving an insight into user preferences and how users perceive things [1, 2]. This power is critical in all industries, particularly marketing, policymaking, and online content curation, as an insight into the audience's feelings may directly affect strategy-making. Nevertheless, there is a tendency to model sentiment incorrectly because conventional machine learning techniques like Support Vector Machines (SVM), Naïve Bayes, and Random Forest fail when correctly inferring sentiment. Their weaknesses come due to complications in understanding the surrounding contexts, dealing with sarcastic utterances and/ the only way to overcome linguistic ambiguity is to classify them, most of the time incorrectly [3, 4]. These issues emphasise the need for more intuitive, context-sensitive models that would be able to accommodate the complexity of human language as presented within the context of a contemporary digital platform. Recently, deep learning methods, especially the use of Long Short-Term Memory (LSTM) networks, have met these deficiencies in the traditional sentiment analysis concerning exploiting the long-term dependency in opinionated texts [5, 6]. Nevertheless, LSTM is quite sensitive to the input feature of swarm intelligence. To assist, the Particle Swarm Optimisation (PSO) algorithm selects the most pertinent features, down dimensions, and improves the classification [7-9]. This paper proposes Particle Swarm-Accelerated Model (PSAM), an ensemble of PSO to find feature selection and LSTM to accomplish sentiment classification, which provides a mixture of techniques capable of solving the drawbacks of either deep learning or machine learning approaches. The model PSAM was also tested with the comments of a movie called Pushpa 2 on YouTube and could successfully label the sentiments as positive, neutral or negative replies [10, 11]. The findings indicate that the model is solid and shows an excellent outcome in processing user-generated content.

2. Literature Review

As more profound deep learning and complex feature selection data emerged, Sentiment Analysis (SA) has advanced amazingly. Though simple methods of the past, including lexicon-based models and well-known classifiers such as Naïve Bayes and Support Vector Machine (SVM), were more popular because of ease of implementation, they perform poorly when used. The traditional models are challenging to deal with complex natural language elements, such as sarcasm, vague terms, and contextual meaning, which decreases the overall accuracy of sentiment recognition tasks in a real scenario [12, 13].

In recent studies, the power of the deep learning model is emphasized, such as the "Long Short-Term Memory" (LSTM) model, which is effective in modelling long-range dependencies and complicated contextual dependencies in the text. Consequently, LSTM-like neural networks perform better than conventional methods in sentiment classification tasks [14, 15].

Based on the movement of fish and bird flocks, Particle Swarm Optimization (PSO) has become one of the most popular methods for finding an optimal feature set. It improves classification accuracy and reduces dimensionality by guiding particles using individual and global best experiences [16, 17]. Numerous studies confirm PSO's ability to enhance the performance of classifiers like SVM and Neural Networks [18, 19]. Conversely, the LSTM networks are well-suited to capturing long-range text dependency; thus, they lend themselves to sentiment analysis tasks [20].

However, limited research has investigated the integration of PSO with LSTM to optimise feature selection and sentiment classification. This study addresses that gap by introducing the PSAM hybrid model, which combines the strengths of PSO and LSTM. Experimental comparisons with traditional classifiers such as SVM, Naïve Bayes, and Random Forest highlight PSAM's superior performance and its potential for scalable sentiment analysis applications across social media platforms [21-23].

3. Proposed Hybrid Model

The Particle Swarm-Accelerated Model (PSAM) incorporates PSO to select features and LSTM to categorise sentiments, improving performance and efficiency. The model gives better sentiment predictions by decreasing the dimensionality of inputs. It starts with good text preprocessing, such as tokenisation, cleaning and normalising. It is an approach that couples the potential of deep learning and the lexicon-based methods to enhance sentiment detection.

PSO processes the features with the highest relevance and redundancy, making the LSTM process classify sentiments, either positive, negative, or neutral, and it does not affect the meaning since it retains context. The hybrid Framework is well-suited to combine swarm intelligence and deep learning to perform strong sentiment analysis.

