scientific reports

OPEN

A swarm-optimization based fusion model of sentiment analysis for cryptocurrency price prediction

Dimple Tiwari[1], Bhoopesh Singh Bhati[2], Bharti Nagpal[3], Amal Al-Rasheed[4]✉, Masresha Getahun[5]✉ & Ben Othman Soufiene[6]

Social media has attracted society for decades due to its reciprocal and real-life nature. It influenced almost all societal entities, including governments, academics, industries, health, and finance. The Social Network generates unstructured information about brands, political issues, cryptocurrencies, and global pandemics. The major challenge is translating this information into reliable consumer opinion as it contains jargon, abbreviations, and reference links with previous content. Several ensemble models have been introduced to mine the enormous noisy range on social platforms. Still, these need more predictability and are the less-generalized models for social sentiment analysis. Hence, an optimized stacked-Long Short-Term Memory (LSTM)-based sentiment analysis model is proposed for cryptocurrency price prediction. The model can find the relationships of latent contextual semantic and co-occurrence statistical features between phrases in a sentence. Additionally, the proposed model comprises multiple LSTM layers, and each layer is optimized with Particle Swarm Optimization (PSO) technique to learn based on the best hyperparameters. The model's efficiency is measured in terms of confusion matrix, weighted f1-Score, weighted Precision, weighted Recall, training accuracy, and testing accuracy. Moreover, comparative results reveal that an optimized stacked LSTM outperformed. The objective of the proposed model is to introduce a benchmark sentiment analysis model for predicting cryptocurrency prices, which will be helpful for other societal sentiment predictions. A pretty significant thing for this presented model is that it can process multilingual and cross-platform social media data. This could be achieved by combining LSTMs with multilingual embeddings, fine-tuning, and effective preprocessing for providing accurate and robust sentiment analysis across diverse languages, platforms, and communication styles.

Keywords Deep Learning, Ensemble Learning, Swarm Optimization, Sentiment Analysis, Cryptocurrency

With the rapid development of the Internet of Things (IoT), SN attracted billions of people to share their emotions and opinions beyond the limitations of geographical distance. Social platforms connect society worldwide and create faith among the religion of different cultures and countries. Integrating the IoT, cloud computing, and social media accelerates the reciprocal and communication quality of the SN. Due to the ripple effect of the SN, it has been working in various fields of real-life. A broad definition of a "social network" refers to any system or platform that helps in facilitating social interaction and connecting individuals or groups. All the social media sites, including Facebook, Twitter (X), Instagram, LinkedIn, and others, through which one can create a profile, share content, and connect with other users, are all social networks[1]. Reciprocal and real-life nature means the interactive, mutual, and authentic aspects of social media interactions that seem to have the features of real-world social dynamics. Social media provides for two-way communication where users can interact with one another's content, in the context of likes, comments, sharing, or direct messages, among others. Relationships are fostered by reciprocity as users respond more often to people who interact with their content. Social media often represents real-life ties, such as friendships, family relationships, professional networks, and shared communities. Users share aspects of their daily lives, opinions, or milestones, thereby

[1]School of Engineering & Technology, Vivekananda Institute of Professional Studies - Technical Campus, Delhi, India. [2]Indian Institute of Information Technology, Sonepat, India. [3]NSUT East Campus (Formerly Ambedkar Institute of Advanced Communication Technologies & Research), Delhi, India. [4]Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, 11671 Riyadh, Saudi Arabia. [5]Department of Computer Science and Information Technology, College of Engineering and Technology, Kebri Dehar University, Kebri Dehar, Ethiopia. [6]PRINCE Laboratory Research, ISITcom, University of Sousse, Hammam Sousse, Tunisia. ✉email: aaalrasheed@pnu.edu.sa; masreshaggetahun@gmail.com

creating an authentic and relatable experience. Thus, the "reciprocal and real-life nature" together brings out the way social media integrates elements of authentic, mutual interaction into the online experience as a natural extension of human communication. Social media has profoundly impacted many sectors, changing the way they function, communicate, and interact with their followers. Governments use social media as a source of public information, implementing policies, and taking direct action with citizens. Politicians and parties use platforms to communicate with voters, present agendas, and build support for their efforts. Governments will face challenges such as miscommunication, cyber-attacks, and social sites being used to spread propaganda or extremist ideologies. Academics and researchers use the channels of ResearchGate, Twitter, and LinkedIn to disseminate their work, collaborate worldwide, and share ideas. Social media aids online learning; teachers use platforms to share resources, host live discussions, and reach a wide audience of learners[2]. Businesses utilize social media for focused advertising, customer interaction, and brand creation. Social media channels allow for direct communication with customers, which means real-time feedback and support. Social media campaigns enhance awareness about diseases, preventive care, and healthy lifestyles. Social media helps patients connect with doctors and create forums for peer support, particularly mental health and chronic illness. Social media serves as a source for the government to spread vital health information, as during the COVID-19 outbreak. The governments of all countries monitor the SN for applying the new policies and predicting the trends and facts related to global pandemics; the COVID-19 outbreak is the most prominent example where social platforms play a vital role.

Furthermore, customer-generated reviews for products and services are valuable sources for stock market analysis, brand monitoring, cryptocurrency predictions, and financial decision-making. Therefore, the researchers are interested in SNs and their supporting fields[3]. Social platforms generate a massive amount of unstructured content; SA is used to mine reliable information from that content. Thus, sentiment analysis is an emerging field of research that utilizes the SN's range differently for society. SA can perform mainly using the lexicon-based approach, machine learning-based approach, and hybrid approach. Among the several approaches, AI-based machine learning algorithms can handle noise in unstructured content and provide reliable results. The unstructured information generated by social media has greatly impacted sentiment analysis, bringing about both opportunities and challenges. Opportunities: Social media offers insights into public opinions, emotions, and trends across various demographics. Analysis can cover several topics, industries, or geographies, allowing businesses, governments, and researchers to make better decisions. Real-time sentiment tracking allows businesses to monitor brand perception or assess public reaction to events. Challenges: Most of the posts contain irrelevant or incomplete information like typos, repeated characters, and non-standard grammar. Sentiment is mostly conveyed by images, videos, GIFs, and emojis with text. The full sentiment is captured by multimodal analysis. The sarcastic or ironic posts are challenging to identify since they convey the opposite sentiment of the literal meaning. Social media data is particularly noisy compared to other data sources because of its unique characteristics, which introduce inconsistencies, ambiguities, and irrelevant information. Frequent use of informal expressions like "lol," "omg," or regional slang complicates understanding and standardization. Errors in spelling, such as "gr8" for "great" or repeated letters ("sooo"), create inconsistencies in the data. Sentences may appear disorganized and unpunctuated or grammatically incorrect, rendering them more complicated to understand. Twitter (X) has imposed a character limitation, which might result in disorganized or fragmented expressions of an idea. Generally, short postings lack context for the interpretation of meaning or tone. Jargon, abbreviations, and reference links in social media data provide unique challenges for natural language processing and data analysis. These bring about complications in interpretation, processing, and extraction of meaningful insights. Jargons evolve rapidly with new terms emerging all the time, requiring on-going updates on language models. Some of the jargons look like regular language but have a different meaning when the context is interpreted. Frequently, an abbreviation may carry a different meaning depending on the context. This is heavily preprocessed by expanding abbreviations, recognizing and interpreting jargon in specific contexts resolving shortened links to extract their original content. Traditional AI algorithms such as Naïve Bayes (NB), Support Vector Machine (SVM), K-Neighbors (KNN), and Decision Trees are less potent in handling complications of social media content due to their single learner scheme. To overcome the vulnerabilities of traditional AI algorithms, ensemble methods are introduced with the combined capabilities of multiple learners[4]. The ensemble model integrates several models to build one optimal predictive model. These models are top-rated for SA and handle the various applications of SA, including recommender systems, brand monitoring, customer support, product analysis, stock market analysis, and outbreak predictions. The importance of sentiment analysis is particularly realized in cryptocurrency markets due to their specific characteristics and dynamics. Cryptocurrency markets are very volatile; hence, their prices change greatly in a very short period. Sentiment analysis captures the emotional and psychological state of investors, which drives sudden price swings. Prices in cryptocurrency markets are highly driven by speculative trading and market sentiment rather than intrinsic value or established fundamentals, as seen in traditional markets. This means that, through social media, news, and forums, sentiment analysis will be able to predict short-term price movements as a result of crowd behaviour. Since cryptocurrencies operate without centralized entities controlling them, the community-driven sentiment and perception take over. Trends in decentralized discussions on platforms such as Twitter, Reddit, or Telegram are detected. Influential voices of crypto influencers whose opinions may sway the market sentiment are monitored[5].

The researchers have also applied ensemble techniques to handle the COVID-19 outbreak and Indian farmer protests[6]. Despite the broad applicability of AI-based ensemble techniques, there is a lack of cross-domain comparative study to select the best ensemble approach for SA. On these grounds, we present the comparative analysis of ensemble-based techniques for SA. Ensemble techniques are broadly categorized into two parts: bagging-based ensemble and boosting-based ensemble. Bagging is a machine learning ensemble technique known as bootstrap aggregation, generally used to decrease the variance in noisy datasets. In this, the same weak learner is trained independently on multiple subsets of data and combined with the averaging process to generate an optimal solution[7]. In contrast, boosting ensemble trains the same learner on multiple subsets of data sequentially so that each next learner compensates for the previous learner's weaknesses to reduce the training errors[8]. Although voting and stacking ensembles are two other ensemble approaches, these ensembles combine traditional algorithms and select one of the best[9,10].

