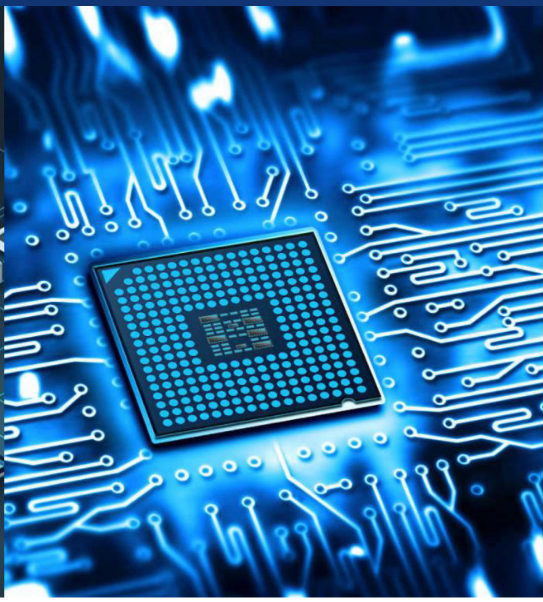
IJIRCCE

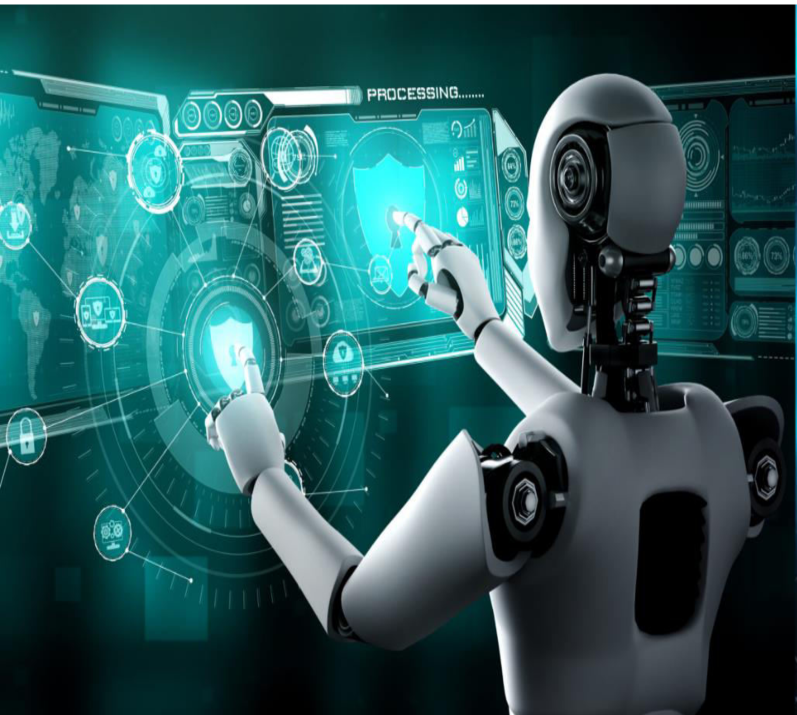
ISSN(O): 2320-9801

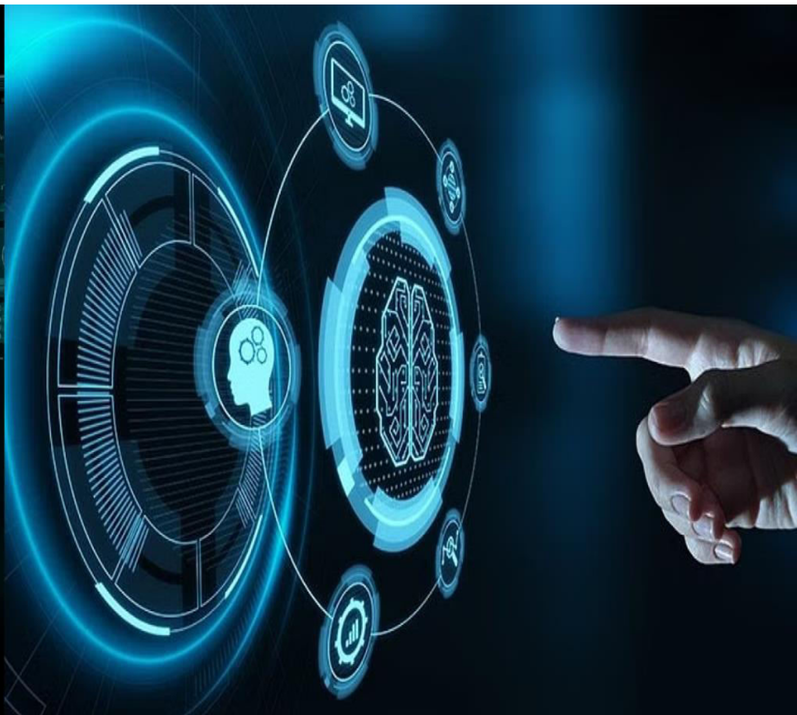
ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