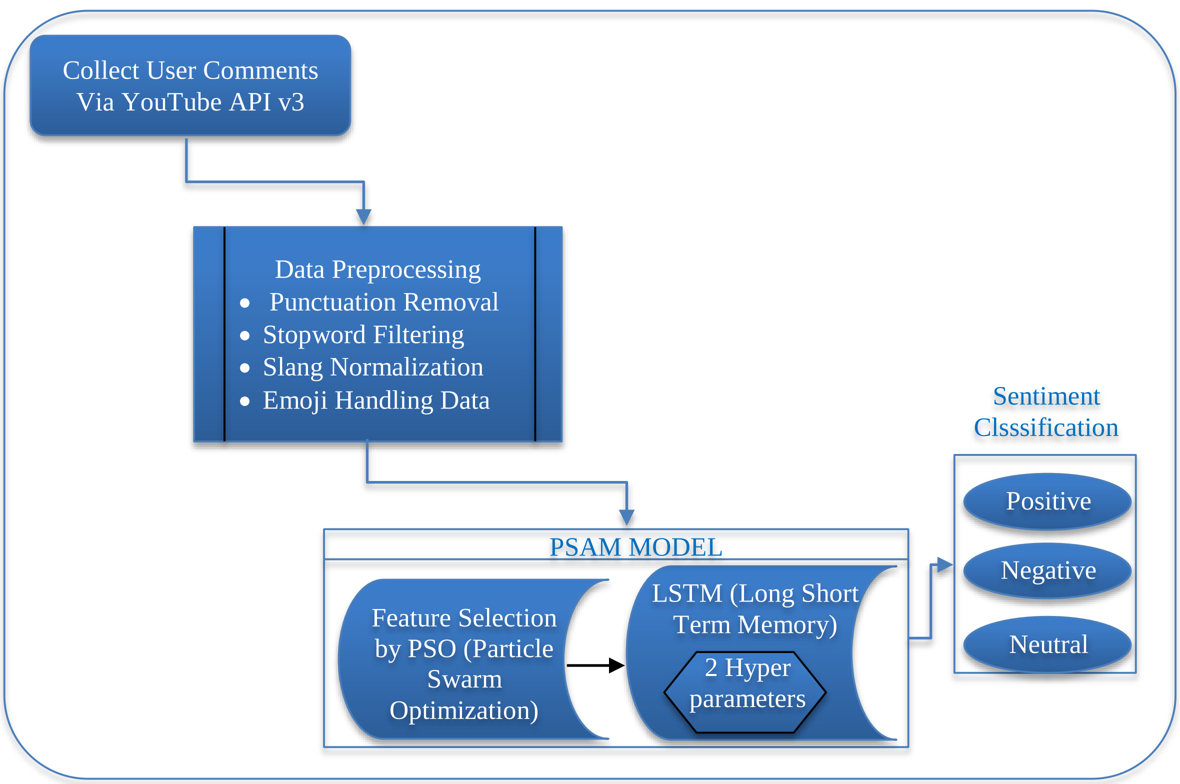

Fig. 1 The PSAM model architecture combines PSO and LSTM to select features and sentiment classification, respectively

3.1. PSAM Workflow Design

In this research, the complete methodological Framework follows some steps, which are as follows:

3.1.1. Data Collection

In this study, it is important to collect user-generated comments connected to the movie Pushpa-2 using the YouTube Data API as the first step.

The relevant content, including trailers, reviews, and discussion videos, was found through the search. List API. The mechanism behind extracting the comments using the commentThreads. The list API is provided and shown in Figure 2 as the API mechanism's request and response API. With this process, several attributes such as comment text, the id of the user, date, like count, and replies are retrieved. The data set includes various emotional expressions, negative, positive, and neutral states, showing viewers' wide range of opinions. These comments are a good alternative to sentiment analysis since these are informal comments that usually consist of slang, misspellings, use of emojis, and phrases, as well as sarcasm. The complexity of the data set renders it both problematic and useful in deriving superior sentiment classification courses [24, 25].

