National Research Journal of Information Technology & Information Science
Volume No: 13, Issue No: 2, Year: 2026 (July-December)
PP: 103-122
ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

UNDERSTANDING HUMAN EMOTIONS THROUGH MACHINE LEARNING-BASED MULTIMODAL SENTIMENT ANALYSIS

Anjali

Research Scholar

CT University, Ludhiana (Punjab)

Ankita Makker

Department of Computer Applications, CT University, Ludhiana (Punjab)

ABSTRACT:

In today's digital world, people share their thoughts and feelings through social media, online platforms and digital content. They express emotions not only through written text but also through voice messages, images and videos. Because of this, understanding human emotions has become very important. Earlier, sentiment analysis was done only using text data. However, text alone cannot always show true emotions. Sometimes a person writes something positive but their voice or facial expression shows a different feeling. To solve this problem, Multimodal Sentiment Analysis was introduced. Multimodal Sentiment Analysis uses more than one type of data such as text, audio and video at the same time. This helps in understanding emotions more clearly and accurately. In this approach, models like SVM are used for text analysis, CNN is used for facial expression recognition and RNN is used for analyzing speech and sentence patterns. Recently, transformer models have also become popular because they understand emotions and context better. This technique is widely used in areas such as customer feedback analysis, online education, healthcare and social media monitoring. It helps organizations understand people's emotions and make better decisions. In conclusion, Multimodal Sentiment Analysis is a useful and powerful technology. In the future, it will play an important role in improving human–computer interaction and making digital systems more intelligent.

Keywords: multimodal sentiment analysis, emotion detection, text audio video analysis, machine learning, deep learning, human emotion recognition, sentiment classification, and human–computer interaction

1. INTRODUCTION

In today's digital world, communication has become very fast and easy. People share their thoughts, feelings and opinions every day through social media, online reviews, blogs, videos and voice messages. These platforms create a large amount of emotional information that shows how people feel about products, services and experiences. This information helps us understand human behavior such as happiness, anger, satisfaction or disappointment. Knowing these emotions is very important for businesses, schools, hospitals and online services because it helps them improve their work, understand users better and make better decisions.

© Copyright @ Authors

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in



Fig 1: An image showing people using mobile phones and social media apps with different facial expressions such as happiness, sadness, anger and surprise, representing various human emotions.

Human communication is not only about words. People express their feelings in many ways, such as through their facial expressions, tone of voice, hand movements and body language. Often, what a person says is not exactly what they feel inside. For example, a person may write a good or positive comment, but their face or voice may show sadness, stress or disappointment. Because of this difference between words and real emotions, understanding feelings only through written text is not always correct. Text alone cannot fully show how a person actually feels. As a result, traditional emotion analysis methods that depend only on written words sometimes fail to understand true emotions. This is why a deeper and more natural way of emotion understanding is needed.

.



Fig 2: A face showing mixed emotions, where the person is smiling but their eyes look sad, showing hidden feelings

To overcome this problem, researchers developed a better and more practical way to understand human emotions. This new method tries to understand feelings in the same way humans understand each other in real life. When people talk, their emotions are not shown only through words. Their tone, expressions and behavior also play an important role. That is why this approach does not depend on just one type of information. Instead, it looks at different signs together to understand what a person is really feeling.

By using more than one source of information, emotions can be identified more clearly and correctly. This helps in reducing mistakes that usually happen when emotions are judged

© Copyright @ Authors

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

using only one method. As a result, emotion detection becomes more accurate, reliable and closer to real human understanding. This method helps systems understand people in a more natural and realistic way.



Fig 3: An image showing a person using a computer or mobile, where facial expressions are being analyzed by a digital system.

Today, advanced emotion analysis is widely used in many fields. It helps companies understand customer feedback and know whether people are happy or unhappy with their products and services. In online education, it helps teachers understand students' interest, attention and learning difficulties. In the healthcare field, especially in mental health support, emotion analysis helps doctors and counselors understand a person's feelings and mental state more clearly. It is also used in marketing to study customer behavior and create better advertisements.

By understanding real human emotions, organizations can improve their services, solve problems more effectively and increase user satisfaction. This also helps in building better relationships between people and technology. As digital communication continues to grow every day, emotion-based analysis is becoming a very important part of modern technology and research. It helps make systems more human-friendly and intelligent.

Multimodal sentiment analysis is a method used to understand human emotions by using more than one type of information. Instead of depending only on written text, it also uses voice and video. In today's world, people do not express their feelings only through words. Their tone of voice, facial expressions and body language also show what they are actually feeling. Because of this, text alone is often not enough to understand emotions correctly.