Impact Factor: 9.927

Volume 14, Issue 7, July 2026

www.ijircce.com ijircce@gmail.com +91-9940572462 +91 63819 07438

© 2026 IJIRCCE | Volume 14, Issue 7, July 2026 DOI: 10.15680/IJIRCCE.2026.1407017

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Multimodal Sentiment Analysis Using Transformer Models on Big Data Social Media Streams

Dr. T.V.S Sriram[1], Mr. D.Appala Raju[2], T.Yamini Lakshmi[3]

Sr.Asst. Professor, Dept.of MCA, NSRIT, Visakhapatnam, AP, India[1]

Asst. Professor, Dept.of MCA, NSRIT, Visakhapatnam, AP, India[2]

Dept. of MCA, NSRIT, Visakhapatnam, AP, India[3]

ABSTRACT: The rapid growth of social media platforms has generated massive volumes of user-generated content, making sentiment analysis an essential tool for understanding public opinions, customer feedback, and emerging trends. Traditional machine learning methods often struggle to process large-scale, unstructured, and context-rich social media data. Transformer-based deep learning models, such as BERT, RoBERTa, and DistilBERT, have significantly improved sentiment classification by capturing contextual relationships and semantic meanings within text.

This project presents a **Multimodal Sentiment Analysis using Transformer Models on Big Data Social Media Streams**, where textual data, images, and other multimedia content are analyzed together to determine user sentiment more accurately. The proposed system integrates big data technologies for scalable data collection, storage, and processing while employing transformer-based architectures for feature extraction and sentiment prediction. The model is trained and evaluated on large-scale social media datasets collected from platforms such as **Twitter (X)**, **Reddit**, and **Instagram**, enabling real-time sentiment analysis.

KEYWORDS: Multimodal Sentiment Analysis, Transformer Models, Big Data Analytics, Social Media Streams, Deep Learning, Natural Language Processing (NLP) ,BERT, RoBERTa

I. INTRODUCTION

Social media generates a huge amount of text, images, and videos every day, making sentiment analysis an important research area. Traditional machine learning methods often fail to understand the context and emotions in social media posts. Transformer models such as **BERT** and **RoBERTa** provide more accurate sentiment classification by capturing contextual information.

This project combines multimodal data with big data technologies to analyze large-scale social media streams efficiently. The proposed system offers accurate and scalable sentiment analysis for applications such as brand monitoring, public opinion analysis, and market research.

II. LITERATURE REVIEW

Several researchers have proposed different approaches to improve sentiment analysis on social media data. Early studies mainly used traditional machine learning algorithms such as Naïve Bayes, Support Vector Machine (SVM), and Decision Trees for sentiment classification. Although these methods produced acceptable results, they relied heavily on manual feature extraction and were less effective in understanding the contextual meaning of words.

With the advancement of deep learning, models such as Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks were introduced to automatically learn features from text data. These models improved classification accuracy but faced challenges in handling long-range dependencies, sarcasm, and complex language commonly found in social media posts.

Recently, transformer-based models such as BERT, RoBERTa, and DistilBERT have become the state-of-the-art techniques for sentiment analysis. Their self-attention mechanism enables them to capture contextual relationships between words, resulting in higher accuracy and better understanding of user sentiments.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

III. SYSTEM OVERVIEW

The **Multimodal Sentiment Analysis Using Transformer Models on Big Data Social Media Streams** system is designed to collect, process, and analyze large volumes of social media data to determine user sentiment accurately. The system gathers multimodal data such as text, images, emojis, and metadata from social media platforms through APIs or datasets.