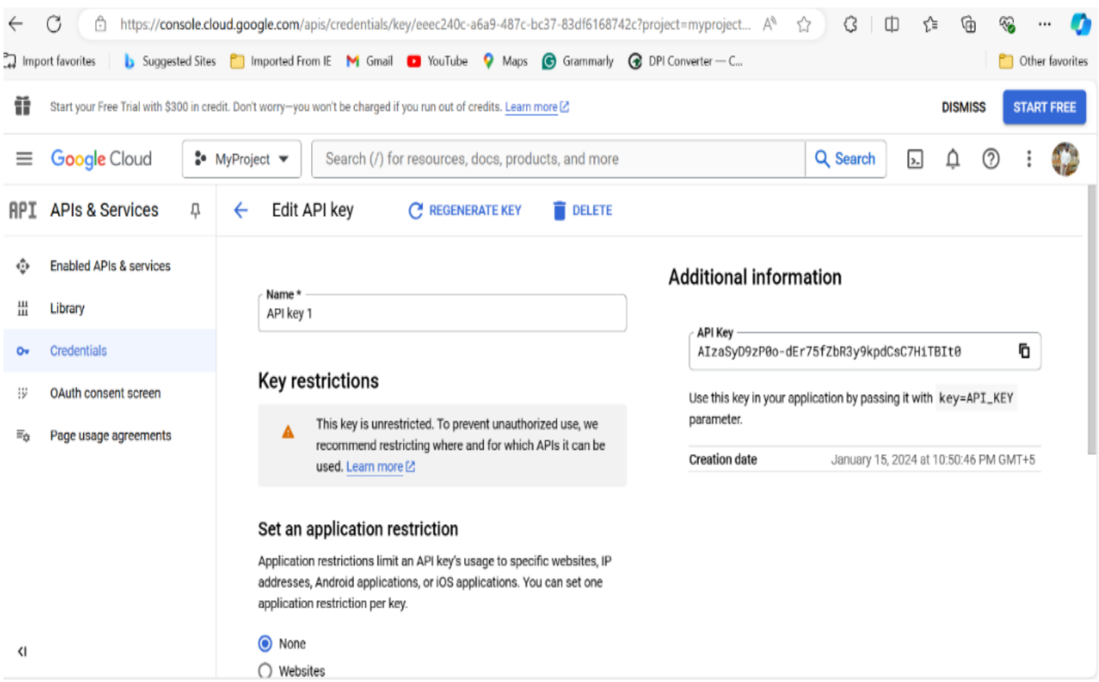

Fig. 2 API key setup via google cloud

3.1.2. Data Acquisition Process

The preprocessing stage is concerned with enhancing the quality of the data by text cleaning (such as standardizing input and tokenizing it). URLs, emoticons, and special characters are stripped by the Regular Expressions (RegEx), resulting in cleaner input.

To analyze text, the SpaCy tokenizer divides it into individual words. Then the stopwords, such as the, are eliminated with the NLTK corpus [26], and the lemmatization, using WordNetLemmatizer, cuts words to stem form, thus increasing sentiment classification accuracy.

Table 1 illustrates the frequently occurring keywords extracted from movie reviews, highlighting their emotional relevance. This table is instrumental in training sentiment analysis models, as it quantifies significant word usage, aiding in accurately identifying sentiments expressed within the reviews.

Table 1. Word cloud frequency

Word	Count
Allu	890
puspa	400
arjun	240
wild	125
fire	110
super	91
fight	82
bad	15
boring	6
---	--

3.1.3. Feature Selection

To increase classification accuracy, feature subsets are optimized using the Particle Swarm Optimization (PSO) algorithm. In this algorithm, every particle indicates a feasible conjunction of features of the model, and its fitness is determined based on the aspect of model performance in classifications. The motion of the particles through the solution space is modelled as an update of the position and the velocity based on specified equations, as shown in Figure 3. It is an iterative update scheme that relies on the best-known location of the particle (personal best) and the best-known location (global best). The update rule has Formula-1 as its basis, whereby the dynamics of updating the movement of particles is in a way that it converges to an optimal set of features in the feature space that increases the overall predictive power of the model.

