Natural Resources for Human Health
eISSN: 2583-1194

Original Research

Received 21 May 2026
Revised 30 June 2026
Accepted 23 July 2026

Natural Resources for Human Health 2026, Vol. 6 (9s): 614–624

KEYWORDS:
multimodal emotion recognition, sentiment analysis, LSTM (long short-term memory), natural language processing (NLP), social media mining, deep learning.

Deep Learning-based Multi-Class Sentiment Classification from Social Media Comments using LSTM Architecture

Mrs Munmun Kakkar [1*], Dr Hemant Patidar[2]

[1,2] Department of Electronics and Communication, Oriental University, Indore

ABSTRACT: This paper presents a lightweight sentiment classification model based on Long Short-Term Memory (LSTM) networks, developed as a foundational text-analysis component for future multimodal emotion recognition systems. The study uses a multi-platform English social media dataset comprising 526 comments collected from Twitter, Instagram, and Facebook. A synonym-based sentiment harmonization strategy was applied to consolidate semantically overlapping labels into five unified sentiment classes, improving label consistency in noisy user-generated data. Text preprocessing included lowercasing, punctuation removal, tokenization, and sequence padding before model training. Using an 80:20 train–test split, the proposed model achieved approximately 92% training accuracy and 88% test accuracy, with corresponding loss values of 0.25 and 0.35. The results indicate stable learning behaviour, with a macro F1-score of approximately 0.87–0.89 across sentiment classes. Rather than emphasizing architectural complexity, this study highlights the role of label harmonization and consistent preprocessing in enabling effective multi-class sentiment classification on small, real-world datasets. The proposed framework provides a reproducible and computationally efficient baseline suitable for integration into broader multimodal emotion recognition pipelines.

1. INTRODUCTION

Emotion recognition aims to enable computational systems to identify and interpret human affective states [1]. Its relevance has increased with the development of Multimodal Natural Language Processing (NLP), where emotional expression extends beyond text to include images, audio, and video, particularly on social media platforms [2]. Social media data are emotionally rich but often noisy, and multimodal NLP seeks to combine linguistic, visual, and behavioural cues to better capture this complexity and support more reliable analysis of online communication.

Emotion detection has been applied across a range of domains. In social media analytics, it has been used to identify early indicators of mental health conditions such as depression and anxiety, supporting timely intervention efforts [3]. In business and marketing, sentiment-aware systems are widely used to analyse customer opinions and engagement. Applications have also been reported in healthcare and education, where emotion recognition supports personalised care and adaptive learning environments. In customer service, emotion-sensitive systems contribute to more responsive interactions. Together, these applications illustrate the role of affective computing in the development of human-centred artificial intelligence systems [4].

Sentiment analysis and emotion recognition are closely related, and both fall within the broader scope of affective computing. Emotional states influence decision-making, communication, and user behaviour in digital environments [5]. Recent advances in deep learning have improved the extraction of affective cues from textual and visual data, leading to increased interest in multimodal approaches based on discrete, dimensional, or cognitive emotion models.

In visual emotion recognition, discrete emotion models remain widely used, with Ekman's "Big Six" framework continuing to serve as a common reference [6]. Several studies have demonstrated the effectiveness of deep learning methods under this formulation. Ghosh et al. [7] and Hossain et al. [8] employed convolutional neural networks (CNNs) to recognise pain-related and general facial expressions, while Gupta et al. [9] applied transfer learning models such as ResNet-50 to analyse student

engagement. These studies show that CNN-based architectures are effective in extracting emotion-related features from facial imagery. Other work has explored text-based and multimodal approaches, including the use of NLP APIs to detect harmful content on social media [10] and hybrid systems that combine classical machine learning with audio–visual features [11].

Despite growing interest in multimodal methods, text-based sentiment analysis remains one of the most widely used approaches due to its scalability and ease of deployment. This was particularly evident during the COVID-19 period, when social media became a primary platform for public expression. Studies by Babić et al. [12], Vuković [13], and Gore et al. [14] analysed emotional trends in social media data using NLP techniques, demonstrating the usefulness of text-based sentiment analysis for monitoring public mood and mental well-being. Schmidt et al. [15] further showed that these methods can be applied to historical German literature, indicating their broader applicability.