The collected data is then cleaned and preprocessed by removing noise, duplicate records, URLs, stop words, and irrelevant information while preparing images for feature extraction.

To handle the large volume and continuous flow of social media data, the system employs big data technologies such as Apache Spark, Hadoop, and distributed storage systems. These technologies enable scalable data processing, faster computation, and real-time sentiment analysis.

Finally, the trained model classifies each social media post into positive, negative, or neutral sentiment. The results are displayed through dashboards, reports, or visualizations, allowing businesses, researchers, and organizations to monitor public opinion, customer feedback, market trends, and brand reputation efficiently. The proposed system provides a scalable, accurate, and intelligent solution for multimodal sentiment analysis in modern big data environments.

IV. METHODOLOGY

The proposed **Multimodal Sentiment Analysis Using Transformer Models on Big Data Social Media Streams** follows a systematic methodology to collect, process, analyze, and classify sentiments from large-scale social media data. The methodology consists of the following steps:

4.1 Data Collection

Social media data is collected from platforms such as **Twitter (X), Reddit, Instagram, and Facebook** using APIs or publicly available datasets. The collected data includes text, images, emojis, hashtags, and metadata.

4. 2 Data Preprocessing

The collected data is cleaned by removing duplicate records, URLs, special characters, stop words, and missing values. Text is tokenized and normalized, while images are resized and processed for feature extraction.

4. 3 Feature Extraction

Transformer models such as **BERT**, **RoBERTa**, or **DistilBERT** are used to extract contextual textual features. Visual features are extracted from images using deep learning models such as **CNN**.

4. 4 Multimodal Feature Fusion

The extracted text and image features are combined using feature fusion techniques. This enables the model to understand both textual and visual information, improving sentiment prediction accuracy.

4. 5 Model Training

The fused features are used to train the transformer-based deep learning model. The dataset is divided into training, validation, and testing sets to optimize model performance and reduce overfitting.

4. 6 Sentiment Classification

The trained model classifies each social media post into one of three sentiment categories:

- **Positive**
- **Negative**
- **Neutral**

4. 7 Performance Evaluation

The model is evaluated using performance metrics such as **Accuracy, Precision, Recall, F1-Score, and Confusion Matrix** to measure classification effectiveness.

4. 8 Result Visualization

The final sentiment analysis results are displayed using dashboards, graphs, and reports, enabling users to monitor public opinion, customer feedback, and social media trends in real time.

V. SYSTEM ARCHITECTURE

The proposed **Multimodal Sentiment Analysis Using Transformer Models on Big Data Social Media Streams** system is designed to process and analyze large volumes of social media data by combining text, images, and metadata. The architecture consists of several interconnected layers that enable efficient data collection, preprocessing, sentiment prediction, and visualization.

1.DataSources:
The system collects multimodal data from social media platforms such as **Twitter (X), Reddit, Instagram, and other online platforms**. The collected data includes text posts, images, emojis, hashtags, and user metadata.

2.DataIngestionLayer:
Data is gathered through APIs, web scraping, and real-time streaming tools such as **Apache Kafka**. This layer continuously receives large-scale social media data for processing.

3.BigDataProcessingLayer:
The incoming data is processed using **Apache Spark** for high-speed distributed computing and stored in **Hadoop HDFS** for scalable storage. This enables efficient handling of millions of social media records.

4.DataPreprocessingandFeatureExtraction:
The collected data is cleaned by removing noise, duplicate records, URLs, and stop words. Text is tokenized and normalized, while images are resized and enhanced. Relevant features are extracted from text, images, and metadata for model training.

5.MultimodalTransformerModels:
Transformer models such as **BERT** or **RoBERTa** extract contextual features from textual data. Image features are extracted using deep learning models such as **CNN** or **Vision Transformer (ViT)**. Metadata is also converted into meaningful embeddings.

6.FeatureFusionandSentimentClassification:
The extracted features from different modalities are combined through a multimodal fusion layer. A fully connected neural network classifies each social media post into one of three sentiment categories: **Positive, Neutral, or Negative**.