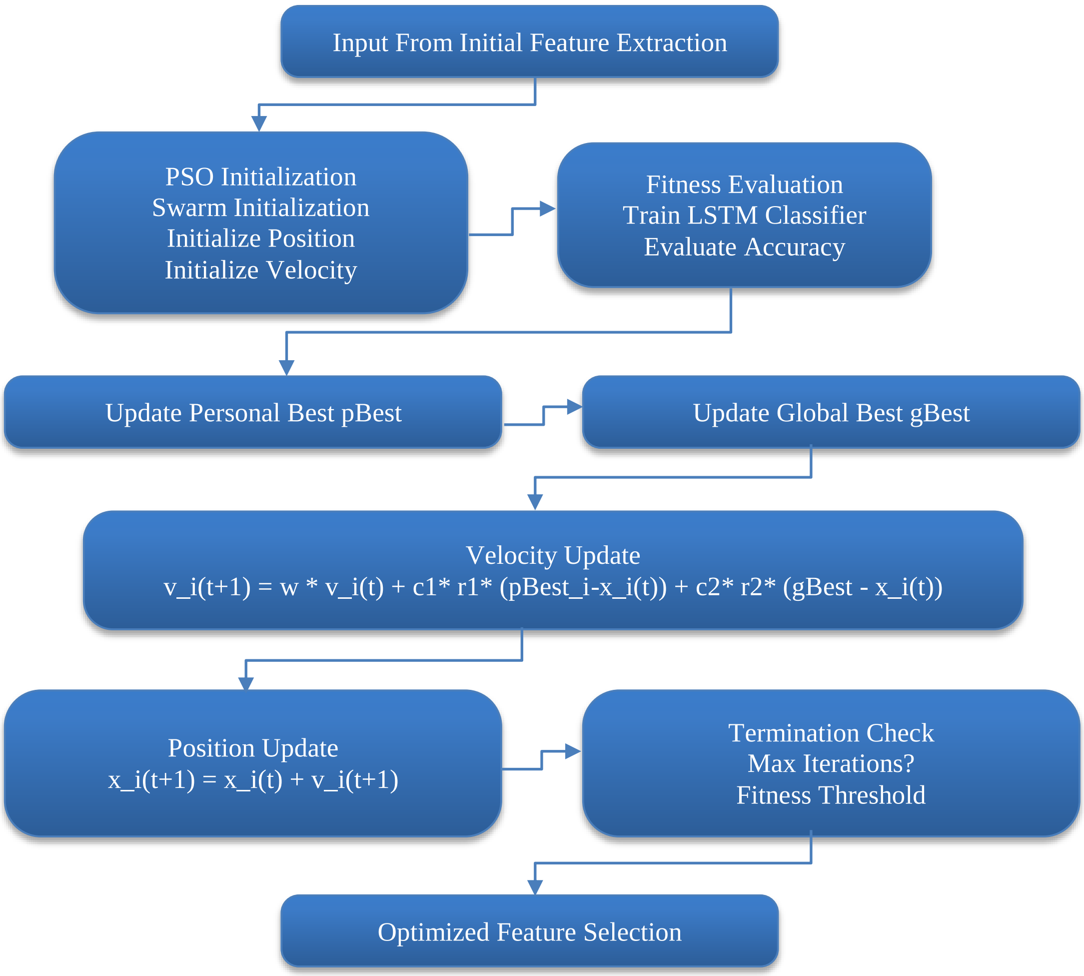

Fig. 3 PSO-based feature selection framework

In the PSO-based feature selection process, each particle represents a potential subset of features. The performance or fitness of a particle is assessed by training an LSTM classifier on the selected features and evaluating it on a validation dataset. A particle updates its personal best position (pBest) when it attains a better classification accuracy.

In the same way, when it achieves higher performance than gBest, which is the best at the moment throughout the swarm, the global best is correspondingly also updated [8, 18]. As particles navigate the feature space, they iteratively update their positions by learning from their experience (pBest) and the swarm's best-known solution (gBest).