That is why multimodal sentiment analysis is used. In this approach, different types of data are studied together to get better results. First, text analysis is done, where messages, comments and reviews are read to find out whether the emotion is positive, negative or neutral. However, sometimes the words look positive but the real feeling is different.The next step is audio analysis, where the person's voice is analyzed. The way someone speaks, their tone, speed and pitch can show emotions like happiness, anger, stress or sadness. For example, a loud voice may show anger or excitement, while a slow and soft voice may show sadness or tiredness.

After that, **video analysis** is used. In this, facial expressions, eye movements and body language are observed. A smiling face usually shows happiness, while a serious or confused face may show tension or disappointment. These visual signals help in understanding emotions more clearly.When text, audio and video are combined, the result becomes more accurate and reliable. This reduces misunderstanding and helps in finding the real emotion of a person. Today, multimodal sentiment analysis is widely used in areas like customer service,

© Copyright @ Authors

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

marketing, healthcare, education and human–computer interaction. It helps organizations understand people's true feelings and take better decisions based on them.

1.1Applications of Multimodal Sentiment Analysis

Today, multimodal sentiment analysis is widely used because it helps in understanding human emotions in a better way. It works by combining text, voice and video instead of using only one type of data. On social media platforms such as Facebook, Instagram and YouTube, people share their thoughts through posts, comments, images and videos. By analyzing all these together, it becomes easier to understand people's feelings, opinions and current trends. This helps companies learn how people feel about their products and services.

In customer support systems, this method plays an important role. Customers express their experiences through reviews, phone calls or video messages. By studying their words, voice tone and facial expressions, companies can easily find out whether customers are happy or unhappy. This helps them improve their services and respond to problems more quickly.

In online education, teachers often cannot see students directly. Multimodal sentiment analysis helps in understanding students' interest level, confusion or lack of attention by analyzing their voice and facial expressions. This makes online learning more interactive and effective.

This technology is also useful in the healthcare sector, especially for mental health. By observing speech patterns and facial expressions, doctors can identify stress, anxiety or depression at an early stage. This helps in providing timely treatment and care.In market research, companies use this approach to understand customer behavior and preferences. It helps them analyze reactions to products, advertisements and brands. Apart from this, multimodal sentiment analysis is also used in security systems to identify unusual or suspicious activities in public places.Overall, multimodal sentiment analysis is a powerful technique because it combines text, audio and video to understand emotions more accurately. Due to its effectiveness, it is now widely used in many real-life applications.

1.2 Advantage of Multimodal Sentiment analysis

Multimodal sentiment analysis has many advantages because it uses text, voice and video together to understand human emotions. This makes emotion detection more accurate than using only one type of data. Sometimes people write positive words but their voice or facial expressions show a different feeling. By analyzing all three together, emotions can be understood in a better way.

One major benefit of this method is that it works in real time. It is very useful in situations like customer support calls, live video chats and online interviews where quick decisions are needed. Companies can understand customer reactions instantly and respond in a better way.This technique also helps businesses in making better decisions. By studying customer feedback through text, voice and video, companies can improve their products and services. It also helps in increasing customer satisfaction because problems can be identified quickly and solved on time.

Multimodal sentiment analysis is also useful in online education. Teachers can understand whether students are interested, confused or bored by analyzing their voice and facial expressions. In healthcare, especially in mental health, it helps doctors understand stress, anxiety or depression by observing speech and facial behavior. Also, it supports automated systems like chatbots, virtual assistants and security systems. These systems work faster and

© Copyright @ Authors

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

more accurately with the help of emotion analysis. To sum up, multimodal sentiment analysis is a powerful and reliable technique that helps in understanding human emotions more clearly and is widely used in modern applications.

1.3 Challenges of Multimodal Sentiment Analysis

Multimodal sentiment analysis is a useful technique, but it also has some challenges. In this method, text, voice and video data are used together, which can create privacy issues because personal information is involved. If the data is not protected properly, it can be misused. Another problem is that this system needs a lot of computing power. Since different types of data are processed at the same time, fast computers, more memory and advanced hardware are required, which can be costly for many organizations.

Another challenge is the quality of data. In real life, data is often not clear. Audio may contain noise, videos may be blurred and text may have spelling mistakes. Such noisy data affects the accuracy of the results. Language difference is also a problem because people speak in different languages and express emotions in different ways. A single model may not work well for all languages and cultures. Another point is that, training multimodal models takes a lot of time because large datasets and multiple models are involved. This increases both time and resource usage. Due to these reasons, even though multimodal sentiment analysis is very effective, it needs better data, strong systems and proper handling to work efficiently in real-world applications

2. Literature Review

Existing research has widely explored the use of machine learning techniques in the field of sentiment analysis. Many studies focus on how different machine learning models can automatically classify opinions and emotions from textual data such as reviews, social media posts and user feedback. These studies explain that machine learning helps in identifying patterns in text and understanding whether the sentiment expressed is positive, negative or neutral.