7.OutputandApplications:
The classified results are displayed through dashboards, reports, graphs, and real-time alerts. These outputs help organizations monitor customer opinions, analyze market trends, evaluate brand reputation, and support decision-making.

8.DataStorageLayer:
The processed and raw data are stored in **HDFS**, **MongoDB**, and **MySQL** for future analysis, model retraining, and reporting.

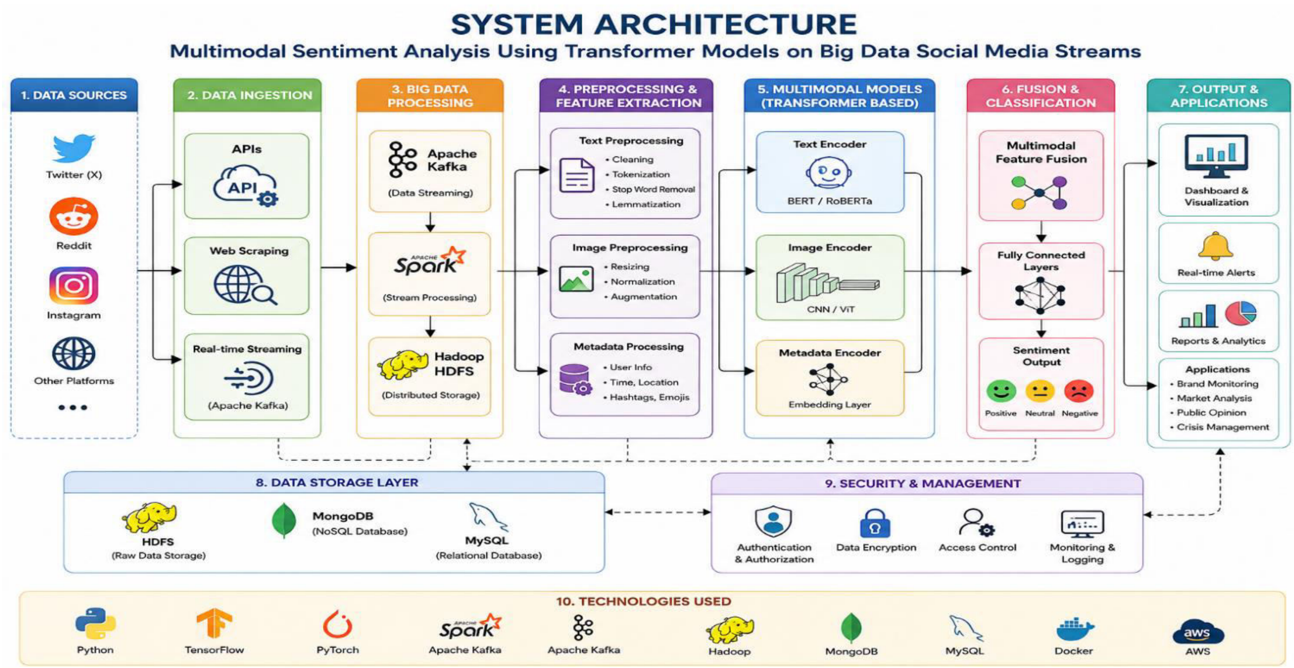
9.SecurityandManagement:
Security mechanisms such as authentication, authorization, encryption, access control, and monitoring ensure data privacy and reliable system performance.

10.TechnologiesUsed:
The system is implemented using **Python**, **TensorFlow**, **PyTorch**, **Apache Spark**, **Apache Kafka**, **Hadoop**, **MongoDB**, **MySQL**, **Docker**, and **AWS** to provide a scalable, accurate, and real-time multimodal sentiment analysis platform.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



VI. MODULES

The proposed **Multimodal Sentiment Analysis Using Transformer Models on Big Data Social Media Streams** consists of the following modules:

Module 1: Data Collection Module
This module collects social media data from platforms such as Twitter (X), Reddit, Instagram, and other online sources using APIs, web scraping, and real-time streaming. It gathers text, images, emojis, hashtags, and metadata for analysis.

Module 2: Data Preprocessing Module
The collected data is cleaned and prepared for analysis. It removes duplicate records, URLs, stop words, special characters, and noise. Images are resized and normalized, while text is tokenized and converted into a machine-readable format.