This dynamic ensures a balance between exploration and exploitation, leading to effective feature subset optimization that enhances classification accuracy while minimizing dimensionality.

(Formula 1), which controls how particles explore the search space:

$$v_i(t+1) = w \cdot v_i(t) + c_1 \cdot r_1 \cdot (pBest_i - x_i(t)) + c_2 \cdot r_2 \cdot (gBest - x_i(t)) \quad (1)$$

Where:

- v i(t): Velocity of the ith particle at iteration t
- x i (t): Present location (position) of particle i at time t
- w: Weight of inertia governing the effect of previous velocity
- c_1, c_2: Learning factors or acceleration constants guiding personal and global learning

- r_1, r_2: Random values between 0 and 1 introduced to maintain diversity in the swarm
- $pBest_i$: The best solution or position reached by particle i so far
- gBest: The general optimal state of any of the particles in the swarm that has been attained so far in the iteration

Each particle updates its position based on the velocity computed in the previous step to explore new potential feature subsets. This adjustment allows particles to move through the search space toward more optimal solutions. The following equation governs the position update, referred to as Formula 2:

Formula–2 (Position Update Equation):

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (2)$$

Through this iterative process of velocity and position updates, the PSO algorithm effectively searches for the optimal feature subset that maximizes the classification performance of the sentiment analysis model.

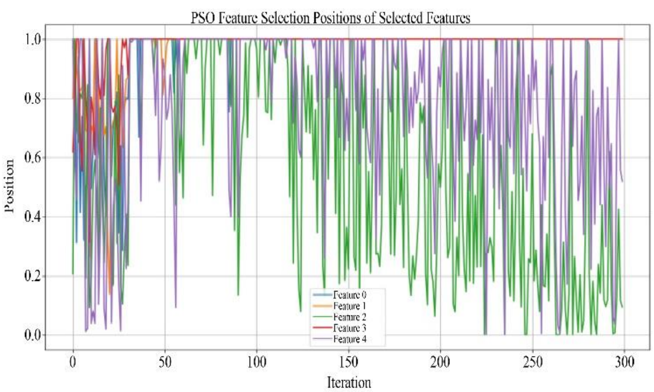

Fig. 4 Feature selection by PSO

Figure 4 indicates the Feature Selection step that incorporates the PSO algorithm. The x-axis displays the position of the particles (0 to 300), and the y-axis displays the fitness values related to classification accuracy (0.0 to 1.0). Particles check characteristics (Feature 0 to 4) and refresh their costs with individual best (pBest) and global best (gBest). PSO finds the best feature subset, improves performance, and decreases redundancy with iterative updates.

3.1.4. LSTM Classifier: Architecture and Functionality

The Long-Short term memory (LSTM) model is most suitable for sentiment analysis because such a model captures long-range dependencies and prevents the Vanishing gradient problem in RNNs. LSTM is used in this research to work on optimised features using Particle Swarm Optimization (PSO). Architecture entails embedding layer, LSTM layers, a dropout layer to discourage learning of overfitting, and a dense layer on output to classify the sentiments as positive, negative, and neutral. Hyperparameters of vital importance, hidden unit size and learning rate, are optimized to enhance prediction performance. Such additions show better accuracy, precision, recall, and F1-score (Formula 3) when using the Adam optimizer and categorical cross-entropy loss [5, 14].

$$F1=2\times Precision+RecallPrecision\times Recall \quad (3)$$

3.1.5. Model Evaluation

In order to prove the effectiveness of the PSAM model, such calculations as precision, recall, and F1-score were employed.. These ensured accurate sentiment detection across positive, negative, and neutral classes. A Comment Clustering graph (Figure 5), generated using K-means and visualized via PCA, reveals distinct sentiment-based clusters. Each shade represents a different group of similar comments, showcasing the model's ability to capture sentiment patterns effectively.