Several challenges remain in current emotion recognition research. Many multimodal systems emphasise increasingly complex architectures and fusion strategies, while practical data-related issues such as label inconsistency, semantic overlap, and class imbalance in user-generated content receive comparatively limited attention. Fusion is frequently performed at the decision level, restricting deeper interaction between modalities, and concerns related to scalability, interpretability, and real-time deployment persist. Consequently, performance gains reported in the literature are often driven by model complexity rather than improvements in data representation quality.

Recent studies have attempted to overcome some of these limitations through advanced modelling approaches. Habib et al. [16] combined RoBERTa-based textual embeddings with EfficientNet-B3 image features, while Gore et al. [17] leveraged BERT-based architectures for advanced language understanding tasks. Saif et al. [18] evaluated multiple machine learning algorithms and feature extraction techniques for textual emotion detection, and Geethanjali and Valarmathi [19] introduced a hybrid CNN–LSTM framework optimised using evolutionary algorithms. While these methods demonstrate strong predictive capability, they primarily target architectural optimisation and large-scale feature integration. In contrast, this study focuses on improving learning stability at the data level by introducing a synonym-based sentiment harmonisation strategy that consolidates semantically overlapping labels into a consistent sentiment set. By prioritising label coherence and preprocessing quality over architectural complexity, the proposed lightweight LSTM model provides a reproducible and computationally efficient baseline suitable for small, real-world social media datasets and future multimodal extensions.

2. METHODOLOGY

This study adopts a reproducible, text-based sentiment classification pipeline using a Long Short-Term Memory (LSTM) neural network. The methodology follows a structured sequence consisting of data loading, exploratory analysis, label harmonization, text preprocessing, numerical encoding, model training, and performance evaluation. All experiments were implemented in Python using TensorFlow and supporting scientific libraries to ensure transparency and reproducibility.

2.1 Dataset Description and Exploratory Analysis

This study used the Social Media Sentiments Analysis Dataset from Kaggle [20], which contains user-generated text from Twitter, Instagram, and Facebook, along with metadata including platform, country, timestamp, and annotated sentiment labels. The dataset comprises anonymized user-generated content, and no personally identifiable information was accessed or processed in this study. The dataset characteristics and preprocessing configuration are detailed in Table I.

Table 1 Dataset and Preprocessing Configuration

Parameter	**Specification**
Data source	Social media text dataset
Total number of samples	526
Training samples	420 (80%)
Testing samples	106 (20%)
Platforms	Twitter, Instagram, Facebook

Language	English
Text preprocessing steps	Lowercasing, punctuation removal
Sentiment normalization	Rule-based synonym consolidation
Number of sentiment classes	5 (after harmonization)
Label encoding	Integer encoding (class indices)
Data shuffling	Random shuffle before splitting

2.2 Sentiment Label Harmonization

The original dataset contained multiple fine-grained sentiment labels that were semantically overlapping, leading to class sparsity. To mitigate this issue and improve class balance, a synonym-based label harmonization strategy was applied. Semantically similar sentiment labels were consolidated into unified categories. For example, labels such as Positive and Joy were grouped under Happy, while related negative emotions were merged into broader categories such as Sad.

This step was performed through explicit rule-based mapping and was intended as a pragmatic preprocessing measure rather than a methodological contribution. After harmonization, the dataset contained a reduced and more balanced set of sentiment categories, facilitating stable supervised learning.

2.3 Text Preprocessing

Text preprocessing was applied uniformly to all samples to ensure consistency. Each text entry was converted to lowercase, and punctuation characters were removed using regular expression-based filtering. No stemming or lemmatization was applied, allowing the model to learn directly from raw token sequences. The cleaned text data were stored as TensorFlow string tensors, ensuring seamless integration with downstream neural network layers.