Existing research gaps

Identify gaps in current research on sentiment analysis a paper aims to bridge. This will often involve going through the flaws and limitations in current methods and approaches. Currently, many popular models of sentiment analysis fail to identify sarcasm, irony, and ambiguous messages because they lack a deeper level of contextual or cultural understanding. There is therefore a need for developing preprocessing techniques and robust models that can manage noisy, informal, and unstructured text more effectively. Further, it can also create domain-adaptive models or fine-tune existing ones to perform better in certain areas. Previous sentiment analysis models typically focus on isolated text snippets without taking into account the larger conversational or situational context. Furthermore, many existing models are assessed using a limited set of benchmarks, which may not accurately represent the complexities of real-world scenarios.

Contribution

Here, in this paper, an optimized stacked-LSTM model is proposed to predict cryptocurrency prices. Extensive experiments have been conducted on various Bitcoin-related tweets. Multiple learners learn together in an ensemble approach to get more accurate and efficient results than individual learners. The primary objective of this research is to provide the best ensemble approach for sentiment analysis. It will be helpful for society to take advantage of the best ensemble technique in various applications of sentiment analysis. The significant contribution of the work is as follows:

- A stacked LSTM model is proposed with multiple LSTM layers, including PSO hyperparameter optimization, to produce an effective attention-oriented sentiment classification tool.
- Bitcoin-related tweets have been collected to investigate the proposed model's effectiveness for future cryptocurrency price prediction. The PSO-optimized greedy parameter selection of the LSTM network is advantageous for solving discrete problems.
- The standard parameters, namely training accuracy, testing accuracy, weighted Recall, weighted Precision, and weighted-f1-score have been calculated to check the authenticity of the ensemble techniques.
- An extensive comparative analysis is performed to present the contribution of the proposed optimized stacked-LSTM, PSO hyperparameter optimized model. Performance evaluation results show the effectiveness of the proposed PSO-optimized stacked-LSTM model over the existing LSTM network.

The further subsections of this paper are organized as follows: Section "Literature" presents the literature work done in the field of sentiment analysis. Section "Methodology" presents the methodology and its step-by-by procedure. Section "Result analysis" discusses the evaluation criteria and results of the proposed model. Section "Comparative analysis" presents the massive comparative analysis. Section "Findings of the study" shows the finding of the study. At last, Section "Conclusions and future scope" provides the conclusion of the work with future directions.

Literature

Since the late 1990s, SA has been an eminent research topic in Natural Language Processing (NLP), data mining, information retrieval, and AI, attracting researchers due to their vast adaptability in various real-life situations[11]. This literature work is categorized into social SA and ensemble techniques for SA.

Social SA

Several researchers have examined the SA from different perspectives, including knowledge representation, word or phrase extraction, feature selection, and typical linguistic specification. Intelligent cities connect information technology, social networks, and business environments to improve the city's lifestyle. Hence, the SA model has needed for real-time opinion mining[12]. These are executed in the cloud environment to enhance the SN system's performance. Integrating cloud computing and SN platforms increases reachability and connectivity among people[1]. Many researchers have been working in the area of text processing and text analytics. Therefore, Alam et al.[13] examined the impact of different pre-processing techniques on three machine learning algorithms (NB, SVM, and Maximum Entropy (ME)) for checking the variation in their accuracies. Alaoui et al.[14] proposed an adaptable SA approach for real-time user opinion extraction. First, they constructed the word dictionary based on the hashtags collected and then classified the tweets into different classes by fine-tuning the polarity degrees of posts. Researchers are actively working on developing novel sentiment analysis techniques. Here, sentiment classification is a more widespread mechanism that can be performed at the sentence, document, and aspect levels[15]. Prediction of the stock market price is profitable but difficult to identify. Machine learning and statistical models have been widely adopted to measure the variations in stock market prices[16].

Ensemble techniques for SA

Feature selection and generation are two pillars for boosting the performance of SA-related tasks. A combined feature selection methods (Information Gain, Chi-Square, and Gini Index) approach has been introduced to enhance the capability of machine learning algorithms for sentiment classification[17]. A combined strategy of Bootstrap Aggregating (Bagging) and Synthetic Minority Oversampling Techniques (SMOTE) have been proposed for imbalanced sentiment polarity detection[18]. Consumer-generated reviews on different websites help analyze the quality of products and services. The analysis and mining of the studies are very much essential. A hybrid boosted model was proposed for sentiment classification. The proposed model combines the potential of SVM and boosting techniques for the sentiment classification of online reviews[19]. Aspect-based sentiment analysis extracts the context from the text. A cascade framework of classifier ensemble and feature selection has been presented using Particle Swarm Optimization (PSO) for aspect-based sentiment analysis. Conditional Random Field (CRF), ME, and SVM have been selected as base learners. Experiments on two different domain datasets show the efficiency of the proposed cascade framework[20]. The high-dimensional feature of the text documents is a big limitation of SA that minimizes the efficiency of algorithms. An Auto Encoder-based Bagging Prediction Architecture (AEBPA) was proposed by inspiring auto-encoder feature reduction and dimensionality reduction[21]. Ensemble models have widely been used in mining social media data due to their capability to combine the strengths of several algorithms to gain better performance. There are some commonly used ensemble models namely Gradient Boosting Machines, Random Forest, Adaptive Boosting Classifiers, Voting and Stacking that are widely adopted in social mining researches.The higher computational cost of boosting algorithms for text classification is a more significant challenge in front of machine learning techniques. Most boosting techniques investigate all the dataset features that increase the process's learning time. RFBoost was proposed with the rank and filter strategy to overcome this limitation, where first ranking has assigned to each feature and then filters each learning iteration. This step enhances the learning time of RFBoost more than AdaBoost, comparatively[22].

There are various text pre-processing techniques available for feature selection and extraction. Term Frequency-Inverse Document Frequency (TF-IDF) and N-Gram were compared using six machine learning algorithms (SVM, Decision Tree, K-Nearest Neighbors, Logistic Regression, Random Forest, and NB). Demonstration reveals that TF-IDF increases with approximately 3–4% accuracy more than N-Gram[23]. An enhanced XGBoost classifier was proposed named SentiXGBoost by utilizing the capabilities of multiple feature sets, including TF-IDF, N-Gram, Part-of-Speech (PoS), Opinion Lexicon, and Term Frequency. The experiments confirm the outperformance of the enhanced ensemble model[24]. A gradient-boosting ensemble has been applied to the Greek text that learns by different hash functions[25]. SNs are the daily venue of unstructured content from other domains, from sports to human rights. Specific resource generation for every discipline is expensive. Therefore, a Light Gradient Boosting Machine (LGBM) handled the high dimensional and imbalanced data. Comparison with other machine learning algorithms affirms the credibility of the LGBM for SA[26].

Massive literature work and the potential of ensemble-based techniques motivate us to develop a new deep-learning ensemble model for efficient sentiment analysis and prediction.

Methodology

The proposed optimized stacked-LSTM model's primary objective is to accurately identify people's opinions regarding cryptocurrency investment. The presented model has the capability of multiple stacked LSTM layers with optimized parameters that can effectively predict people's sentiment for future crypto investments. Existing models are generally labeled as "less-generalized" since they tend to have difficulty applying what they learned to other scenarios that are not enclosed within their specific training data or domain. This further emphasizes the need for a better, more flexible and context-aware model that can effectively manage the variety and complexity of real-world data, especially in dynamic environments like social media. Some reasons that describe these models as less general are overfitting training data, narrow data reliance, insufficient contextual understanding, limited robustness, domain-specific limitation, and biases in training data, and challenges with transfer learning.

Figure 1 shows the deep architecture of the proposed model with a step-by-step procedure. This model has completed in various steps. Each step is filled with advanced Artificial Intelligence (AI) associated computations such as feature extraction, dimensionality reduction, multiple layering of the neural network, and nature-inspired PSO optimization. Initially, live Tweets are collected regarding the latest cryptocurrencies. Before applying the neural model to the $Tweets\,[t_1, t_2, t_3.................t_n]$ corpus, noise reduction has been performed, like punctuation, whitespace, numbers, and URL removal. After removing noise from the dataset, it is tokenized in the form of Corpus $[w_1, w_2, w_3.................w_n]$. The essential feature engineering has been applied, namely N-Gram, Part-of-Speech (PoS) tagging, and sentiment word extraction to focus on the utility features. Dimensionality reduction is necessary for solving the AI model's overfitting problem. Therefore, Principle Component Analysis collects the critical Principle Components (PCs) from the large sample. Finally, Glove word embedding is applied to generate the global word vectors vect $[v_1, v_2, v_3.................v_n]$. After that, the multiple layers of the LSTM neural network have trained on the worldwide vector vect $[v_1, v_2, v_3.................v_n]$ with PSO-optimized hyperparameters. Ultimately, the knowledge-enriched model has been produced to mine the opinion and predict Cryptocurrency's future. The proposed model input sentences sequentially, one token at a time. Each word or phrase is represented as a vector (via embeddings), and the LSTM updates its internal state based on the current input and its previous state. This enables it to accumulate information about preceding words, phrases, and their relationships. This means that, by incrementally updating its internal state and learning what to retain or discard, an LSTM can capture meaningful relationships between phrases, even across long sentences. LSTMs are much better than traditional RNNs at maintaining relationships over longer distances, such as connecting "If–then." structures in a sentence.