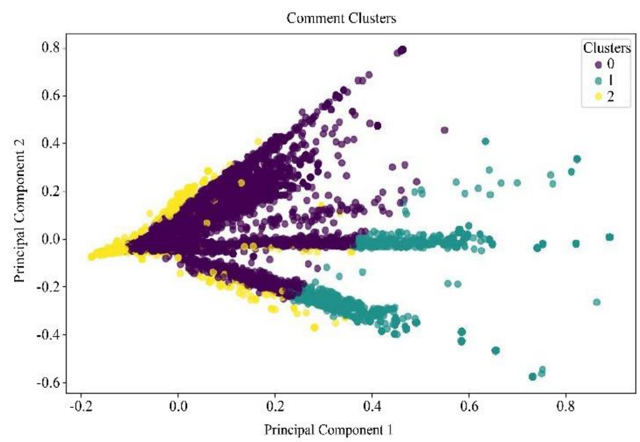

Fig. 5 Comment clusters

4. Result and Discussion

Analysis of YouTube comments has been displayed through bar graphs (Figure 6(a)) and word clouds (Figure 6(b)). The bar chart indicates negative, neutral and positive sentiment distribution, offering a cut-and-dried view of the audience's emotions.

The word cloud gives prominent words such as excellent, love, awesome, and the reference to Housefull 5, and it gives us a clue about the emotional expression in the commentary.

As the pie chart in Figure 6 (a) demonstrates, 70.2 percent of users have positive sentiment, 10.3 percent were negative, and 19.5 percent were neutral. These images and the word cloud illustrated in Figure 6(b) give quantitative and qualitative details about audience views, as indicated in the YouTube comments. PSAM model also exhibits its robustness in Table 2 with 95.3 percent accuracy, surpassing the traditional classifiers, such as Random Forest, SVM, and Naïve Bayes and therefore the strong ability to deal with sentiment analysis tasks.

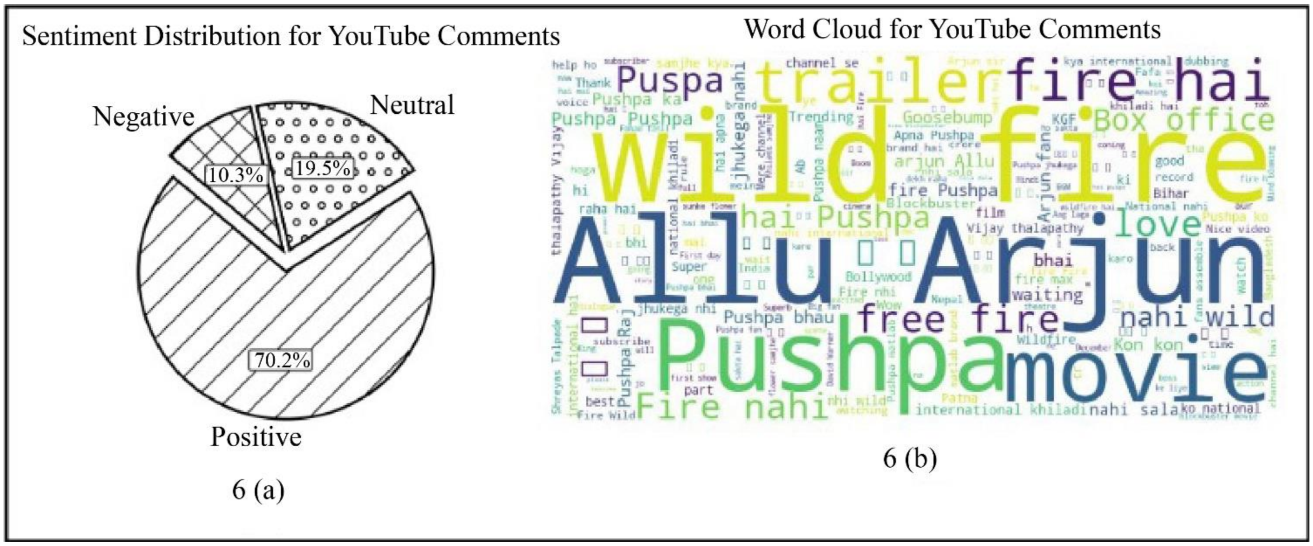

Fig. 6(a) Pie chart of sentiments, and (b)Word cloud representation.