The findings of previous research clearly show that sentiment analysis has become increasingly important in modern computing applications. It plays a significant role in supporting decision-making processes by helping organizations understand public opinion, customer satisfaction and user behavior. Because of its ability to process large amounts of data efficiently, sentiment analysis is now widely used in areas such as business intelligence, customer service systems, social media monitoring and online recommendation systems. This growing importance highlights the relevance of sentiment analysis as an essential tool in today's data-driven world.

Author/ Source /Ref No. /Year	Title	Techniques	Accuracy	Conclusion
Saravanan Arumugam/PETI/[1]/2025	Deep Learning-based smart invigilation System for Enhanced Exam Integrity	bLLMs, BERT RoBERT	98.8%	It minimizes the risk of human error in Higher education and effectively improves the detection of dishonest behaviours.
C.Joseph& P.UmaMaheswari /Scientific Reports/ [2]/2025	Facial emotion based Smartphone addiction detection and prevention using	MnasNet, CNN, TLBO	93%	It identifies the cause why students are becoming more addicted to Smartphone's using deep learning

© Copyright @ Authors

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

	Deep Learning and Video Based Learning			algorithms.
NitinShivsharan, Vikrant Kambliet al. / Procedia Computer Science/ [3]/2025	Evaluating Machine Learning Models for Sentiment Analysis of YouTube Comments and Creating an Accessible Web Application for Comment Analysis.	Extreme gradient boosting (XGB), Random Forest(RF), Logistic Regression (LR), AdaBoost (AdaB) Bagging Classifier (BgC)	95.8%	Development of user-friendly web application for classifying live YouTube comments accurately.
Harshal Pawar et al./IJCT/[4]/2025	Video Sentiment Analysis Using Deep Learning	Haar cascade algorithm, VADER,SentiWordNet,CNN,VAE	Accuracy upto 94%	It introduces a web based system that combines facial recognition, natural language processing and audio analysis.
Hayette Hadjar et al. / electronics/ [5]/2025	TheraSense: Deep Learning for Facial Emotion Analysis in Mental Health Teleconsultation	CNNs	96%	TheraSense system used CNN-based models to accurately detect emotions using audio and video sensesors.
Mahender kumar et al./PeerJ/[6]/2025	Evolving techniques in sentiment analysis: A Comprehensive Review	GPT,BERT, SVM	99.5%	Comparison of the results from four approaches that are lexicon-based, machine learning-based, deep learning-based, transformer based.
Mohammed Aly/ Multimedia Tools and Application/ [7]/2024	Revolutionizing online education: Advanced facial expression recognition for real-time student progress tracking via deep learning model.	ResNet50, CBAM,TCNs	91.08%	This paper focused on the detection of student's facial expression in online education.
Antonio Fernandez, Suzan Awinat/ ScienceDirect/[8]/2024	Multimodal Sentiment Analysis based on video and Audio Inputs	Wav2vec2, Vivit on CREMA-D and RAVDESS	Accuracy of audio model is 72.6% and accuracy of video model is 88%	Emotion recognition using audio and video inputs. It also tests transformer-based models for higher accuracy.
Prathwini , Prathyakshini / IJACSA/ [9]/2024	DeepEmo Vision: Unveiling Emotion Dynamics in Video Through Deep Learning Algorithms	CNN, SVM RNN	95%	With the combinations of machine learning approaches like KNN, RNN we can identify emotions in dynamic visual material in real time.
Md.Shofiqul Islam et al./Springer/ [10]/2024	Challenges and future in deep learning for sentiment analysis: a comprehensive review and a proposed novel	LSTM,CNN, RNN,GRU, BiLSTM	98.2%	A novel hybrid model CRDC(Capsule with deep CNN and Bi-structured RNN) Is proposed for better performance on multiple datasets.