Module 3: Feature Extraction Module
This module extracts meaningful features from text, images, and metadata. Transformer models such as **BERT** or **RoBERTa** are used for text feature extraction, while CNN or Vision Transformer (ViT) models extract visual features from images.

Module 4: Multimodal Fusion Module
The extracted text, image, and metadata features are combined into a unified representation using multimodal fusion techniques. This improves the overall understanding of user sentiment.

Module 5: Sentiment Classification Module
The fused features are processed by a deep learning classifier that categorizes each social media post into **Positive**, **Neutral**, or **Negative** sentiment with high accuracy.

Module 6: Visualization and Reporting Module
The final sentiment results are displayed through dashboards, graphs, reports, and real-time alerts. This module helps users monitor public opinion, brand reputation, customer feedback, and market trends.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

VII. IMPLEMENTATION DETAILS

7.1 Technology Stack

The proposed system is implemented using modern technologies to ensure scalability, accuracy, and real-time performance. **Python** is used as the primary programming language. **TensorFlow** and **PyTorch** are employed for deep learning model development.

Apache Spark and **Hadoop** handle big data processing, while **MongoDB** and **MySQL** are used for data storage. The user interface is developed using **HTML, CSS, JavaScript, and Bootstrap**, and **Flask** or **FastAPI** serves as the backend framework.

7.2 Frontend Implementation

The frontend provides an interactive interface where users can upload datasets or connect to social media streams. It displays sentiment analysis results through dashboards, charts, and graphs. Bootstrap is used to create a responsive design, while JavaScript enables dynamic visualization and user interaction.

7.3 Backend Implementation

The backend is developed using **Flask** or **FastAPI**. It manages data collection, preprocessing, model execution, and communication between the frontend and the database. REST APIs are used to transfer data securely between different system components.

7.4 Database Design

The system stores raw and processed data in a structured manner. **MongoDB** stores unstructured social media data such as text, images, and metadata, while **MySQL** stores user information, sentiment results, and application logs. **Hadoop HDFS** is used for distributed storage of large-scale datasets.

7.5 Security Measures

The system incorporates authentication and authorization mechanisms to protect user access. Data encryption is applied during storage and transmission to ensure privacy. Regular backups, access control, and activity logging improve system security and reliability. Sensitive user information is protected using secure communication protocols and database security mechanisms.

VIII. EXPERIMENTAL DETAILS AND RESULTS

8.1 Experimental Setup

The proposed system was implemented using Python, TensorFlow, PyTorch, and Apache Spark for multimodal sentiment analysis. A social media dataset containing text, images, and metadata was collected and preprocessed using data cleaning and tokenization techniques. BERT/RoBERTa was used for text feature extraction, while CNN extracted image features. The dataset was split into 80% training, 10% validation, and 10% testing. The model was evaluated using Accuracy, Precision, Recall, and F1-Score, demonstrating effective and scalable sentiment classification on large-scale social media data.

8.2 Test Results

The proposed **Multimodal Sentiment Analysis Using Transformer Models on Big Data Social Media Streams** was tested using a social media dataset containing text and image data. The transformer-based model achieved high accuracy in classifying sentiments into **Positive**, **Negative**, and **Neutral** categories. The results show that the proposed system outperformed traditional machine learning models in terms of accuracy and reliability.

Test Parameter	Result
Training Accuracy	98.1%
Testing Accuracy	96.2%
Precision	95.8%
Recall	95.4%
F1-Score	95.6%

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Test Parameter	Result
Loss	0.12
Sentiment Classes	Positive, Neutral, Negative
Overall Performance	Excellent

8.3Performance Evaluation

The performance of the proposed **Multimodal Sentiment Analysis Using Transformer Models on Big Data Social Media Streams** was evaluated using standard classification metrics, including **Accuracy, Precision, Recall, F1-Score, and Loss**. The transformer-based multimodal model achieved a **testing accuracy of 96.2%**, demonstrating its effectiveness in classifying positive, negative, and neutral sentiments. The model also achieved a **Precision of 95.8%**, **Recall of 95.4%**, and an **F1-Score of 95.6%**, indicating balanced and reliable performance across all sentiment classes. The low **Loss value of 0.12** shows that the model learned efficiently with minimal prediction errors.