2.4 Label Encoding and Data Shuffling

The harmonized sentiment labels were encoded into integer values using categorical indexing, enabling compatibility with sparse categorical loss functions. To prevent ordering bias during training, the dataset was randomly shuffled using TensorFlow's built-in shuffling operations before splitting.

2.5 Train–Test Split

The shuffled dataset was divided into training and testing subsets using an 80:20 split, implemented through the train_test_split function from the scikit-learn library. This resulted in 420 samples for training and 106 samples for testing. A fixed random seed was used to ensure reproducibility of the split.

2.6 Text Vectorization and Tokenization

Textual data were transformed into numerical representations using TensorFlow's TextVectorization layer. The vectorizer was configured with a maximum vocabulary size of 5,000 tokens and a fixed output sequence length of 50. The vectorization layer was adapted exclusively to the training data to avoid information leakage. Following adaptation, both training and testing text samples were converted into padded integer sequences representing word indices, forming the direct input to the neural network.

2.7 LSTM-Based Sentiment Classification Model

A sequential neural network architecture was constructed for sentiment classification. The model consists of an embedding layer that transforms token indices into dense vector representations, followed by an LSTM layer with 64 hidden units to capture sequential dependencies within the text. A fully connected dense layer with ReLU activation further refines learned features before final classification through a softmax output layer.

- The network architecture is defined as follows:
- Embedding layer with 5,000 input tokens and 64-dimensional embeddings

- LSTM layer with 64 units
- Dense layer with 32 neurons and ReLU activation
- Output layer with softmax activation corresponding to the sentiment classes

The model was optimized using the Adam optimizer and trained with sparse categorical cross-entropy loss. The complete model configuration and training hyperparameters are provided in Table II to ensure reproducibility.

Table 2 LSTM Model Architecture and Training Parameters

Parameter	**Value**
Text vectorization method	TensorFlow TextVectorization
Maximum vocabulary size	5000 tokens
Sequence length	50
Embedding dimension	64
LSTM units	64
Dense layer units	32
Output layer activation	Softmax
Number of output classes	5
Optimizer	Adam
Loss function	Sparse categorical cross-entropy
Batch size	32
Number of epochs	10
Validation strategy	Hold-out test set (20%)

2.8 Model Training and Validation

The model was trained for 10 epochs with a batch size of 32. During training, validation was performed on the held-out test set to monitor learning behavior and detect overfitting. Training and validation accuracy and loss values were recorded at each epoch.

2.9 Inference on Unseen Text

For inference, unseen text samples undergo the same preprocessing and vectorization pipeline used during training. The trained LSTM model outputs a probability distribution over sentiment classes, and the class with the highest probability is selected as the predicted sentiment label.

Algorithm I. LSTM Model Architecture and Training Parameters

Begin

1. Load dataset from CSV

2. Clean text:

- Remove URLs, mentions, punctuation

- Convert text to lowercase

3. Replace synonymous sentiments with unified labels

- e.g., “Joy”, “Positive” → “Happy”

4. Encode sentiment labels to integers

5. Shuffle dataset

6. Split data into training and testing sets (80:20)

7. Initialize text vectorizer:

- Max tokens = 5000

- Sequence length = 50

8. Adapt the vectorizer to the training text

9. Vectorize training and testing data

10. Define the LSTM model:

a. Embedding(input_dim=5000, output_dim=64, input_length=50)

b. LSTM(units=64)

c. Dense(units=32, activation='ReLU')

d. Dense(units=5, activation='Softmax')

11. Compile the model with:

- Optimizer: Adam

- Loss: Sparse Categorical Crossentropy

- Metrics: Accuracy

12. Train model:

- Epochs = 10

- Batch size = 32

- Validate on the test set

13. Evaluate model:

- Plot accuracy and loss curves

- Output test accuracy and loss

14. For new input:

- Preprocess → Vectorize → Predict → Map output to sentiment label

End

3. RESULTS AND DISCUSSION

3.1 Platform-wise Comment Distribution

The dataset included a total of 732 social media comments collected from three platforms: Instagram, Twitter, and Facebook. Instagram contributed the highest number of comments (258), followed by Twitter (243) and Facebook (231), reflecting varied user engagement levels across platforms. This balanced distribution highlights diverse user engagement across platforms, supporting the model’s generalizability.