Table 2. PSAM vs. Standard sentiment classification techniques

“Model”	“Precision”	“Recall”	“Accuracy/Result”	“F1-Score”
“Naïve Bayes”	81.9	80.4	83.0%	0.78
“SVM”	88.9	86.4	90.3%	0.80
“CNN”	83.6	84.1	85.2%	0.84
“Proposed Model”	93.7	94.4	95.2%	0.92

5. Conclusion

The specified research suggests the application of PSAM, that is, a hybrid solution involving Particle Swarm Optimization (PSO) and “Long Short-Term Memory (LSTM)” networks to retrieve the best features and examine the sentiment accordingly. The model produces high-quality sentiment analysis results by optimizing feature subsets via PSO and narrow instances of important LSTM parameters, such as the number of hidden units and learning rates. PSAM exceeds the classic algorithms, including Naïve Bayes, SVM and CNN, with an impressive accuracy of 95.3%. Dimensionality reduction based on the PSO makes classification easy and increases accuracy. Robust results observed on all the evaluation measures (precision, recall, F1-score, and accuracy) point to the strength of the model. Proposed extensions include real-time tracking of sentiment, multimediatization (text, audio, video) and more complex neural architectures to improve the performance across larger databases.

Acknowledgments

I wish to acknowledge and thank Dr. Saurabh Dhyani, as my mentor under whom I have learned a lot, who has guided me to make many contributions, and who was of utmost help to me, time and again, as the first author of the work. I also wish to honor the academic society and academic resources provided by Uttaranchal University, which helped to complete the research successfully. We also appreciate the helpful suggestions of our mentors and colleagues, and the constructive criticism and discussion of the manuscript with our colleagues contributed greatly to the clarity and effectiveness of the manuscript.

References

[1] Nhan Cach Dang, María N. Moreno-García, and Fernando De la Prieta, “Sentiment Analysis Based on Deep Learning: A Comparative Study,” *Electronics*, vol. 9, no. 3, pp. 1-29, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[2] Oscar Araque et al., “Enhancing Deep Learning Sentiment Analysis with Ensemble Techniques in Social Applications,” *Expert Systems with Applications*, vol. 77, pp. 236-246, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[3] Chenggang Li et al., “Impacts and Pathways of the Belt and Road Initiative on Sustainable Development Goals of the Involved Countries,” *Sustainable Development*, vol. 32, no. 4, pp. 2976-2989, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[4] Juana Du, Nadeem Akhtar, and Yulei Dou, “Editorial: Towards 2030: Sustainable Development Goal 9: Industry, Innovation and Infrastructure. A Communication Perspective,” *Frontiers in Communication*, vol. 8, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[5] Arash Hajikhani, and Arho Suominen, “Mapping the Sustainable Development Goals (SDGs) in Science, Technology and Innovation: Application of Machine Learning in SDG-Oriented Artefact Detection,” *Scientometrics*, vol. 127, no. 11, pp. 6661-6693, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[6] Md. Shofiqul Islam et al., "Challenges and Future in Deep Learning for Sentiment Analysis: a Comprehensive Review and a Proposed Novel Hybrid Approach," *Artificial Intelligence Review*, vol. 57, no. 3, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[7] Joanne Westwood, *Social Media in Social Work Education*, 1st ed., London: Routledge, pp. 1-104, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[8] Marouane Birjali, Mohammed Kasri, and Abderrahim Beni-Hssane, "A Comprehensive Survey on Sentiment Analysis: Approaches, Challenges and Trends," *Knowledge-Based Systems*, vol. 226, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[9] Jose Ramon Saura, Ana Reyes-Menendez, and Stephen B. Thomas, "Gaining a Deeper Understanding of Nutrition Using Social Networks and User-Generated Content," *Internet Interventions*, vol. 20, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[10] Yeamduan Narangajavana Kaosiri et al., "User-Generated Content Sources in Social Media: A New Approach to Explore Tourist Satisfaction," *Journal of Travel Research*, vol. 58, no. 2, pp. 253-265, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[11] Anisha P Rodrigues et al., "Real-Time Twitter Spam Detection and Sentiment Analysis using Machine Learning and Deep Learning Techniques," *Computational Intelligence and Neuroscience*, vol. 2022, no. 1, pp. 1-14, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[12] Barrie Gunter, Nelya Koteyko, and Dimitrinka Atanasova, "Sentiment Analysis: A Market-Relevant and Reliable Measure of Public Feeling?," *International Journal of Market Research*, vol. 56, no. 2, pp. 231-247, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[13] Jiabei Liu et al., "Application of Deep Learning-Based Natural Language Processing in Multilingual Sentiment Analysis," *Mediterranean Journal of Basic and Applied Sciences*, vol. 8, no. 2, pp. 243-260, 2024. [Google Scholar] [Publisher Link]