© Copyright @ Authors

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

	hybrid approach.			
Hongyu Shi/Systems and Soft Computing /[11]/2024	A short video sentiment analysis model based on multimodal feature fusion	BERT, MTCNN	F1-score: 98 at training size 500. Accuracy up to 0.96	It shows strong recognition of emotions like anger, Happiness sadness and neutrality.
Md Shad Akhtar et al./ journal of autonomous intelligence/ [12]/2024	Multi-task Learning for Multi-modal Emotion Recognition and sentiment Analysis	Bag-of-Words, N-grams,SVM, Neural Network etc.	97%.	
GauravMeena et al. /IJIM/ [13]/2023	Sentiment analysis on images using convolution neural networks based Inception-v3 transfer learning approach	CNN based Inception-v3 architecture	99.5%	Focus on large portions of the body such as person's face.
Sharma Vishal Kumar Sureshbhai, TulsidasNakrani/[14]/2024	A Literature Review : Enhancing Sentiment Analysis of Deep Learning Techniques Using Generative AI Model	Systematic survey of existing research, analysis of deep learning algorithms, evaluation of datasets, assessment of performance metrics, and identification of research gaps.	90%	Improving detection expressions like implicit sentiment.
Ahmad ShobriHashim et al./Journal of System and Management Sciences. /[15]/2024	Leveraging Social Media Sentiment Analysis for Enhanced Disaster Management: A Systematic Review and Future Research Agenda.	BERT,Bi-LSTM, SVM,LEXICON+ LDA	98%	Use of Sentiment analysis to transform disaster management, especially when it comes to using social media data to gain real time insights.
Konstantinos I. Roumeliotis et al./Big data and Cognitive Computing/ [16]/2024	LLMs and NLP Models in Crypto currency Sentiment Analysis	GPT-4, BERT and FinBERT.	86.7%.	Support trading strategies, risk assessment and market forecasting. But remain challenge with volatile crypto, sarcasm and data quality.
Ringki Das et al./ACM Computing Surveys/[17]/2023	Multimodal Sentiment Analysis :A Survey of Methods, Trends, and Challenges	Naive Bayes , SVM ,CNNs ,LSTMs ,BERT , RoBERT	Accuracy= 82% F1-score=76%	Modern approaches integrate multimodal inputs to capture richer sentiment signals.
GauravMeena et al./ScienceDirect/[18]/2022	Sentiment Analysis From Images using VGG19 based Transfer Learning Approach	Deep CNN, VGG-19	Accuracy upto 99%	Transfer learning approach is applied to identify end to end image analysis.

© Copyright @ Authors

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

SerapCoban et al./IJEFI/ [19]/2024	Investigating the Impact of Social Media Driven Sentiment on the Electric Vehicle stock Market Using Quartile Regression	VADER, TextBlob, bibliometrics.	No numeric accuracy.	Positive sentiment toward ESS is increasing; ESS plays a growing role in energy transition, with clear research themes and gaps identified.
Steffen T. Eberhardt, et al./Psychotherapy Research/ [20]/2024	Decoding Emotions: Exploring the validity of sentimental Analysis in Psychotherapy.	NLP	90%	Decode emotions and support therapists. Overall, it is promising tool if combined with psychological insights and validated clinical data.
Wang Hanmei/Sage Journals/ [21]/2024	Deep Learning Based Analysis of Emotional and Content Relevance Between Bullet Screen and Subtitles As Movie Narrative Medium	BERT,LDA, TF-IDF Word2Vec	Like emotion is very high pre-set Subtitles (53.85%)	The Research identifies the difference in emotional expression between Subtitles and bullet screens as well as their correlations.
Xiang Liu et al./applied sciences/[22]/2023	MTR-SAM: Visual Multimodal Text Recognition and Sentiment Analysis in Public Opinion Analysis on the Internet	LK-PAN network, CML Distillation, GTC.	96%.	Sarcasm, Domain-shift, Multilingual issues, High resources, Imbalanced data.
Tejashwini S.G., Aradhana D / IJACSA/ [23]/2023	Multimodal Deep Learning Approach for Real –Time Sentiment Analysis in Video Streaming	CNN	84.6%	It identifies emotions from visual data using multi-view deep learning frame work.
TehseenMazhar, Muhammad Amir Malik et al. / Symmetry/ [24]/2022	Movie Reviews classification through Facial image Recognition and Emotion Detection Using Machine Learning Methods	Naive Bayes, SVM, Random Forest , CNN	84.72	Aspect oriented emotion classification for movie reviews.
Chamandeepkaur et al./ [25]/2022	Incorporating sentimental analysis into development of a hybrid classification model: A Comprehensive Study	Random Forest, Light Gradient Boosting Machine, logistic Regression, Support Vector Machine (SVM), and Naive Bayes classifers.	95%	new technique is more accurate and takes less time to execute than its predecessor. .

OVERVIEW OF LITERATURE REVIEW

The table presents an overview of previous research studies related to sentiment analysis. It highlights the methods used and the key findings of each study. This helps in understanding the research trends and identifying gaps for further study.

© Copyright @ Authors

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

This Research investigates how bigger Large Language Models (bLLMs) can improve Sentiment analysis for Software Engineering [1]. It aims to overcome challenges like data scarcity, imbalanced datasets, and cross platform generalization. comparision between zero-shot and few-shot prompting of bLLMs with fine-tuned smaller LLMs like BERT and RoBERTa. First step in understanding the possibility of using prompted bLLMs to identify sentiment in SE domain. our test demonstrates that bLLMs perform better than sLLMs when there is a lack of annotated data, and zero-shot learning is a practical strategy.