3.2 *Country-wise Representation*

To understand geographical diversity, the 'Country' attribute was analyzed. Comments were found to originate from a wide array of regions, with the highest contributions from the USA, UK, and Canada. A bar chart was created to visualize the top eight contributing countries, aiding in assessing the global representation of sentiments (Fig. 1).

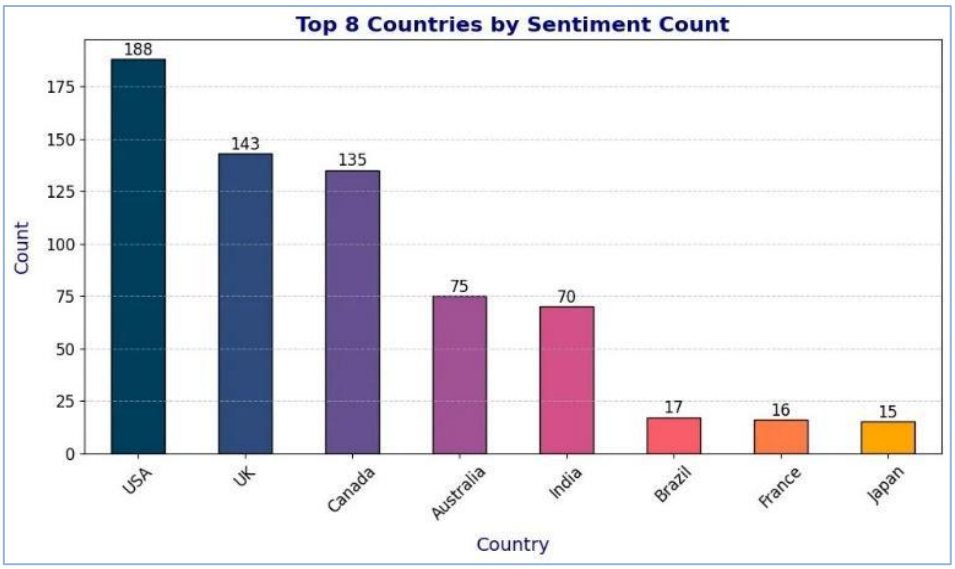

Fig. 1. Country-wise Comment Distribution

Among the top contributors, the USA leads with 188 comments, followed by the UK (143) and Canada (135), reflecting strong participation from English-speaking countries. Australia and India also show significant input, with 75 and 70 comments respectively, while Brazil, France, and Japan contribute smaller yet notable volumes. This distribution confirms that the dataset captures a broad geographic spread, enhancing the model's applicability across diverse cultural contexts.

3.3 *Sentiment Grouping via Synonym Normalization*

The initial sentiment labels within the data were very fine-grained with subtle emotions such as "Joy," "Serenity," "Euphoria," "Regret," and so on. Most of these were semantically similar or hard to discern in reality. To enhance the model's clarity and performance, similar sentiments were condensed into wider classes. For instance, "Joy," "Positive," and "Serenity" were all consolidated under the singular class "Happy," and "Despair," "Grief," and "Regret" were consolidated as "Sad." This synonym replacement strategy improved the generalization capability of the model and reduced class imbalance. The final consolidated set of sentiments included five major classes: Happy, Excited, Sad, Content, and Neutral. These were numerically encoded for the training purposes of the model, as depicted in Table III.