Data engineering

Data elicitation

The 9,998 cryptocurrency-associated unlabelled tweets have been collected from Kaggle.com to predict people's sentiment towards Cryptocurrency. The Text-Blob corpus method is exploited to label the tweets as Positive,

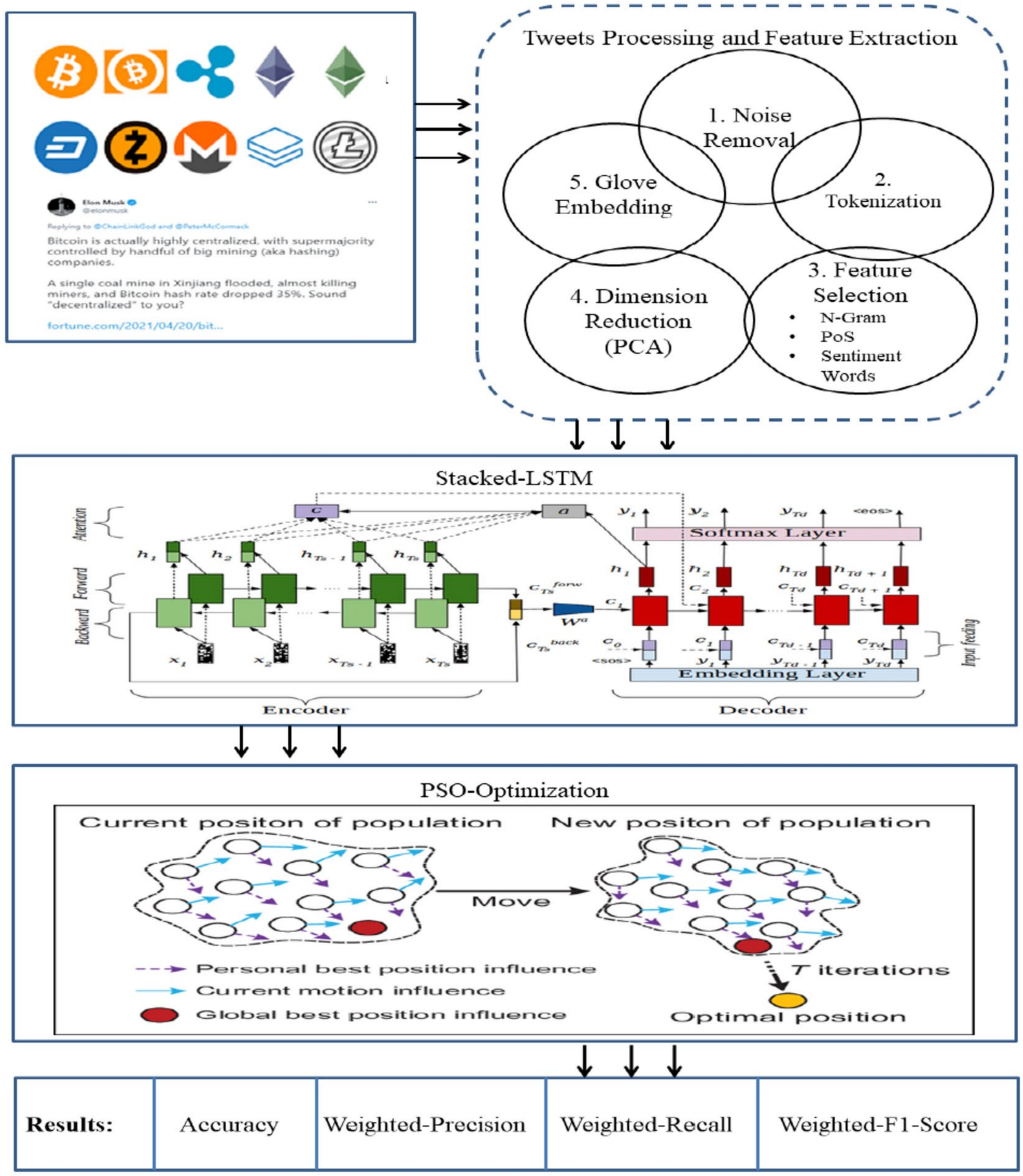

Figure1. The Architecture of the proposed PSO-optimized staked-LSTM model for cryptocurrency price prediction.

Negative, and Neutral. The polarity and subjectivity score are also calculated to get a deeper insight into the tweets. Table 1 presents the sample of the selected cryptocurrency tweets dataset.

The sentences in the text that relate to Twitter data seem clear, but can be improved for readers with different expertise levels. However, informal language that defines a tweet may be full of hash tags, mentions, and abbreviations, making it unclear for a reader with little technical background. Several pre-processing steps have been used in this work to enhance the clarity of the sentences including text cleaning, hash tag removal, emoji to text conversion handling abbreviations and slangs.

Data processing

Tweets are generally a composition of incomplete expressions containing noise, poorly organized sentences, irregular grammar, frequent acronyms, and non-dictionary phases. The unstructured content degraded the performance of sentiment classification. Therefore, refining the tweets before training the model is required, which is done in various stages.

- **Noise removal:** This phase includes numbers, Non-ASCII characters, URL links, punctuation removal, and harmful reference replacement.

User_Name	Text	Hashtag	Polarity-Score	Subjectivity-Score	Sentiment-Type
DeSota Wilson	Blue Ridge Bank shares halted by NYSE after #bitcoin ATM announcement	['bitcoin']	0	0.1	NEUTRAL
Tdlmatias	Guys evening, I have read this article about BTC and would like to share with you all—https://t.co/o6wn7ppkVY	['Bitcoin]	0	0.1	NEUTRAL
Alex Kirchmaier	This network is secured by 9 508 nodes as of today. Soon, the biggest bears will recognise	['BTC']	0	0.1	NEUTRAL
DeSota Wilson	investment is revolutionary for #crypto but other firms may not do the same just yet -	['bitcoin', 'crypto']	0.0625	0.25	NEGATIVE
#Mailey	Free #Mailey Advice: Imagine in 2019 there was a #Bitcoin mining company that mined 20,000	['Mailey', 'Bitcoin', 'BTC']	0.4	0.8	POSITIVE

Table 1. Sample of the experimented cryptocurrency tweets dataset.

Step-1: Define Hyperparameter Space • Define the ranges of each hyperparameter (number of layers: [1,5], learning rate: [0.0001, 0.01])
Step-2: Initialize Particles: • Generate random initial positions and velocities for a swarm of particles within the defined hyperparameter space
Step-3: Train Stacked-LSTM Model • For each particle, train an LSTM model using the hyperparameters represented by the particle's position. Evaluate the model's performance on a validation dataset
Step-4: Evaluate Fitness: • Compute the fitness of each particle (e.g., validation accuracy)
Step-5: Update Personal and Global Bests • Update- PBest for each particle if its current position yields better performance • Update- GBest if a particle's fitness is better than the global best obtained so far
Step-6: Update Velocity and Position • Update velocity and position of all particles according to update equations of PSO
Step-7: Iterate • Steps 3–6 are repeated for a predefined number of iterations or until convergence
Step-8: Return Optimal Hyperparameters • After convergence, the GBest is used as the optimal set of hyperparameters

Table 2. Steps in PSO for LSTM Hyperparameter Optimization:

- **Part-of-Speech (PoS) Tagging:** Provides elementary information to identify aspect terms. It includes categorizing the words in a text corpus with a particular part of speech based on their definition and context. Table 2 describes the tagging of speech presented in sentences.

Part-of-Speech	Tag
Noun	n
Verb	v
Adjective	a
Adverb	r

Word Cluster is a hierarchal clustering technique of assigning words to a cluster according to their context. A feature vector of length five (binary) denote the cluster identifiers.

Review:	But	the	staff	was	very	horrible	to	us
Cluster-ID:	00101	01010	11,010	01100	00101	11,101	01110	01111

Semantic Orientation (SO): It measures how much the word is associated with positive and negative sentiment. PointWise Mutual Information (PMI), a measure of association of token t with respect to positive and negative reviews.

$$SO(t) = PMI(t, pos\text{Rev}) - PMI(t, neg\text{Rev}) \tag{1}$$

$$PMI(t, neg\text{Rev}) = \log \frac{freq(t, neg\text{Rev}) * N)}{freq(t) * M} \tag{2}$$

where $freq(t, neg\text{Rev})$ is the frequency of term t in a negative tweet, $freq(t)$ is the frequency of t in the corpus, M is the number of tokens in a negative review, and N is the number of tokens in the corpus [27].**GloVe Embedding:** For generating the co-occurrence matrix, GloVe incorporates global statistics to obtain word vectors. The central principle of GloVe is that the co-occurrence ratios between two words in a context are strongly connected to form a meaning.

- **Normalization:** For scaling the vectors in uniformity, min–max normalization is performed on the dataset. [Eq. 3] presents the formulation of min–max normalization.

$$V' = \frac{(V - \min)}{(\max - \min)} \tag{3}$$

V The set of attributes $\max$ denotes the maximum value from the attributes, and $\min$ represents the minimum value in the dataset's attributes V'. Represents the normalized data that holds values from 0 to 1.

Figure 2 presents the text length of positive, negative, and neutral tweets. The above pre-processing steps enrich the tweets for building an efficient sentiment classification model. The text content is further converted into a co-occurrence matrix as an input for the Stacked-LSTM model[28].

Stacked LSTM

LSTM is the extended version of RNN introduced by Hochreiter and schmidhuber[29]. It overcomes the problem of RNN and can handle the long dependencies between phrases. The stacking of multiple layers of LSTM arranged hierarchically in a single model allows one layer to pass the output to the next in sequence, with each layer being able to detect more complex patterns and temporal dependencies at different levels of abstraction. The first LSTM layer receives the input sequence, such as a time series or a sentence, directly and analyzes this input to capture fundamental dependencies and patterns like trends or short-term fluctuations in time-series data or local context in text. The output from this layer includes hidden states and cell states for each timestep, which carry information on the dynamics of the sequence. Each subsequent LSTM layer takes hidden states from the previous layer rather than raw input data as an input. Therefore, these layers can process information already analyzed, meaning they can be able to learn higher-level abstract features. In this case, while the first layer may depend on local dependency, the second layer could point out longer dependency. The deeper layers usually focus on more complex or long-range patterns in the data, which improves the model's ability to understand intricate relationships within the sequence.