This study identifies the cause why teenagers are becoming more addicted to Smartphone's [2]. Conventional therapies are not enough to address these behavioral issues.

AI-based face expression recognition and video-based behavioral treatments are combined in TMVM (Theory of Mind Video Modeling).it also identifies emotions such as joy, sorrow or rage in student's films using the cameras on Smartphone's. BHP (Behavioral Parameters) assessed using SIDE theory like social identity, self-awareness, loss of individuality, responsibility and accountability. Emotion detection are conducted using MnaNET, TLBO (Teaching learning Based Optimization), CNN, CSO (Cuckoo Search Optimization).

In this paper a thorough examination of the live YouTube comments using machine learning models has beendone [3]. An efficient sentiment analysis algorithm and web application has also been developed using Random forest, extreme gradient boosting, and adaptive boosting techniques. This study also suggests focusing on imbalanced datasets in future.

This paper introduces a multimodal approach which combines facial recognition, natural language processing and audio analysis to detect sentiment more accurately [4]. Deep Learning techniques such as variation auto encoder, adversarial training and attention mechanisms proposed modal achieve higher accuracy. This research highlights some challenges also such as dealing with neutral sentiments, dataset imbalance and privacy concerns.

The main objective of this research is to analysis emotions using CNN [5]. TheraSense system integrates CNNs with user friendly interface to support more informed decision making in mental health care. Its architecture is based upon Node.js and cloud hosting.TheraSense shows that emotion aware teleconsultation can enrich remote therapy.

This Research uses elements of film text from the field of film studies to explore the emotional and content relevance between the two narrative mediums in film, subtitles and bullet screens, as well as the differences in narrative effects that they produce [6]. It also aims to develop the potential of data analysis using AI algorithms and tools. BERT, LDA, Word2Vec techniques are used in this study. It also identifies differences and correlations in emotional expression between subtitles and bullet screens.

This paper focus on facial expression recognition using proposed model during online classes [7]. ResNET50 model is used for the extraction of facial features. By combining ResNET, CBAM and TCNs achieved impressive accuracy across multiple datasets.

In this study researchers mainly focused on the development and the evaluation of multimodal to explore human emotions by using audio and video inputs [8]. Two large datasets are used in this research CREMA-D for voices and RAVDESS for facial expressions. The researchers also fine-tuned advanced transformer models: wav2vec2 for audio and vivit for video. Each model predicts emotions like anger, happiness or sadness and their outputs are merged through different decision making techniques.

© Copyright @ Authors

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

This paper aims to enhance multimodal emotion measurement in therapy by integrating automated text analysis [9]. it is considered as a valid and scalable tool for tracking emotional dynamics in therapy. it also supports data-informed psychotherapy, therapist training and real time feedback system. This study also suggests combing sentiment analysis with facial and vocal emotion recognition for better insights. it can meaningfully reflect emotional tone is psychotherapy sessions and may enhance emotion tracking when integrated with other modalities.

A Novel hybrid model is proposed in this study to achieve better performance [10]. Sentiment analysis is performed at three levels. Document level identifies general polarity. Sentence level performs emotion detection on individual sentence and Aspect level achieve fine grained sentiment tied to specific feature. Popular datasets like IMDB, YELP and Amazon are used in this research.Pre-processing steps include tokenization, stemming, normalization, noise removal and stop words filtering. It combines capsule networks, Deep CNN, and Bi-RNN in spite of multiple applications some of them are most popular like fake news detection, product reviews, emotion analysis.

This study identifies the different emotions like anger, happiness, sadness and neutrality [11]. MTCNN is used face detection. Future research can be done on more diverse emotional states. This research proposes a multimodal sentiment analysis modal for short videos.

This research mainly focuses upon classifying call center notes into positive, negative and neutral sentiments for better understanding of customer's feedback [12]. Up to 3000 notes has been collected from telecommunication companies in Jordan. Arabic BERT tokenize was used for classification BiLSTM (bidirectional long short terms memory) a type of RNN that is used to process data in forward and backward directions. Kaggle platform was used for training and testing data.

This study proposes a new method for software engineering task Automatic Semantic Augmentation Prompts (ASAP) [13]. Main objective of this research is to improve performance on code summarization and code completion tasks. Dataset used and tested is CodeSearchNet. Code-davinci-002, Text-davinci-003 and GPT-3.5 turbo models are used in this study. ASAP-enhanced prompts led to consistent and statistically significant improvements in BLEU scores across six programming languages

This paper explores Generative AI models like GPT, BERT, GANs to enhance sentiment analysis by generating and interpreting emotionally rich text [14]. these Generative models (GANs, GPT) are effective in handling imbalanced datasets and generating sentiment aware text. It also suggests future work on efficiency, bias reduction and interpretability of the models.