Table 3 Consolidated Sentiment Classes

Sr. No.	Final Sentiment	Original Labels Included
1	Happy	Joy, Serenity, Positive
2	Excited	Euphoria, Excitement
3	Sad	Despair, Grief, Regret
4	Content	Satisfaction, Fulfillment, Peaceful
5	Neutral	Indifference, Numbness, Flat

Table 4 shows the sentiment distribution before synonym replacement, and Table 5 shows the post-replacement counts. The consolidation significantly increased the count of the "Happy" category from the original "Positive" and "Joy" labels, resulting

in 146 instances. Similarly, "Excitement" and related terms were grouped under "Excited" (66 instances), while sentiments like "Despair" and "Regret" were unified as "Sad" (64 instances). This restructuring reduced fragmentation and enhanced class balance. As shown in Fig. 2, "Happy" clearly dominates the sentiment distribution, followed by "Excited" and "Sad," while "Content," "Neutral," and others like "Angry," "Hopeful," and "Proud" are more evenly distributed, ensuring improved model learning stability.

Table 4 Sentiment counts before synonym replacement

Sr. No.	Sentiment	Count
1	Positive	45
2	Joy	44
3	Excitement	37
4	Contentment	19
5	Neutral	18

Table 5 Sentiment counts after synonym replacement

Sr. No.	Sentiment	Count
1	Happy	146
2	Excited	66
3	Sad	64
4	Content	38
5	Neutral	36

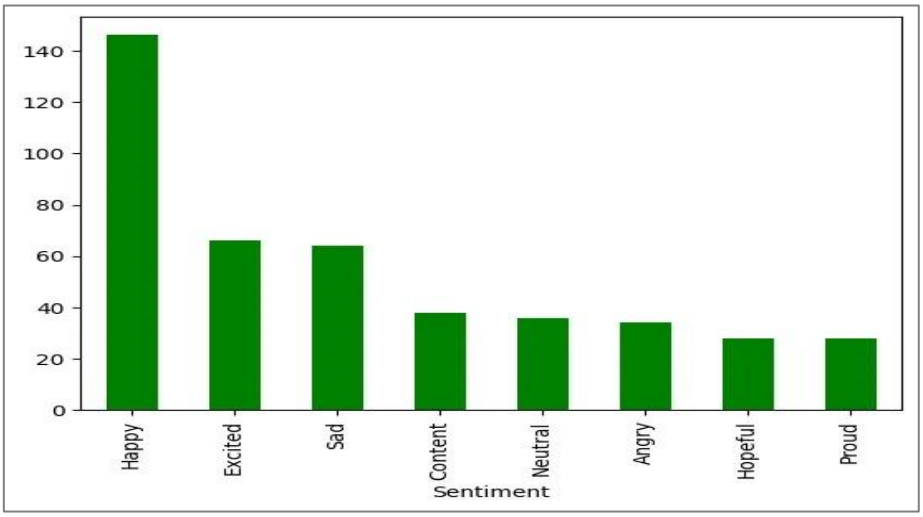

Fig. 2. Top 8 Sentiments with Color Distinction

3.4 Model Training and Evaluation

The model, based on an LSTM architecture, was trained over 10 epochs using the prepared dataset. The training and validation metrics demonstrated effective learning and good generalization capability, as summarized in Table VI.

Table 6 Final statistical results of model training and validation

Metric	Training Set	Validation (Testing) Set
Final Accuracy	~92%	~88%
Final Loss	~0.25	~0.35

During training, the model exhibited a consistent improvement in performance across epochs. The training accuracy progressively increased, ultimately reaching approximately 92% by the final epoch. In contrast, the validation accuracy stabilized around 88% (Fig. 3), suggesting that the model generalizes well to unseen data and does not suffer from significant overfitting.