The LSTM embedded three major gates, namely forget gate (f_t), input gate (i_t), and output gate (o_t). These manage the flow of information for updating, writing, and reading in the network. [Eqs. 4, 5, 6, 7, 8, 9] presents the mathematical evaluation of LSTM[30].

$$f_t = \sigma(W_f[h_{t-1}; x_t]) + b_f \tag{4}$$

$$i_t = \sigma(W_i[h_{t-1}; x_t]) + b_i \tag{5}$$

$$o_t = \sigma(Wo[h_{t-1}; x_t]) + b_o \tag{6}$$

$$\tilde{c}_t = \tanh(Wc[h_{t-1}; x_t]) + b_c \tag{7}$$

$$c_t = i_t \otimes \tilde{c}_t + f_t \otimes c_{t-1} \tag{8}$$

$$h_t = o_t \otimes \tanh(c_t) \tag{9}$$

$$H = [\vec{h}_t : \underset{t}{\overleftarrow{h}}] \tag{10}$$

where,h_t, c_t, and x_t are the hidden vector, cell state, an input vector. $W_f, W_i, W_o, W_c, b_f, b_i, b_o, b_c$ Are the weight matrix and bias vectors for each gate. For encoding, the sequence in backward and forward Bidirectional-LSTM is formulated in [Eq. 10]. The proposed model stacked multiple LSTM layers for effective result manipulations.

Due to the ability of this architecture to tackle complex sequential data is impressive, using this stacked-LSTM architecture in sentiment analysis has several benefits.

- **Deeper Understanding of Text:** A stacked-LSTM, with multiple LSTM layers, enables the model to learn hierarchical features of text data. Lower layers capture simpler patterns, while upper layers capture higher-level abstractions.

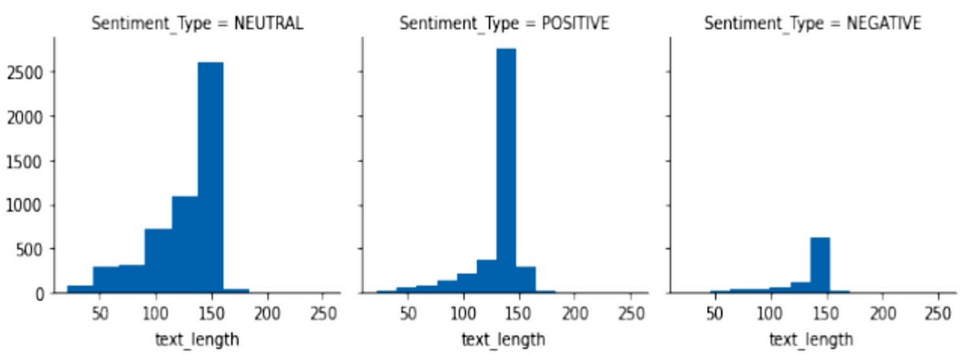

Fig. 2. The Text Length Count of Neutral, Positive, and Negative Tweets.

- **Improved Contextuality:** Hierarchical representation facilitates the model to pick up subtle sentiment shifts in long texts or complex sentences.
- **Improved Expressive Power:** Every level in an LSTM is a degree of non-linearity, allowing stacked architecture to express complex word phrase relationships which underlie sentiment.
- **Flexibility and Scalability:** Stacked-LSTMs may be adapted based on the nature of the complexity of the data by changing the number of layers and hidden units. For very large datasets or tasks with rich sentiment patterns, deeper architectures could be more helpful.

Latent contextual semantic features refer to the underlying or implicit meanings of words that can be understood from their context within a sentence, paragraph, or larger document. These features are crucial because the meaning of a word often relies on the words around it and the broader context in which it appears. Words can take on different meanings based on their context. For instance, “cool” can signify “cold” or “impressive” depending on how it's used in a sentence. Contextual semantic features assist the model in grasping these nuances. The sentiment of a sentence can be shaped by the context in which a word is used. For example, "I love the concept but hate the implementation" includes words like “love” and "hate," yet the overall sentiment is complex and requires context to fully understand the mixed feelings. By employing techniques such as word embeddings (like Word2Vec, GloVe, or contextualized embeddings like BERT), the model can uncover the deeper semantic connections between words. For example, it can learn that “happy” and “joyful” are semantically similar and convey a positive sentiment.

Attention mechanism

The attention mechanism is applied to gather the more profound context information present in the tweets. It focused on the local level and global levels. The local level pays attention to only a few words of sequence, and the global level pays selective attention to all the words available in the sequence. This paper applied local attention and focused on the small width of sequence $S = \{x_1, x_2, ...x_n\}$ and $h_i = [h1, h2, ..., h_n]$ hidden word vectors calculated from sentences.

$$h_{t,t^t} = \tanh(x_t^T * W_t + x_{t_t}^T * W_t + b_t) \tag{11}$$

$$e_{t,t^t} = \sigma(x_t^T * W_a + x_t * W_t + b_a) \tag{12}$$

$$a_t = soft\max(e_t) \tag{13}$$

$$l_t = \sum_{t^t} a_{t,t^t * x_{t^t}} \tag{14}$$

The multiplicative attention of the given sequence is generated in [Eqs. 11, 12, 13] and normalized by [Eq. 12]. After that, the context vectors of the normalized sequence are calculated in [Eq. 14]. The main differences between latent contextual semantics and co-occurrence statistical features lie in how they model relationships between words, the kind of information that they capture, and their applications in NLP.

- **Latent Contextual Semantics:** The hidden meaning or relationship that is implied by the context in which words or phrases are used. In it, the goal is to understand the semantic relationship between words through the use of their context. Captured using advanced models such as embeddings (word2vec, glove, BERT), or other contextual representations.
- **Co-occurrence statistical features:** These are statistical relationships between words based on the frequency of their simultaneous appearance in a given corpus. They focus on the surface patterns of word usage without attempting to model deeper semantics; thus, they are often captured using techniques such as co-occurrence matrices, Point-wise Mutual Information (PMI), or word frequency counts.

PSO optimization

The next step is featuring selection, which is done by applying the greedy PSO algorithm. PSO is one of the effective techniques for calculating the optimized features from noisy subsets. It is biologically motivated and inspired by a horde of birds searching for food. Each individual in the dataset represents a bird that holds the fitness value in the search space. In this case, it has a best-labeled lexicon, and the possible solutions are called the particles that are the parts of the population. Every particle keeps its best solution called as $pbest$. At the same time, the best value selected from the whole swarm is called as $gbest$. Figure 3 presents the velocity and position relationship in PSO.

Majorly, PSO works on the principle of updating position and changing velocity. In the first step, every particle changes its velocity towards its $pbest$ and $gbest$. In the second step, the particle updates its position. Every particle in space is represented as $D - \dim ension$ a vector. The $_{\mathrm{nth}}$ vector can be represented as $A_n = (a_{n1}, a_{n2}, ...a_{nD})$. The velocity of the nth particle can be represented as $V_n = (v_{n1}, v_{n2}, ...v_{nD})$ where the best previous position of a particle can be represented as $P_n = (p_{n1}, p_{n2}, ...p_{nD})$.

Particle Swarm Optimization is a population-based optimization algorithm based on the social behavior of birds and fish. It is applicable for searching through the space of hyperparameters in LSTM layers in a systematic manner to get the best values together. Table 2 represents the steps of PSO for LSTM hyperparameter tunning.

The PSO has a number of advantages over traditional optimization techniques, especially in complex and high-dimensional problems. Some of the key advantages are as follows:

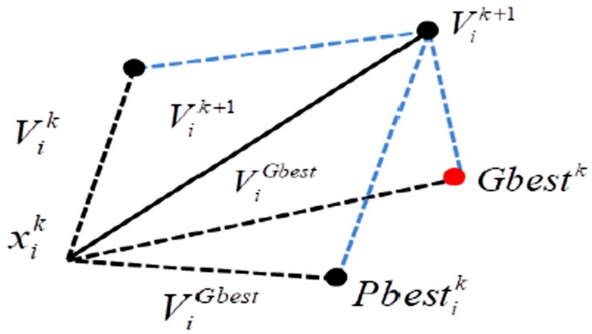

Fig. 3. Relationship of PSO within velocity and position vectors.

- Ease of Implementation and Simplicity: PSO is relatively simpler to implement and understand compared to other optimization algorithms, such as GA or SA.
- Robustness to Non-Linear, Multimodal, and Complex Objective Functions: PSO can solve complex, non-linear, and multimodal optimization problems effectively, such as those difficult to solve by the traditional gradient-based algorithms.
- Parallelism and Global Search Ability: In PSO, every particle is an independent agent that computes its position and updates its state according to the experiences it gains and its neighbors. In this manner, a large number of particles can be concurrently exploring the search space in the pursuit of a global optimum.
- Low Requirement of Derivatives: Unlike gradient-based optimization methods, PSO does not require the derivatives of the objective function to be known or continuous.

Sentiment classification using aspect category

The Softmax classification layer is applied in the proposed model to classify cryptocurrency tweets. This layer receives the input from the local attention layer and calculates the probability score of each sentiment label and aspect category. After that, the aspect category and sentiment label are predicted as follows in [Eq. 15].

$$\hat{y} = soft\max(W * l_t + b) \tag{15}$$

$$w_{t+1} = w_t - \alpha m_t \tag{16}$$

$$m_t = \beta m_{t-1} + (1-\beta)\left[\frac{\delta L}{\delta w_t}\right] \tag{17}$$

Here b W and are the bias vectors and weight matrix. The categorical cross-entropy loss function is exploited with Adam weight optimizer to measure the loss of the model in Eqs. 16, 17. Where, m_t is the gradient at time t, m_{t-1} is at time $t-1.w_t$ is the weights at time t, α_t is the learning rate, and β represents the moving average parameter [constant 0.9].

Challenges in model training

The Particle Swarm Optimization (PSO) process for a stacked LSTM architecture has its own problem in hyperparameter optimization.