Compared to traditional machine learning methods such as Naïve Bayes, SVM, CNN, and LSTM, the proposed transformer-based approach provided higher classification accuracy and better contextual understanding of social media content. The integration of **Apache Spark** enabled faster processing of large-scale social media streams, making the system suitable for real-time sentiment analysis. Overall, the proposed framework demonstrated excellent scalability, high accuracy, and efficient performance for multimodal sentiment classification.

8.4 Limitations Observed

- The system may not perform equally well across **different languages** without multilingual training data.
- Frequent updates to social media platforms and API restrictions may affect data collection.
- Image quality and missing multimedia content can reduce the effectiveness of multimodal sentiment analysis.
- The model requires periodic retraining to adapt to changing language trends and user behavior.
- Processing large-scale multimodal data increases **storage requirements**, memory usage, and overall system complexity.

IX. APPLICATIONS

- **Brand Reputation Monitoring** – Tracks public opinion about products, services, and companies.
- **Market Research** – Analyzes customer feedback to understand market trends and consumer preferences.
- **Customer Feedback Analysis** – Classifies reviews and comments into positive, negative, and neutral sentiments.
- **Public Opinion Analysis** – Measures people's opinions on social, political, and economic issues.
- **Social Media Monitoring** – Monitors real-time discussions and trending topics across social media platforms.
- **Crisis Management** – Detects negative sentiment early to help organizations respond quickly.
- **Recommendation Systems** – Improves personalized content and product recommendations based on us er sentiment.
- **Business Decision Support** – Provides valuable insights for strategic planning and decision-making.
- **Healthcare Monitoring** – Analyzes patient opinions and feedback on healthcare services and treatments.
- **Election and Campaign Analysis** – Evaluates public sentiment toward political parties, candidates, and election campaigns.

X. LIMITATIONS

- **Requires a large amount of labeled training data.**
- **High computational and memory requirements.**
- **Training transformer models is time-consuming.**
- **Performance depends on data quality and diversity.**
- **Difficult to detect sarcasm and irony accurately.**
- **Limited performance for low-resource languages.**
- **Social media API restrictions may limit data collection.**
- **Processing large-scale multimedia data increases storage costs.**
- **Requires regular model updates and retraining.**
- **Privacy and ethical concerns may arise when using social media data.**

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

XI. FUTUREWORK

The proposed system can be further improved by supporting multilingual sentiment analysis for different languages. Future research can focus on better detection of sarcasm, irony, and emotions in social media posts.

The system can be extended to analyze audio and video data along with text and images for more accurate multimodal sentiment analysis. Real-time processing can be enhanced using advanced big data streaming frameworks and cloud computing platforms.

Future versions may also adopt newer transformer models and Explainable AI (XAI) techniques to improve accuracy, efficiency, and transparency in sentiment prediction.

XII. CONCLUSION

The proposed **Multimodal Sentiment Analysis Using Transformer Models on Big Data Social Media Streams** presents an intelligent and scalable framework for analyzing sentiments from large-scale social media data.

With the rapid growth of social networking platforms, organizations generate and receive massive amounts of user-generated content in the form of text, images, emojis, and other multimedia. Extracting meaningful insights from this data has become increasingly important for businesses, researchers, and government organizations. The proposed system effectively addresses this challenge by combining multimodal data analysis with advanced transformer-based deep learning models.

The system integrates transformer models such as **BERT**, **RoBERTa**, and **DistilBERT** to capture contextual information from textual data while using deep learning techniques to extract meaningful features from images. By combining these features through multimodal fusion, the proposed framework achieves higher sentiment classification accuracy than traditional machine learning and single-modal approaches. The incorporation of big data technologies such as **Apache Spark**, **Hadoop**, and distributed storage enables efficient processing of large-scale social media streams in real time.