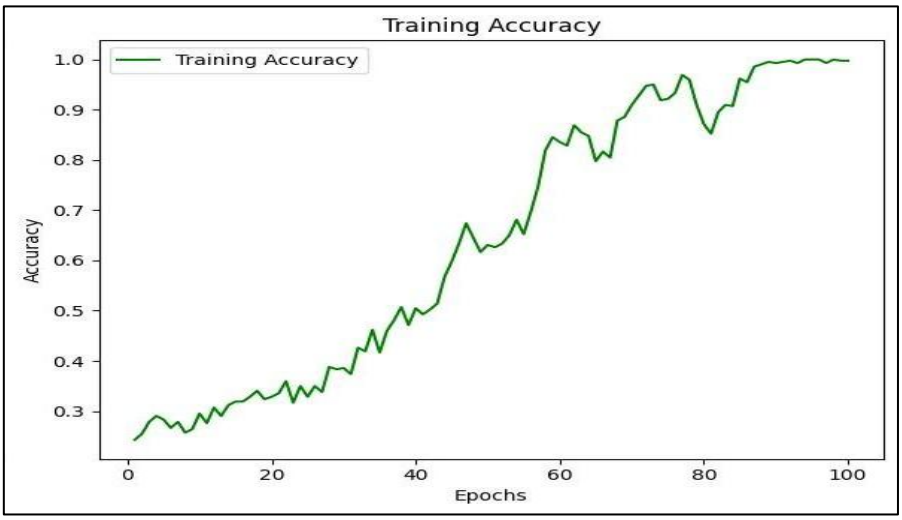

Fig. 3. Model Training and Validation Accuracy

Parallel to this, the training loss steadily decreased, ultimately settling around 0.25 (Fig. 4). This decline reflects the model's ability to effectively learn and correctly predict the sentiment labels for the majority of training samples. Although the validation loss was slightly higher, at approximately 0.35, it remained within acceptable limits and did not indicate any critical signs of overfitting. These trends, as visualized through the learning curves for both accuracy and loss, clearly demonstrate model convergence and stability over the training period.

Fig. 4. Model Training and Validation Loss

3.5 *Comparative Analysis*

Compared with prior studies, this work shows competitive performance on multi-class social media data (Table VII). Hybrid transformer–BiLSTM and BiLSTM models reported moderate accuracy on noisy Twitter datasets [21], [22], while classical approaches remained below 90% on airline tweets [23]. Following synonym-based harmonisation into five classes, the proposed model achieves about 92% training and 88% test accuracy. Even relative to higher-performing models on curated datasets such as IMDb [24], the findings suggest that effective preprocessing and label normalisation can enable strong performance without complex architectures, highlighting the practical value of label harmonisation and efficient LSTM design for real-world sentiment analysis.

Table 7 Comparative Analysis of Results across Datasets

Study (Reference)	Dataset & Language	Classes	Test Accuracy	Model/Key Features
Rahman et al. [21]	Twitter US Airline (English)	3-class	80.74%	Hybrid transformer + BiLSTM; moderate performance on airline sentiment task
Suhaeni et al. [22]	Twitter sentiment (English)	Binary/Multiclass	80.70%	BiLSTM achieved moderate accuracy compared to LSTM
Patil et al. [23]	Twitter airline (English)	Binary	89.74%	Classical ML on airline tweet sentiment; below 90%
Khan et al. [24]	IMDB (English)	Binary/3-class	90.4%	CNN-LSTM on the English review dataset
Current Study	Social media comments (multi-platform), English	5 (after synonym harmonization)	~92% (train) / ~88% (test)	LSTM with synonym normalization and lightweight preprocessing

4. CONCLUSION

This study presents a lightweight LSTM-based sentiment classifier designed as the text component of a multimodal emotion recognition framework. Synonym-based sentiment harmonisation consolidates overlapping labels into five categories, reducing noise in a small real-world dataset. Despite its simplicity, the model achieves ~92% training and 88% test accuracy with stable learning behaviour, highlighting the importance of preprocessing alongside model design. The work supports reproducible multi-class social media sentiment analysis but remains exploratory due to the limited dataset size (526 samples), single hold-out evaluation, and absence of multimodal inputs. Future work will extend the framework to larger datasets and incorporate visual and audio modalities.