- **High-Dimensional Search Space:** Hyperparameter optimization for stacked LSTMs has a high-dimensional search space. It includes parameters such as the number of layers, number of units per layer, learning rate, dropout rate, and batch size. PSO finds it difficult to explore and exploit such a large space efficiently.
- **Particle Initialization:** Ineffective initialization of particles may cause them to not explore the search space effectively.
- **Loss Function Sensitivity:** The optimization process is highly dependent on the loss function. PSO may have difficulty with noisy or non-smooth loss landscapes.
- **Stochastic Behaviour:** The PSO is a stochastic algorithm. Results between runs may not be the same due to randomness in initialization and updating. Run multiple times and take the average value to ensure convergence.

Result analysis

This section presents the result and discussion of the proposed optimized stacked-LSTM model. To evaluate the performance of model, cryptocurrency Tweets have been selected and different evaluation measures are experimented to test the results critically. Focusing on cryptocurrency price prediction as the target for a sentiment analysis model presents several compelling reasons, particularly due to the unique characteristics of cryptocurrencies and their markets. Cryptocurrency prices are often heavily influenced by market sentiment, news, and public opinion rather than solely relying on traditional financial metrics. Unlike stocks, which are typically more affected by company performance and financial reports, cryptocurrencies like Bitcoin and Ethereum are significantly swayed by the collective sentiment expressed on social media, forums, and news outlets. Such an increase in positive sentiment or a major regulatory announcement can create huge price fluctuations. Thus, this makes sentiment analysis an important tool in predicting market movements. Moreover, cryptocurrency markets run 24/7, so shifts in sentiment can occur at any time. Therefore, the need for real-time prediction models is inevitable. Traders and investors in the cryptocurrency space often make decisions based on the market sentiment and news within minutes. Real-time sentiment analysis makes these predictions dynamic, thereby potentially allowing faster identification of price changes than would be possible using traditional models. A positive tweet about a cryptocurrency can lead to a price increase, and the immediate detection of such shifts in sentiment can give one an edge in competition.

Evaluation measures

The selection of critical evaluation measures is very much required for effective testing of the results. Training Accuracy, Testing Accuracy, Precision, Recall, and F1-Score are calculated to check the authenticity of the results confusion matrix. Table 3 presents the description of evaluation measures used to check the authenticity of proposed model. Metrics used for evaluation could be weighted F1-score, precision, or recall, amongst others, primarily according to the particular problem being targeted and the characteristic features of the used dataset. Recall and Precision refer to two metrics that measure slightly different aspects of how well a model is performing. Precision is how many of the predicted positive instances are actually correct, which becomes important in those scenarios where a false positive costs a lot. For example, spam detection or medical diagnosis. Recall is how many of the actual positive instances were found, which is important when not finding a positive instance has huge consequences, like cancer detection. F1-Score balances the precision and recall and delivers a harmonic mean of precision and recall, in this way balancing two metrics within a single number considering false positives as well as false negatives. n many real-world situations, the dataset is imbalanced, in that some classes have low representation against the others. Weighted F1-Score adjusts for this imbalance by weighting the F1-Score of each class according to its proportion in the dataset, ensuring that the performance of minority classes is not overshadowed by majority classes. Weighted metrics provide some fine-grained insights into performance, especially if the class distributions are imbalanced, whereas for non-weighted metrics, their picture of a general performance might either be less nuanced or even false. Weighted metrics adjust their contribution to the class's proportion of the dataset that each class contributed to the general score. This ensures the performance on the minority classes—those with a few samples not being overshadowed by the majority class. However, non-weighted metrics do adjust the contribution of each class toward the overall score based on its proportion in the dataset. In this way, the performance on the minority classes—those with a few samples—is not overshadowed by the majority class.

Accuracy is the ratio of correct prediction to total prediction, Precision is the ratio of actual positive to total positive predictions, Recall is the ratio of actual positive truly positive predictions, and f1 score is the harmonic mean of Precision and Recall, which is used to present conclusive result of Precision and recall [31]. MAE and MSE are the measures used for regression analysis. MAE is the measure of error between paired observations expressing the same occurrence. MSE measures the average of the squares of the errors. Both training accuracy and testing accuracy should be assessed to understand the performance and generalization capabilities of a machine learning model. Training Accuracy: Measures how well the model fits the training data. High training accuracy indicates that the model has learned patterns in the training set. Testing Accuracy: Measures how well the model generalizes to unseen data. Significant degradation in the testing accuracy from training accuracy may indicate overfitting, whereby a model memorizes the training data rather than truly learning general patterns. A model suffering from overfitting will give higher accuracy on training data compared to new real-world data.

Result discussion

The experiments are conducted on Cryptocurrency related Tweets, where proposed model achieves higher accuracy in terms of standard classification and regression measures. The proposed optimized stacked LSTM model aims to achieve the best accuracy for both the classification and regression problems. Table 4 presents

Evaluation Measure	Formula
Accuracy	$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$
Weighted-Precision	$Weighted - \Pr ecision = \frac{\sum_{i=1}^{n} \lvert y_i \rvert \frac{TP_i}{TP_i+FP_i}}{\sum_{i=1}^{n} \lvert y_i \rvert}$
Weighted-Recall	$Weighted - \text{Recall} = \frac{\sum_{i=1}^{n} \lvert y_i \rvert \frac{TP_i}{TP_i+Fn_i}}{\sum_{i=1}^{n} \lvert y_i \rvert}$
Weighted F1-Score	$Weighted - F1 - Score = \sum_{i=1}^{n} w_i \times F1 - Score_i$
Mean Absolute Error (MAE)	$MAE = \frac{\sum_{i=1}^{n} \lvert y_i - x_i \rvert}{n} \left\{ \begin{array}{l} \#\ \mathrm{y_i}\ \text{is prediction} \\ \#\ \mathrm{x_i}\ \text{is original value} \end{array} \right\}$
Mean Squared Error (MSE)	$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 \left\{ \begin{array}{l} \#Y_i\ is\ observed\ values \\ \#Y_i\ is\ predicted\ values \end{array} \right.$

Table 3. Description of the Calculated Measures for Result Evaluation.

	Precision	Recall	F1-Score	Support
Negative	0.81	0.61	0.70	184
Positive	0.95	0.93	0.94	1025
Neutral	0.88	0.95	0.91	782
Macro-AVG	0.88	0.83	0.85	1991
Weighted-AVG	0.91	0.91	0.90	1991
Training Accuracy	98%			
Testing Accuracy	91%			

Table 4. Confusion Matrix of the Proposed Optimized Stacked-LSTM Model.

the classification report of the proposed model on Cryptocurrency associated Tweets. It is visualized that the proposed model achieves 98% training and 91% testing accuracy on Tweets. The training and testing accuracy shows the goodness of fit. The proposed model acquires the 91% weighted Precision, 91% weighted recall, and 90% F1-score that represents the model's effectiveness on all the measurement units. In sentiment analysis, the confusion matrix is interpreted to determine how well the model classifies sentiments, for example, positive, negative, and neutral. This interpretation gives a detailed breakdown of correct and incorrect predictions across different sentiment classes, which enables insights into the strengths and weaknesses of the model. A confusion matrix can reveal if the model is biased toward predicting more frequent sentiments. The terms "Neutral" and "Positive" are often used interchangeably in language, causing confusion. The confusion matrix can be used to identify which classes are most often misclassified and guide efforts such as better feature engineering or dataset cleaning. In sentiment analysis, the confusion matrix provides a granular view of the model's predictions, highlighting specific areas for improvement. The proposed optimized Stacked-LSTM model can serve as a benchmark for sentiment analysis in cryptocurrency price prediction if it effectively addresses the challenges of the cryptocurrency domain and offers superior performance.

- **Temporal Sequence Modeling:** The proposed model can effectively capture temporal dependencies and changing patterns in sentiment over time, which is critical in the cryptocurrency market, as they are dynamic responses to changes in sentiment.
- **Cryptocurrency Context Understanding:** The LSTM is trained and fine-tuned on datasets that include cryptocurrency-specific jargon (e.g., "FOMO," "HODL") and informal language, capturing the nuances of this domain. The model is designed to adapt to the high volatility and sentiment-driven price swings that characterize cryptocurrency markets, making it more suitable for this use case than generic sentiment models.
- **Feature Engineering:** Evaluations made from the proposed model showed that the sentiment scores were strongly related to cryptocurrency price movements and proved to be predictive relevance. Integrating these with other market indicators such as historical prices or trading volume into a predictive price movement model delivers much more accurate predictions.
- **Noise Resilience:** This Stacked-LSTM is robust to noisy, unstructured, and informal text data, which is commonly found in social media platforms when discussing cryptocurrencies. The model captures temporal sentiment trends, thus it can respond well to the fast-changing market conditions.
- **Evaluating its performance over diversified datasets:** It consistently runs strong over diversified datasets such as the ones taken for social media news articles and also over trading forums and diverse timelines and cryptocurrencies outperforms others even more competitive with simpler versions of RNN and CNN, compared with the previous best benchmark sentiment analysis statistical model.

The proposed model is generalizable as it is trained on many parameters. The presented model uses GloVe word embedding feature rich in domain adaption for solid foundation to understand language. The stacked-LSTM is pre-trained on a large general corpus to minimize over-fitting on domain-specific data. This Dropout layer was added to prevent the model from memorizing the patterns of its training data, making the model be a better generalizer. The model allows the integration of attention mechanisms in order to focus on the most relevant parts of the input sequence. An attention layer is used in the proposed Staked-LSTM for capturing universal domain-independent contextual relationships.