This study focus on Synthesize techniques, ethical considerations and integration of Sentiment Analysis in Disaster response and recovery [15]. High classification accuracy was demonstrated by models such as BERT, Bi-LSTM and SVM when it came to tweets about disaster. Sentiment Analysis aids in crisis communication, resource allocation and situational awareness. Disaster management is a global priority due to increasing frequency and severity of catastrophes. It involves strategic planning for response and recovery. The future agenda of this research is to create standardized crisis data sets and to extend this study beyond twitter to other platforms.

The ability of Large Language models (LLMs) and natural language processing (NLP) models to analyze sentiment from news stories about crypto currencies is examined in this Research [16]. It also compares and evaluate the performance of state of art models like GPT-

© Copyright @ Authors

4, BERT and FinBERT in the categorization of Sentiment. This study has implication for both academic research and real-world financial market applications. GPT-4, BERT, LLM, NLP and finBERT models are used for crypto currency sentiment analysis. Combinations of NLP and LLMs models represent best results in crypto currency market.

The main goal of this research is to improve the performance of multimodal sentiment analysis tasks to understand human emotions [17]. Machine learning and deep learning techniques like Naïve Bayes, SVM, CNN, BERT, RoBERT has expanded the scope and effectiveness of sentiment detection. This study focuses on images and videos to detect emotions. Applications are advertising, Social Media monitoring and multimedia content analysis.

The purpose of this study is to classify different emotions.VGG-19 modal performed better than other models like ResNet, DenseNet[18]. Transfer learning approach improves the performance. Continuous retraining on new data this modal can also be used in future.VGG-19 architecture consists of 16 convolutional layers and 3 fully connected layers. the convolutional layers extract deep visual features

Quantile regression analysis is used in this study to investigate the connection between sectoral stock returns and sentimental influenced by social media [19]. Quantile regression provides deeper insights into market volatility by capturing the distributional impacts of sentiment especially at the extremes, in contrast to classic models that concentrate on average market behavior. Using text mining techniques based on comments from the social media site Reddit, it also analyzes sentiment around the need for electric vehicles and looks at how it interacts with top indexes for the electric vehicle industry.NRC lexicons are used to classify emotion.Time –series quantile regression model is used to capture effects of different market conditions.

This study mainly evaluate the validity of measuring emotional expressions in psychotherapy sessions by sentiment analysis utilizing transformer based natural language processing model [20]. Emotional response from the patients and therapists were in good alignment with sentiment derived from NLP. This study also suggests future research that can combine sentiment analysis with automatic emotion reorganization in facial expression and audio streaming using NOVA (Nonverbal behavior Analyzer. XLM-T a transformer based multilingual sentiment analysis model. Sentiment scores aligned well with patient'sself reports of positive and negative emotions. Therapist rating of negative emotions showed weaker correlations. It also suggests the better model that can capture patient expressed affect than therapists-perceived affect.

This paper thoroughly explores the developing methods in sentiment analysis, identifying fundamental procedures such as feature extraction and data collection [21]. State-of-the-art transformer models are used in this research like Generative Pre-Trained Transformer (GPT), Bidirectional Encoder Representations from Transformers (BERT) are used to achieve accuracy. This study seeks to investigate current patterns and problems using sentiment analysis. It incorporates current research to provide a thorough summary of the discipline and to pinpoint areas that need to investigate. It also provides deep explanation of the various approaches of the Sentiment analysis including lexicon-based, machine learning, deep learning, transformer-based and hybrid methods.

This survey paper offers a thorough introduction to sentiment analysis, covering its uses, methods for classifying sentiment and widely used for assessment measures [22]. The main focus of this study is on the difficulties that come with various types of sentiment classification data such as small-scale data, cross-domain data, and multimodal data and cross

© Copyright @ Authors

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

lingual data. Techniques such as SVM, GLM, Bayes, Ensemble methods, CNN, RNN, LSTM, BiL STM.

This study develops a multi-view deep learning framework for the recognition of different types of emotions. CNNs are used for the extraction of features from visual inputs [23]. ADAM optimizer is used to build sequential model with four CNN layers and two fully connected layers. It incorporates multimodal and cross lingual analysis to interpret sentiments across diverse media formats and languages.

In this study authors reveal that video Analysis behind the feelings of the viewers of the movie through facial recognition [24]. For deeper emotion classification on different movie review a lightweight machine learning model is proposed in this study. In this study. techniques particularly Naive Bayes, Random Forest and Support Vector Machines are widely used for Sentimental Analysis tasks. There are various challenges identified like detection of datasets of different kind of emotions with same or different characteristics.