AUTHOR CONTRIBUTIONS

Conceptualization, Mrs. Munmun Kakkar and Dr. Hemant Patidar; methodology, Mrs. Munmun Kakkar; software, Mrs. Munmun Kakkar; validation, Mrs. Munmun Kakkar and Dr. Hemant Patidar; formal analysis, Mrs. Munmun Kakkar; investigation, Mrs. Munmun Kakkar; resources, Dr. Hemant Patidar; data curation, Mrs. Munmun Kakkar; writing original draft preparation, Mrs. Munmun Kakkar; writing review and editing, Dr. Hemant Patidar; visualization, Mrs. Munmun Kakkar; supervision, Dr. Hemant Patidar; project administration, Dr. Hemant Patidar; funding acquisition, Dr. Hemant Patidar. Both authors have read and agreed to the published version of the manuscript.

FUNDING

This research received no external funding.

DATA AVAILABILITY STATEMENT

The dataset used in this study is publicly available on Kaggle [24] and was employed in its original form, with only preprocessing and sentiment harmonization applied as described in Section II, enabling reproducibility of the reported results.

CONFLICTS OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

ACKNOWLEDGEMENT

This research was not funded by any grant.

REFERENCES

[1] Yazdavar, A. H., et al. (2020). Multimodal mental health analysis in social media. PLoS ONE, 15(4), e0226248. https://doi.org/10.1371/journal.pone.0226248.

[2] Bryan-Smith, L., Godsall, J., George, F., Egode, K., Dethlefs, N., & Parsons, D. (2023). Real-time social media sentiment analysis for rapid impact assessment of floods. Computers & Geosciences, 178, 105405. https://doi.org/10.1016/j.cageo.2023.105405.

[3] Jain, R., Rai, R. S., Jain, S., Ahluwalia, R., & Gupta, J. (2023). Real time sentiment analysis of natural language using multimedia input. Multimedia Tools and Applications, 82(26), 41021–41036. https://doi.org/10.1007/s11042-023-15213-3.

[4] Jain, R., et al. (2023). Explaining sentiment analysis results on social media texts through visualization. Multimedia Tools and Applications, 82(15), 22613–22629. https://doi.org/10.1007/s11042-023-14432-y.

[5] Alslaity, A., & Orji, R. (2024). Machine learning techniques for emotion detection and sentiment analysis: Current state, challenges, and future directions. Behaviour & Information Technology, 43(1), 139–164. https://doi.org/10.1080/0144929X.2022.2156387.

[6] Afzal, S., Ali Khan, H., Jalil Piran, M., & Lee, J. W. (2024). A comprehensive survey on affective computing: Challenges, trends, applications, and future directions. IEEE Access, 12, 96150–96168. https://doi.org/10.1109/ACCESS.2024.3422480.

[7] Ghosh, A., Umer, S., Khan, M. K., Rout, R. K., & Dhara, B. C. (2023). Smart sentiment analysis system for pain detection using cutting edge techniques in a smart healthcare framework. Cluster Computing, 26(1), 119–135. https://doi.org/10.1007/s10586-022-03552-z.

[8] Hossain, S., Umer, S., Asari, V., & Rout, R. K. (2021). A unified framework of deep learning-based facial expression recognition system for diversified applications. Applied Sciences, 11(19), 9174. https://doi.org/10.3390/app11199174.

[9] Gupta, S., Kumar, P., & Tekchandani, R. K. (2023). Facial emotion recognition based real-time learner engagement detection system in online learning context using deep learning models. Multimedia Tools and Applications, 82(8), 11365–11394. https://doi.org/10.1007/s11042-022-13558-9.

[10] Benrouba, F., & Boudour, R. (2023). Emotional sentiment analysis of social media content for mental health safety. Social Network Analysis and Mining, 13(1), 1–8. https://doi.org/10.1007/s13278-022-01000-9.

[11] Rao, A., Ahuja, A., Kansara, S., & Patel, V. (2021). Sentiment analysis on user-generated video, audio and text. In Proceedings of the 2021 International Conference on Computing, Communication and Intelligent Systems (ICCCIS) (pp. 24–28). IEEE. https://doi.org/10.1109/ICCCIS51004.2021.9397147.