Figures 4 and 5 present the closing and scaled closing price of the Bitcoin cryptocurrency. Figure 6 illustrates the future prediction of Cryptocurrency by the proposed model, which is significantly closer to actual values. It shows the accuracy and effectiveness of the proposed optimized stacked LSTM model. Table 5 presents the regression outcome of the proposed model on the Bitcoin time-series dataset. The price fluctuations based on time per second are crawled from the "CoinMarketCap" price tracking website using "cryptcmd" scraper. It has been shown that the model performs accurately and produces very less amount of error as MAE = 0.0441, MAE Percentage 8.658. Mean Absolute Scaled Error = 0.004, and MSE = 0.003. It shows the outperformance the proposed model on both classification and regression problems.

There is great trade-offs concerning the performance versus computational efficiency in the proposed model of Stacked-LSTM. Such trade-offs often depend on the application and/or dataset under study and desired outcome.

- **Use of Truncated Backpropagation through Time (TBPTT):** TBPTT enables LSTMs to learn from long sequences by breaking them into smaller chunks while retaining essential temporal dependencies. The com-

Fig. 4. Closing Price of the Bitcoin Cryptocurrency.

Fig. 5. Scaled Closing Price of the Bitcoin Cryptocurrency.

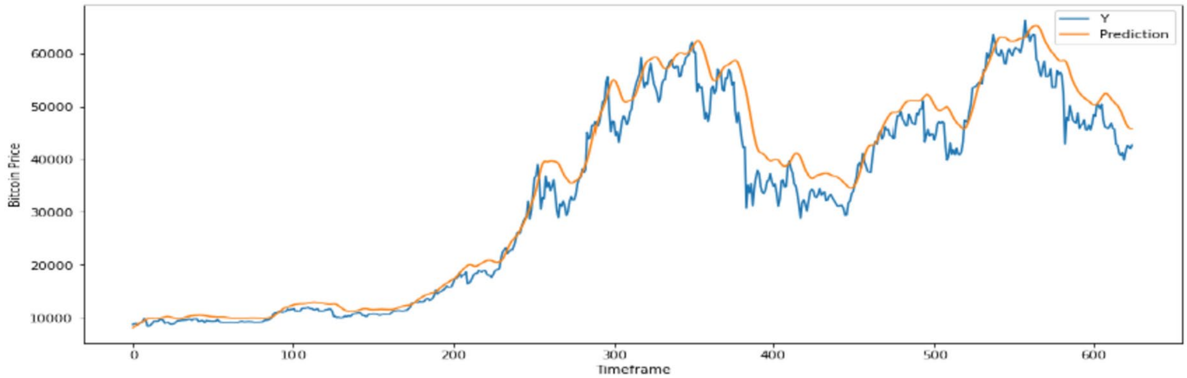

Fig. 6. Cryptocurrency Price Prediction by Proposed Optimized Stacked LSTM.

Measure	Value
Mean absolute error	0.0441660947681782
The systematic mean absolute percentage error	8.658282573123673
Mean_absolute_scaled_error	0.004965664833348569
Mean square error	0.003961048288506163

Table 5. Regression Results of the Proposed Optimized Stacked-LSTM Model.

putational requirements are reduced by truncating the sequence length for backpropagation, and training is faster without significant performance loss for many tasks.
- **Reduced Model Complexity:** Even if the number of layers or hidden units in the LSTM is reduced, it can still capture the important patterns in the data, especially for simpler tasks or datasets that are less complex. Fewer parameters lead to faster training and inference, reduced memory usage, and lower computational costs.
- **Attention Mechanisms:** Adding an attention mechanism to LSTMs makes the model more attentive about parts of the sequence it should pay attention to, and thus, improving performance on tasks such as translation and summarization. Attention reduces the burden on the LSTM for processing all time steps equally; results become much better with smaller and shallower models.
- **Layer-Freezing during Fine-Tuning:** Freezing the lower layers of an LSTM-based model during task-specific fine-tuning preserves general features learned, ensuring that this method focuses on task-specific learning. As a result, gradient updates for frozen layers are avoided, reducing computational overhead during training.

The proposed Optimized Stacked-LSTM model will be effective in applying it to the non-financial domains, as it enables the processing of sequential data coupled with the ability to capture long-term dependencies. It is significantly flexible and can be applied to many tasks from assorted domains. It is very generalizable and can be applied to several non-financial domains. Stacked-LSTM models can achieve strong performance in tasks ranging from healthcare to NLP and beyond by tailoring the preprocessing, architecture, and training process to the specific requirements of the target domain. The proposed stacked-LSTM model would solve most the problems faced during real-time sentiment analysis. Real-time sentiment analysis needs to constantly ingest and process vast amounts of data coming from varied sources, including social media, news feeds, and forums. Stacked-LSTM models are ideal for time-series data; therefore, it can sequentially process the streams of incoming data. Real-time analysis should be processed with little or no latency at a very high speed. Although LSTMs are sequential, preprocessing steps (such as tokenization and embedding) can be batched or parallelized to minimize latency. Real-time data is often noisy (e.g., typos, slang, and spam) and often irregular or incomplete. Preprocessing pipelines clean, tokenize, and normalize text before feeding it into the model. A pretty significant thing for this presented model is that it can process multilingual and cross-platform social media data. It supports sequence and noisy text data processing, learns contextual dependencies over time, can operate with mixed-language or code-switched data, adapts to platform-specific linguistic features, scales to large datasets, and adjusts to evolving trends. This could be achieved by combining LSTMs with multilingual embeddings, fine-tuning, and effective preprocessing for providing accurate and robust sentiment analysis across diverse languages, platforms, and communication styles.

Comparative analysis

This section presents a deep comparative analysis of the proposed model with existing ensemble techniques and previous state-of-the-art models. The structure, training methods, and optimization techniques used by an optimized stacked Long Short-Term Memory model distinguish it mainly from the traditional LSTM models.

- A general LSTM is normally a single LSTM layer where the model gets to treat the whole input sequence in one go. That said, it successfully captures temporal dependencies and patterns but restricts its ability toward learning complex features because of only one layer. An optimized stacked LSTM has multiple LSTM layers stacked together. This design allows the model to capture more intricate hierarchical features at different levels of abstraction. Each layer learns higher-level representations of the input data, allowing the model to grasp complex patterns over time.
- Traditional LSTMs rely mostly on traditional training algorithms, like BPTT, to optimize the weights and use a simple fixed learning rate and gradient descent techniques, which suffer from vanishing gradients when applied to really long sequences. On the other hand, optimized stacked LSTMs use advanced optimization techniques, like learning rate schedules, gradient clipping, or adaptive learning rates (such as Adam or RMSprop), which can solve problems such as vanishing gradients or exploding gradients and accelerate the convergence process. Dropout can be applied between the layers to decrease overfitting, and various other regularization techniques can be added.
- Traditional LSTMs are more appropriate for less complex sequences or tasks where the temporal relationships are not too deep or complicated. They are not efficient at capturing long dependencies across huge datasets or very complex sequences. Optimized stacked LSTMs are more suitable for complex data such as cryptocurrency price prediction or natural language processing, as they have deeper layers and more refined training methods. They have a greater ability to survive against multi-step forecasting or very noisy data since deeper models tend to be more effective at noise filtering.

LSTM models are preferred for implementation on different domains and their respective datasets, especially when applied in tasks that deal with tasks like sentiment analysis, sequence prediction, and other areas of NLP, primarily because of its ability to capture temporal dependencies and long-range relationships in the data. Designed based on sequential input, this particular Staked-LSTM model significantly outperforms other models at large in cases involving datasets where an order of some data points applies, for instance, words in a phrase, phrases in an article, and timestamps. Specifically applied to natural language processing, any context and meanings depend upon what words mean exactly in their contexts. Inputted sequences are assumed to have the same length. This flexibility is important when handling datasets where the length of input data, for example, reviews, sentences, or documents, isn't fixed. It is very capable of learning domain-specific vocabulary and patterns within the text. While traditional models may require manual feature extraction or domain-specific rules, it can learn these patterns directly from the data.

Comparative analysis with popular ensembles

Huge experiments have been conducted on the similar "CryptocurrencyTweetLabelled" dataset to validate the performance of the proposed optimized stacked-LSTM model. Several popular ensemble models have been trained and tested on the same dataset to compare the accuracy with the proposed model. Table 6 presents the comparative results of the ensemble techniques with the proposed model.

Models like AdaBoost, Gradient Boosting, CatBoost, and Linear SVC are the most frequently used models in state-of-the-art research. These models handle complex patterns with high accuracy in a wide range of tasks. These models combine multiple weak learners to produce a strong predictive model that improves the overall performance and robustness of the system. In several domains of research, these models have produced exceptional results, especially for structured data problems, such as classification and regression tasks. In this study, the performance of these ensemble models is compared to the proposed Optimized Stacked-LSTM model, a deep learning-based approach designed to take advantage of sequential dependencies and learn complex patterns within the data. The Stacked-LSTM model involves several layers of LSTM units, which have been shown to be very effective when dealing with time-series data or data where features are highly intertwined. The optimization applied to the Stacked-LSTM model further improves its ability to learn from the data, making it more accurate in tasks like sentiment analysis, sequence prediction, or any problem involving temporal dependencies. Compared to the above ensemble models, the Optimized Stacked-LSTM model outperforms them.

Table 5 presents the comparative Precision, Recall, and F1-Score of the proposed model with popular machine learning ensemble models. In the experiments, Gradient Boosting and Linear SVC achieve higher Precision = 0.85, Recall = 0.86, and F1-Score = 0.85. Still, the proposed model acquires par results than existing ensembles precision = 0.91, Recall = 0.91, and F1-Score = 0.90, which is 6% more for Precision, 5% more for Recall, and 5% for F1-Score.

Figure 7 presents the proposed model's comparative training and testing accuracy with popular ensembles. It is visualized that Linear-SVC acquires higher training = 96% and testing = 86% accuracies but 4% less training and 5% less testing accuracy than the proposed optimized stacked-LSTM. Hence, IT has to be stated that the proposed deep learning-based ensemble outperforms existing machine learning ensembles and introduced the best optimized stacked-LSTM ensemble for cryptocurrency Tweets classification.