A novel hybrid framework that make use of three classifiers (SVM, Logistic Regression and Random forest) to increase the classification accuracy. It compares the model's performance against several matrices, such as accuracy, precision and recalls showing how effective the suggested method is in sentiment based categorization tasks, especially when it comes to analyzing customer feedback on websites. It enhances user sentiment classification from textual data with the goals of increasing recall, accuracy, and precision.

In recent years, many researchers have worked on multimodal sentiment analysis to understand human emotions in a better way. In these studies, text, voice and video data are used together instead of using only one type of data. This helps in getting more accurate results while studying human emotions.

Mohammad Ali (2024) used ResNet50 and Temporal Convolutional Network in his research. His work mainly focused on recognizing facial expressions from videos. ResNet50 was used to find important facial features, while TCN helped in understanding how emotions change with time. The results showed that when both models are used together, emotion detection becomes faster and more accurate. This method is also useful in real-time applications and human–computer interaction.

Antonio Fernandez (2024) worked on both audio and video data in his study. He used Wav2Vec2 to understand voice and ViViT to analyze videos. His work showed that emotions are easier to understand from videos than from voice alone because facial expressions clearly show how a person feels. The study also found that using more than one type of data gives better and more accurate results.

In 2024, Prithvimi carried out a study to understand human emotions using different models. He used CNN to analyze facial expressions, RNN to understand text and spoken words and SVM to classify emotions into different groups. The results of his study were quite good and showed that using more than one model helps in detecting emotions more accurately. Along with this, other studies also support the idea that multimodal sentiment analysis works better than using only one type of data. Models like CNN and ViViT are very useful for analyzing videos, while RNN and Wav2Vec help in understanding text and audio. When all these models are used together, emotions can be identified more clearly and correctly. Because of this, multimodal approaches are considered the best option for real-world emotion detection.

© Copyright @ Authors

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

3. METHODOLOGY

This study follows a well-planned methodology to analyze human emotions using multimodal sentiment analysis. Instead of depending on only one form of data, this research combines text, audio and video to achieve better understanding of emotions. In real-life situations, people express their feelings not only through written words but also through their voice and facial expressions. Therefore, using a single source of data does not provide complete emotional information. Keeping this in mind, multiple data sources were used in this study to improve the accuracy and reliability of emotion detection.



Figure: Proposed Methodology for Multimodal Sentiment Analysis

This study focuses on multimodal sentiment analysis to understand human emotions using text, audio and visual data. First, relevant literature is reviewed to identify existing methods and approaches. Then, different methodologies such as text-based, audio-based, visual-based and multimodal techniques are classified and compared. Finally, the study identifies research gaps, key findings and concludes that multimodal approaches provide more accurate and reliable results than single-mode analysis.

The first step of the study was data collection. Text data was taken from social media posts and online reviews, while audio data was collected from recorded speech samples. Video data included clips showing different facial expressions such as happiness, anger and sadness. After collecting the data, preprocessing was done to improve its quality. In text data, unnecessary words and mistakes were removed. Audio recordings were cleaned to reduce background noise and video data was processed by selecting useful frames and detecting facial areas. This step helped in making the data clear and ready for analysis.

After preprocessing, feature extraction was carried out. From text data, emotional words and sentence patterns were identified. From audio data, features like pitch, tone and voice intensity were studied to understand emotional changes in speech. From video data, facial

© Copyright @ Authors

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

expressions and movements were analyzed to recognize different emotions. These features helped in improving the overall accuracy of emotion detection.

Once the features were extracted, different machine learning and deep learning models were applied. Convolutional Neural Networks (CNN) were used for facial expression analysis because they work well with images and videos. Recurrent Neural Networks (RNN) were used for text and speech data since they are good at handling sequence-based information. Support Vector Machine (SVM) was used for emotion classification as it gives reliable results even when the dataset is small.

Along with these models, transformer-based methods were also used to improve performance. Wav2Vec was used to analyze speech and understand emotional patterns in voice. ViViT was used for video analysis to study facial expressions more effectively. These advanced models helped in improving accuracy and made the system faster and more efficient.

In the final stage, results from all models were combined to classify emotions into positive, negative and neutral categories. The findings showed that using text, audio and video together gives better and more accurate results than using only one type of data. This study proves that multimodal sentiment analysis is a useful and effective method for understanding human emotions and can be applied in areas such as customer feedback analysis, interview evaluation and human–computer interaction.

5. COMPARATIVE ANALYSIS OF SENTIMENT ANALYSIS METHOD

Sentiment analysis helps in understanding what people feel by studying their words, voice and expressions. Different techniques are used for this purpose and each technique works in a different way. Some methods analyze only written text, while others depend on speech or facial movements. Because of these differences, the results produced by each method are not always the same. That is why it becomes necessary to study and compare these methods. Through comparison, it becomes easier to understand which method gives better and more reliable results.