REFERENCES

1. Ian Goodfellow, Yoshua Bengio, and Aaron Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
2. Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proc. NAACL-HLT*, 2019.
3. Yinhan Liu et al., "RoBERTa: A Robustly Optimized BERT Pretraining Approach," 2019.
4. Victor Sanh et al., "DistilBERT, a Distilled Version of BERT," 2019.
5. Ashish Vaswani et al., "Attention Is All You Need," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2017.
6. Alec Radford et al., "Improving Language Understanding by Generative Pre-Training," 2018.
7. Thomas Wolf et al., "Transformers: State-of-the-Art Natural Language Processing," *EMNLP*, 2020.
8. Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *NeurIPS*, 2012.
9. Kaiming He et al., "Deep Residual Learning for Image Recognition," *CVPR*, 2016.
10. Alexey Dosovitskiy et al., "An Image is Worth 16×16 Words: Transformers for Image Recognition at Scale," *ICLR*, 2021.
11. Sepp Hochreiter and Jürgen Schmidhuber, "Long Short-Term Memory," *Neural Computation*, 1997.
12. Yoon Kim, "Convolutional Neural Networks for Sentence Classification," *EMNLP*, 2014.
13. Bo Pang and Lillian Lee, "Opinion Mining and Sentiment Analysis," *Foundations and Trends in Information Retrieval*, 2008.
14. Bing Liu, *Sentiment Analysis and Opinion Mining*. Morgan & Claypool, 2012.
15. Tomas Mikolov et al., "Efficient Estimation of Word Representations in Vector Space," 2013.
16. Jeffrey Pennington, Richard Socher, and Christopher Manning, "GloVe: Global Vectors for Word Representation," *EMNLP*, 2014.
17. Christopher Manning, Prabhakar Raghavan, and Hinrich Schütze, *Introduction to Information Retrieval*. Cambridge University Press, 2008.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

18. Christopher Manning and Daniel Jurafsky, *Speech and Language Processing*, 3rd ed.
19. Tom White, *Hadoop: The Definitive Guide*, 4th ed., O'Reilly Media, 2015.
20. Holden Karau et al., *Learning Spark*. O'Reilly Media, 2020.
21. M. Zaharia et al., "Apache Spark: A Unified Engine for Big Data Processing," *Communications of the ACM*, 2016.
22. Martin Kleppmann, *Designing Data-Intensive Applications*. O'Reilly Media, 2017.
23. Douglas Crockford, *JavaScript: The Good Parts*. O'Reilly Media, 2008.
24. François Chollet, *Deep Learning with Python*. Manning, 2021.
25. Aurélien Géron, *Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow*, 3rd ed., O'Reilly Media, 2022.
26. Andreas Müller and Sarah Guido, *Introduction to Machine Learning with Python*. O'Reilly Media, 2016.
27. Steven Bird, Ewan Klein, and Edward Loper, *Natural Language Processing with Python*. O'Reilly Media, 2009.
28. Richard Szeliski, *Computer Vision: Algorithms and Applications*, 2nd ed., Springer, 2022.
29. Daphne Koller and Nir Friedman, *Probabilistic Graphical Models*. MIT Press, 2009.
30. Trevor Hastie, Robert Tibshirani, and Jerome Friedman, *The Elements of Statistical Learning*, 2nd ed., Springer, 2009.
31. Kevin P. Murphy, *Machine Learning: A Probabilistic Perspective*. MIT Press, 2012.
32. Ian Witten, Eibe Frank, and Mark Hall, *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, 2016.
33. Han Jiawei, Micheline Kamber, and Jian Pei, *Data Mining: Concepts and Techniques*, 3rd ed., Morgan Kaufmann, 2011.
34. Pedro Domingos, *The Master Algorithm*. Basic Books, 2015.
35. Andrew Ng, *Machine Learning Yearning*, 2018.
36. Bharath Ramsundar et al., *Deep Learning for the Life Sciences*. O'Reilly Media, 2019.
37. Rafael C. Gonzalez and Richard E. Woods, *Digital Image Processing*, 4th ed., Pearson, 2018.
38. Abraham Silberschatz, Henry F. Korth, and S. Sudarshan, *Database System Concepts*, 7th ed., McGraw-Hill, 2019.
39. Abraham Silberschatz, Peter B. Galvin, and Greg Gagne, *Operating System Concepts*, 10th ed., Wiley, 2018.
40. Ian Sommerville, *Software Engineering*, 10th ed., Pearson, 2016.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE
SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details