Comparative analysis with existing state-of-the-arts

This section presents the comparative analysis of the proposed Optimized Stacked LSTM-Model with the existing state of the arts sentiment analysis approaches. Table 7 presents the comparative discussion of the models. Optimized Stacked LSTM-Model is robust for complex, noisy datasets in a variety of domain-specific sentiment tasks. Optimization techniques also enhance the robustness and generalizability of the model. The analysis reveals that the proposed approach is more scalable and adaptable for real-world sentiment analysis applications. In a nutshell, the Optimized Stacked LSTM-Model sets a new benchmark for future research and application in this area.

It compares the performance and capabilities of five models designed for emotion recognition and sentiment analysis across diverse applications.

Dataset	Model	Label	Precision	Recall	F1-Score
Cryptocurrency-Tweet-Labelled	**Ada-Boost**	Negative	0.75	0.40	0.52
		Positive	0.73	0.99	0.84
		Neutral	0.96	0.65	0.77
		Weighted-AVG	0.83	0.79	0.79
	Gradient-Boosting	Negative	0.71	0.43	0.54
		Positive	0.84	0.95	0.89
		Neutral	0.90	0.85	0.87
		Weighted-AVG	0.85	0.86	0.85
	Cat-Boost	Negative	0.77	0.29	0.42
		Positive	0.80	0.97	0.88
		Neutral	0.91	0.79	0.85
		Weighted-AVG	0.84	0.83	0.82
	Linear-SVC	Negative	0.73	0.44	0.55
		Positive	0.84	0.95	0.89
		Neutral	0.90	0.85	0.87
		Weighted-AVG	0.85	0.86	0.85
	Proposed Swarm Optimized Stacked-LSTM	Negative	0.81	0.61	0.70
		Positive	0.95	0.93	0.94
		Neutral	0.88	0.95	0.91
		Weighted-AVG	0.91	0.91	0.90

Table 6. Comparative Results of the Proposed Optimized Stacked-LSTM with Ensemble Techniques.

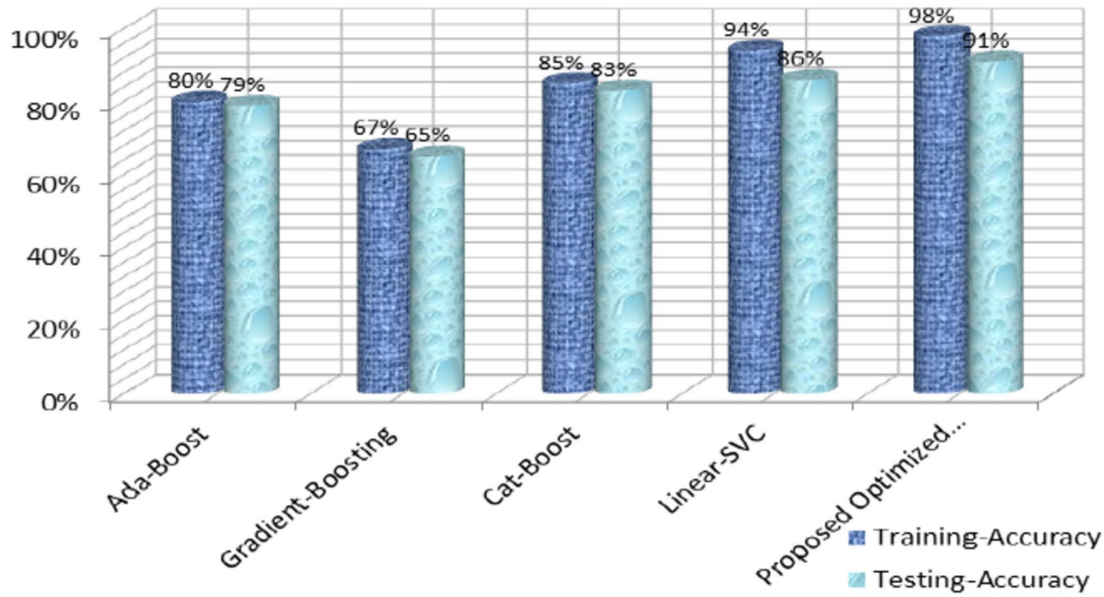

Fig. 7. Comparative Accuracy of the Proposed Optimized Stacked-LSTM with Ensemble Techniques.

Author	Year	Methodology	Model	Results
Ding et al. [32]	2023	Dialogue-INAB is the neural network based on social psychology's attitude-behavior theory, simulating the dynamic interplay between attitudes and speech behaviors in conversations and improving dialogue emotion recognition	DialogueINAB (Interaction Neural Network based on Attitude and Behaviour	**Accuracy**- 60.52% **Weighted F1**- 57.58%
Zhu et al. [33]	2023	An intelligent assistant system is proposed that utilizes emotional intelligence in order to personalize services and recommendations on the basis of users' emotions and behavior data. It will enhance user satisfaction and experience through giving insights into users' needs, both personal and work contexts	Emotion Recognition-Based Smart Assistant System	Accuracy- 85%
Ye et al. [34]	2024	Graph Emotion-Net: A novel EEG emotion recognition model enhances the accuracy using a spatiotemporal mechanism of attention and transfer learning, captures EEG-channel relationships using an adaptive graph, and optimizes spatial–temporal graph convolutions to better classify emotions	Graph Emotion-Net	Accuracy-94.68%
Song et al. [35]	2024	The FMSA-SC dataset offers fine-grained multimodal sentiment annotations for 1,247 stock commentary videos with textual, visual, and acoustic modalities. A novel multimodal multi-task framework is proposed as a baseline for sentiment analysis using this dataset	FMSA-SC (Fine-Grained Multimodal Sentiment analysis for Stock Common Videos	Accuracy- 76.92% F1-Score- 76.62%
Tiwari et al	2025	This paper presents an optimized stacked LSTM model for sentiment analysis to predict cryptocurrency prices, using PSO to optimize hyperparameters. The model outperforms existing approaches and aims to set a benchmark for sentiment analysis in cryptocurrency and broader societal predictions	SOSL (Swarm Optimization Stacked-LSTM	Accuracy- **98%** Weighted F1- **91%**

Table 7. Comparative Results Discussion of the Proposed Model with Existing State-of-the-Arts.

Dialogue-INAB, using the attitude-behavior theory of social psychology, reaches 60.52% accuracy and a weighted F1-score of 57.58%. The fact that it can mimic interactive dynamics of attitudes and speech behaviors makes it promising, though the relatively modest performance indicates the need for further refinement. The Emotion Recognition-Based Smart Assistant System reaches an accuracy of 85%, which demonstrates its practical utility in identifying the emotional states of users and offering personalized services. This system improves satisfaction and enhances the relevance for everyday and professional use by user behavior analysis. GraphEmotionNet outperforms many models with 94.68% accuracy. Through the use of the spatiotemporal attention mechanism and transfer learning, it has been able to capture intrinsic relationships that exist between EEG channels, for emotion recognition through adaptive graph construction. FMSA-SC, a fine-grained multimodal sentiment analysis framework, achieves the accuracy of 76.92% and an F1-score of 76.62%. It is notable for its detailed annotations, aligning textual, visual, and acoustic modalities at the phrase level, making it a valuable dataset for sentiment analysis in stock commentary videos. At last, the proposed SOSL, Swarm Optimization Stacked-LSTM, delivers the best result for sentiment analysis tasks, achieving training accuracy of 98%, a testing accuracy of 91%, and a weighted F1-score of 91%. Hyperparameter tuning by the PSO is applied, and the addition of such technique will make learning even stronger for it, establishing a good benchmark for cryptocurrency price prediction and possible other societal sentiments predictions.

Findings of the study

The findings of this study, especially the use of optimized stacked-LSTM for sentiment analysis in cryptocurrency price prediction, can be integrated into practical applications and decision-making processes. Traders and investors can leverage the model's predictions to make more informed decisions regarding buying, selling, or holding cryptocurrencies. By incorporating sentiment data from social media and news, investors can gain early insights into market movements that could help mitigate risks or maximize returns.

Financial institutions and hedge funds use sentiment analysis to better assess market risk. They follow the flow of sentiment in social media and predict price volatility. Therefore, they can formulate hedging strategies about potential losses within the cryptocurrency market. Research findings can be used by market analysts to track and predict patterns in cryptocurrency market trends. It helps business or investor communities understand shifts in social sentiment about the next direction of markets so that decision making can be made more in advance. Firms related to the cryptocurrency world; including the exchanges and the wallet provider firms, use the sentiment analysis technique to gauge customers' attitudes real-time. They can monitor a possible crisis or customer needs in terms of their opinion over products. Many optimization parameters have been implemented in the proposed optimized stacked-LSTM model, such as hyperparameter tuning, attention mechanism, data augmentation, pretrained embeddings, and regularization. Furthermore, Combining LSTM with other architectures, such as CNNs for feature extraction or transformers for improved context representation, may lead to better performance.