[14] Neeraj Anand Sharma, A.B.M. Shawkat Ali, and Muhammad Ashad Kabir, "A Review of Sentiment Analysis: Tasks, Applications, and Deep Learning Techniques," *International Journal of Data Science and Analytics*, vol. 19, pp. 351-388, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[15] Danfeng Zhao, Xiaolian Chen, and Yan Chen, "Named Entity Recognition for Chinese Texts on Marine Coral Reef Ecosystems Based on the BERT-BiGRU-Att-CRF Model," *Applied Sciences*, vol. 14, no. 13, pp. 1-17, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[16] Erik Cambria et al., "Affective Computing and Sentiment Analysis," *A Practical Guide to Sentiment Analysis*, pp. 1-10, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[17] Yan Wang et al., "A Systematic Review on Affective Computing: Emotion Models, Databases, and Recent Advances," *Information Fusion*, vol. 83-84, pp. 19-52, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[18] Doaa Mohey El-Din Mohamed Hussein, "A Survey on Sentiment Analysis Challenges," *Journal of King Saud University - Engineering Sciences*, vol. 30, no. 4, pp. 330-338, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[19] Galina Ilieva et al., "Social Media Influencers: Customer Attitudes and Impact on Purchase Behaviour," *Information*, vol. 15, no. 6, pp. 1-32, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[20] Hitkul Jangid et al., "Aspect-Based Financial Sentiment Analysis using Deep Learning," *Companion Proceedings of the Web Conference 2018*, New York, USA, pp. 1961-1966, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[21] Prince Shiva Chaudhary, Mohit R. Khurana, and Mukund Ayalasomayajula, "Real-World Applications of Data Analytics, Big Data, and Machine Learning," *Data Analytics and Machine Learning*, pp. 237-263, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[22] Ahmed Sulaiman M. Alharbi, and Elise de Doncker, "Twitter Sentiment Analysis with a Deep Neural Network: An Enhanced Approach Using User Behavioral Information," *Cognitive Systems Research*, vol. 54, pp. 50-61, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[23] Darie Moldovan, "A Majority Voting Framework for Reliable Sentiment Analysis of Product Reviews," *PeerJ Computer Science*, vol. 11, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[24] Fazeel Abid et al., "Sentiment Analysis through Recurrent Variants Latterly on Convolutional Neural Network of Twitter," *Future Generation Computer Systems*, vol. 95, pp. 292-308, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[25] S. Kiritchenko, X. Zhu, and S.M. Mohammad, "Sentiment Analysis of Short Informal Texts," *Journal of Artificial Intelligence Research*, vol. 50, pp. 723-762, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[26] O. Hwang Kwon et al., "Sentiment Analysis of the United States Public Support of Nuclear Power on Social Media Using Large Language Models," *Renewable and Sustainable Energy Reviews*, vol. 200, 2024. [CrossRef] [Google Scholar] [Publisher Link]