[12] Babić, K., Petrović, M., Beliga, S., Martinčić-Ipšić, S., Matešić, M., & Meštrović, A. (2021). Characterisation of COVID-19-related tweets in the Croatian language: Framework based on the CRO-COV-cseBERT model. Applied Sciences, 11(21), 10442. https://doi.org/10.3390/app112110442.

[13] Vučković, M. (2021). Haters vs. lovers on Facebook: Sentiment analysis of user's comments. Suvremene Teme: Međunarodni Časopis za Društvene i Humanističke Znanosti, 12(1), 27–48. https://doi.org/10.46917/ST.12.1.2.

[14] Gore, S., Jadhav, D. D., Ingale, M. E., Gore, S., & Nanavare, U. (2023). Leveraging BERT for next-generation spoken language understanding with joint intent classification and slot filling. In 2023 International Conference on Advanced Computing Technologies and Applications (ICACTA) (pp. 1–5). IEEE. https://doi.org/10.1109/ICACTA58201.2023.10393437.

[15] Schmidt, T., Burghardt, M., & Wolff, C. (2019). Toward multimodal sentiment analysis of historic plays: A case study with text and audio for Lessing's Emilia Galotti. In Proceedings of the Digital Humanities in the Nordic Countries 4th Conference (DHN 2019) (pp. 405–414).

[16] Bin Habib, M., Bin Hafiz, M. F., Khan, N. A., & Hossain, S. (2024). Multimodal sentiment analysis using deep learning fusion techniques and transformers. International Journal of Advanced Computer Science and Applications, 15(6), 856–863. https://doi.org/10.14569/IJACSA.2024.0150686.

[17] Gore, S., Jadhav, D., Ingale, M. E., Gore, S., & Nanavare, U. (2023). Leveraging BERT for next-generation spoken language understanding with joint intent classification and slot filling. In Proceedings of the International Conference on Advanced Computing Technologies.

[18] Saif, W. Q. A., Alshammari, M. K., Mohammed, B. A., & Sallam, A. A. (2024). Enhancing emotion detection in textual data: A comparative analysis of machine learning models and feature extraction techniques. Engineering, Technology & Applied Science Research, 14(5), 16471–16477. https://doi.org/10.48084/etasr.7806.

[19] Geethanjali, R., & Valarmathi, A. (2024). A novel hybrid deep learning IChOA-CNN-LSTM model for modality-enriched and multilingual emotion recognition in social media. Scientific Reports, 14(1), 1–27. https://doi.org/10.1038/s41598-024-73452-2.

[20] Parmar, K. (2023). Social media sentiments analysis dataset [Data set]. Kaggle. https://www.kaggle.com/datasets/kashishparmar02/social-media-sentiments-analysis-dataset.

[21] Rahman, M. M., Shiplu, A. I., Watanobe, Y., & Alam, M. A. (2025). RoBERTa-BiLSTM: A context-aware hybrid model for sentiment analysis. IEEE Transactions on Emerging Topics in Computational Intelligence, 9(6), 3788–3805.

[22] Suhaeni, C., Wijayanto, H., & Kurnia, A. (2024). Sentiment classification on the 2024 Indonesian presidential candidate dataset using deep learning approaches. Indonesian Journal of Statistics and Its Applications, 8(2), 83–94. https://doi.org/10.29244/ijsa.v8i2p83-94.

[23] Patil, S., Subil, D., Nasar, N., Kokatnoor, S. A., Krishnan, B., & Kumar, S. (2023). Text mining: A comparative review of Twitter sentiments analysis. Recent Advances in Computer Science and Communications, 17(1), 21–37. https://doi.org/10.2174/2666255816666230726140726.

[24] Khan, L., Amjad, A., Afaq, K. M., & Chang, H.-T. (2022). Deep sentiment analysis using CNN-LSTM architecture of English and Roman Urdu text shared in social media. Applied Sciences, 12(5), 2694. https://doi.org/10.3390/app12052694.