Conclusions and future scope

With the incremental growth in social media, sentiment analysis is widely adopted by people to make crucial decisions. Purchasing Cryptocurrency is a significant task, where people select sentiment analysis as a tool for better decisions. This paper proposes a novel model for practical sentiment analysis that can produce effective and authenticated results for selecting the best Cryptocurrency. The PSO optimization is applied to select the best parameters of the LSTM neural network. After that, the various LSTM were fine-tuned and stacked to achieve higher accuracy for sentiment classification. Advanced feature extraction as PoS tagging, word cluster calculation, semantic orientation, and min–max normalization, has been performed to enhance the quality of the Tweets before feeding the PSO-optimized Stacked-LSTM model. Glove word embedding is applied to generate the global word's co-occurrence matrix. Experiments are conducted on Cryptocurrency price-associated time-series datasets and cryptocurrency-related Tweets. The proposed model acquires 98% training accuracy, 91% testing accuracy, Precision, recall, and 90% F1-Score for the cryptocurrency Tweets dataset. Additionally, experiments are also conducted on the cryptocurrency price dataset, where the proposed model obtains 0.0441 MAE and 0.0039 MSE. Evaluated results prove the proposed model's effectiveness for classification and regression problems. Existing ensemble models have also been applied to the same dataset to compare the accuracy of the proposed model. Here proposed PSO-optimized Stacked-LSTM model achieves 5% more accuracy, 6% more Precision, 5% more Recall, and 5% more F1-Score from the existing robust ensemble model. Hence, the proposed PSO-optimized Stacked-LSTM model has the capability for better decision-making by selecting sentiment analysis as a tool. Adaptation of the Stacked-LSTM model in various other financial prediction tasks can only take place after careful tuning of its architecture, input features, and training process for a particular requirement of the new domain; hence, the model should effectively generalize and perform well on different scenarios of financial prediction, considering task-specific preprocessing, domain knowledge, and robust evaluation methods. Future work aims to extend the current work and develop the model to classify multilingual Tweets. Additionally, federated learning will incorporate into the present introduced collaborative and secure learning model. Aadditional data sources or metrics would significantly improve the model's accuracy and applicability. More data will integrate in future work that would not only provide more accurate model but also more adaptable across different market conditions and cryptocurrencies.

Data availability

The data that support the findings of this study are openly available at: Cryptocurrencies-related tweets dataset: https://www.kaggle.com/kaushiksuresh147/bitcoin-tweets Cryptocurrency time-series dataset: https://coinmarketcap.com/

Received: 10 December 2024; Accepted: 28 February 2025
Published online: 08 March 2025

References

1. Kim, S. Incentive mechanism-based influential maximisation scheme for social cloud service networks. *Int. J. Ad Hoc Ubiquitous Comput.* **33**(1), 11–21 (2020).
2. Aye, C. M., Wansaseub, K., Kumar, S., Tejani, G. G., Bureerat, S., Yildiz, A. R., & Pholdee, N. Airfoil Shape Optimisation Using a Multi-Fidelity Surrogate-Assisted Metaheuristic with a New Multi-Objective Infill Sampling Technique. CMES-Computer Modeling in Engineering & Sciences, 137(3) (2023)
3. Tiwari, D. & Nagpal, B. Enhanced long short-term memory (ELSTM) model for sentiment analysis. *Int. Arab J. Inf. Technol.* https://doi.org/10.34028/iajit/18/6/12 (2021).
4. Yang, J., Xu, W., Ma, K. & Li, C. A three-stage multi-energy trading strategy based on P2P trading mode. *IEEE Trans. Sustain. Energy* **14**(1), 233–241 (2022).
5. Tiwari, D., Bhati, B. S., Al-Turjman, F. & Nagpal, B. Pandemic coronavirus disease (Covid-19): World effects analysis and prediction using machine-learning techniques. *Expert Syst.* **39**, e12714 (2021).
6. Pan, S., Xu, G. J., Guo, K., Park, S. H. & Ding, H. Cultural insights in souls-like games: analyzing player behaviors, perspectives, and emotions across a multicultural context. *IEEE Trans. Games* https://doi.org/10.1109/TG.2024.3366239 (2024).
7. Liu, X., Wang, G., Cai, Z. & Zhang, H. Bagging based ensemble transfer learning. *J. Ambient. Intell. Humaniz. Comput.* **7**(1), 29–36 (2016).
8. He, Q., Xia, P., Hu, C., & Li, B. PUBLIC INFORMATION, ACTUAL INTERVENTION AND INFLATION EXPECTATIONS. Transformations in Business & Economics, 21 (2022)

9. Tejani, G., Savsani, V. & Patel, V. Modified sub-population based heat transfer search algorithm for structural optimization. *Int. J. Appl. Metaheuristic Comput. (IJAMC)* **8**(3), 1–23 (2017).
10. Pari, R., Sandhya, M. & Sankar, S. A multitier stacked ensemble algorithm for improving classification accuracy. *Comput. Sci. Eng.* **22**(4), 74–85 (2018).
11. Boiy, E. & Moens, M. F. A machine learning approach to sentiment analysis in multilingual Web texts. *Inf. Retrieval* **12**(5), 526–558 (2009).
12. Pasupa, K. & Ayutthaya, T. S. N. Thai sentiment analysis with deep learning techniques: A comparative study based on word embedding, POS-tag, and sentic features. *Sustain. Cities Soc.* **50**, 101615 (2019).
13. Alam, S. & Yao, N. The impact of pre-processing steps on the accuracy of machine learning algorithms in sentiment analysis. *Comput. Math. Organ. Theory* **25**(3), 319–335 (2019).
14. El Alaoui, I. et al. A novel adaptable approach for sentiment analysis on big social data. *J. Big Data* **5**(1), 1–18 (2018).
15. Khoo, C. S. & Johnkhan, S. B. Lexicon-based sentiment analysis: Comparative evaluation of six sentiment lexicons. *J. Inf. Sci.* **44**(4), 491–511 (2018).
16. Roy, S. S., Chopra, R., Lee, K. C., Spampinato, C. & Mohammadi-ivatlood, B. Random forest, gradient boosted machines and deep neural network for stock price forecasting: A comparative analysis on South Korean companies. *Int. J. Ad Hoc Ubiquitous Comput.* **33**(1), 62–71 (2020).
17. Ghosh, M. & Sanyal, G. An ensemble approach to stabilize the features for multi-domain sentiment analysis using supervised machine learning. *J. Big Data* **5**(1), 1–25 (2018).
18. Al–Azani, S., & El–Alfy, E. S. M. Imbalanced sentiment polarity detection using emoji-based features and bagging ensemble. In 2018 1st International Conference on Computer Applications & Information Security (ICCAIS) (pp. 1–5). IEEE (2018)
19. Sharma, A., & Dey, S. A boosted SVM based sentiment analysis approach for online opinionated text. In Proceedings of the 2013 research in adaptive and convergent systems (pp. 28–34) (2013)
20. Akhtar, M. S., Gupta, D., Ekbal, A. & Bhattacharyya, P. Feature selection and ensemble construction: A two-step method for aspect based sentiment analysis. *Knowl. Based Syst.* **125**, 116–135 (2017).
21. Rong, W., Nie, Y., Ouyang, Y., Peng, B. & Xiong, Z. Auto-encoder based bagging architecture for sentiment analysis. *J. Vis. Lang. Comput.* **25**(6), 840–849 (2014).
22. Al-Salemi, B., Ayob, M. & Noah, S. A. M. Feature ranking for enhancing boosting-based multi-label text categorization. *Expert Syst. Appl.* **113**, 531–543 (2018).
23. Ahuja, R., Chug, A., Kohli, S., Gupta, S. & Ahuja, P. The impact of features extraction on the sentiment analysis. *Proc. Comput. Sci.* **152**, 341–348 (2019).
24. Hama Aziz, R. H. & Dimililer, N. SentiXGboost: Enhanced sentiment analysis in social media posts with ensemble XGBoost classifier. *J. Chin. Inst. Eng.* **44**(6), 562–572 (2021).
25. Athanasiou, V. & Maragoudakis, M. A novel, gradient boosting framework for sentiment analysis in languages where NLP resources are not plentiful: A case study for modern Greek. *Algorithms* **10**(1), 34 (2017).
26. Alzamzami, F., Hoda, M. & El Saddik, A. Light gradient boosting machine for general sentiment classification on short texts: A comparative evaluation. *IEEE Access* **8**, 101840–101858 (2020).
27. Tejani, G. G., Mashru, N., Patel, P., Sharma, S. K. & Celik, E. Application of the 2-archive multi-objective cuckoo search algorithm for structure optimization. *Sci. Rep.* **14**(1), 31553 (2024).
28. Tiwari, D., Bhati, B. S., Nagpal, B., Sankhwar, S. & Al-Turjman, F. An enhanced intelligent model: To protect marine IoT sensor environment using ensemble machine learning approach. *Ocean Eng.* **242**, 110180 (2021).
29. Nonut, A. et al. A small fixed-wing UAV system identification using metaheuristics. *Cog. Eng.* **9**(1), 2114196 (2022).
30. Trueman, T. E. & Cambria, E. A convolutional stacked bidirectional LSTM with a multiplicative attention mechanism for aspect category and sentiment detection. *Cognit. Comput.* **13**, 1423–1432 (2021).
31. Tiwari, D. & Nagpal, B. KEAHT: A knowledge-enriched attention-based hybrid transformer model for social sentiment analysis. *New Gener. Comput.* **40**(4), 1165–2120 (2022).
32. Ding, J. et al. DialogueINAB: an interaction neural network based on attitudes and behaviors of interlocutors for dialogue emotion recognition. *J. Supercomput.* **79**(18), 20481–20514 (2023).
33. Zhu, C. Research on emotion recognition-based smart assistant system: Emotional intelligence and personalized services. *J. Syst. Manage. Sci.* **13**(5), 227–242 (2023).
34. Ye, W. et al. Adaptive spatial-temporal aware graph learning for EEG-based emotion recognition. *Cyborg Bionic Syst.* **5**, 0088 (2024).
35. Song, L., Chen, S., Meng, Z., Sun, M. & Shang, X. FMSA-SC: A Fine-grained Multimodal Sentiment Analysis Dataset based on Stock Comment Videos. *IEEE Trans. Multimedia* https://doi.org/10.1109/TMM.2024.3363641 (2024).

Acknowledgements
This research was financially supported by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2025R235), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Author contributions
All authors contributed equally to the conceptualization, formal analysis, investigation, methodology, and writing and editing of the original draft. All authors have read and agreed to the published version of the manuscript.

Funding
This research was financially supported by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2025R235), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to A.A.-R. or M.G.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.