Analysis Dimension	Text-Based Method	Audio-Based Method	Visual-Based Method	Multimodal Method
Data Source	Written words, reviews, posts	Voice signals, tone, pitch	Facial expressions, gestures	Combination of text, voice & visuals
Emotion Detection Ability	Limited to word meaning	Medium – captures emotional tone	High – shows real expressions	Very High – complete emotional understanding
Emotional Depth	Shallow (only what is written)	Moderate (voice reflects feelings)	Deep (face shows emotions)	Deepest (all cues combined)
Accuracy Level	Moderate	Better than text	High	Highest
Context Understanding	Weak (sarcasm ignored)	Partial	Visual context only	Strong contextual understanding
Noise Sensitivity	Low	High (background noise affects)	Medium (lighting, angle)	Balanced (one method compensates another)
Real-world	Limited	Medium	Good	Very high

© Copyright @ Authors

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

Analysis Dimension	Text-Based Method	Audio-Based Method	Visual-Based Method	Multimodal Method
Reliability				
Technical Complexity	Low	Medium	High	Very High
Computational Cost	Low	Medium	High	Highest
Practical Usage	Reviews, feedback analysis	Call centers, voice assistants	Surveillance, emotion tracking	AI systems, smart assistants, research
Major Limitation	Cannot detect real emotion	Affected by noise	Expression may be fake	Requires large data &resources
Overall Effectiveness	Basic level	Improved	Advanced	Most effective & reliable

5.1 Analytical Interpretation

This comparison clearly shows that no single method is sufficient to understand human emotions completely. Text-based analysis only captures surface-level emotions, while audio-based analysis adds emotional tone but lacks semantic clarity. Visual-based methods provide rich emotional information but are highly dependent on external conditions such as lighting and facial visibility. In contrast, the multimodal approach integrates all three sources, making emotion detection more accurate, balanced and reliable. Although it requires higher computational power and complex modeling, its performance surpasses all single-modality methods. Therefore, multimodal sentiment analysis represents the most advanced and realistic approach for understanding human emotions in modern intelligent systems.

6. RESULTS AND DISCUSSION

The results of this study show that emotion recognition works much better when text, audio and video data are used together. During testing, the system was able to identify emotions such as positive, negative and neutral more accurately. It was seen that when only one type of data, like text or audio, was used, the system could not understand emotions properly. But when all three types of data were combined, the results improved clearly. This shows that using a multimodal approach is important for better emotion analysis.

The study also showed that each model played an important role in improving the results. The Convolutional Neural Network (CNN) worked well in detecting facial expressions from video data. It was able to identify expressions like happiness, anger and other emotions correctly. The Recurrent Neural Network (RNN) was helpful in analyzing text and speech because it can understand the order of words and speech flow. This helped in recognizing emotional tone more accurately. The Support Vector Machine (SVM) also gave stable results, especially when the amount of data was limited, making it useful for classification tasks.

To further improve performance, transformer-based models such as Wav2Vec and ViViT were used. Wav2Vec helped in understanding speech by analyzing voice tone, pitch and intensity while ViViT was effective in studying facial movements from video. These models helped increase the accuracy and reliability of the system. Another important point observed was the importance of data preprocessing. Cleaning text, removing noise from audio and selecting useful video frames improved the quality of data and helped the system give better results.

© Copyright @ Authors

The findings of this study clearly show that multimodal sentiment analysis works better than using only one type of data. By combining text, audio and video, the system can understand human emotions in a more natural and accurate way. This approach is very useful for real-life applications such as customer feedback analysis, interview evaluation, emotion-based recommendation systems and human–computer interaction. The study confirms that using multiple data sources together is an effective way to improve emotion recognition.

7. CONCLUSION

This study shows that multimodal sentiment analysis is an important and useful technology in today's digital world. It understands human emotions by using text, voice and video together. This makes emotion detection more accurate than using only text. When voice tone and facial expressions are also considered, emotions can be understood in a better and more natural way. Because of this, the results become more reliable. This technology is very useful in many areas such as business, education and healthcareand customer service. In business, it helps companies understand customer opinions. In education, it helps teachers understand students' feelings and interest. In healthcare, especially mental health, it helps in observing emotional conditions. These uses show that multimodal sentiment analysis is very helpful in real life. There are some challenges like data privacy, high cost and long training time. However, with better technology and improved methods, these problems can be reduced in the future. Overall, multimodal sentiment analysis is a strong and promising technique. In the coming years, it will become faster, more accurate and more widely used, helping machines understand human emotions better and improving interaction between humans and computers.

8. Future Scope of Multimodal Sentiment Analysis

The future of multimodal sentiment analysis is very promising because artificial intelligence and machine learning are developing very fast. In the coming years, this technology will be used in many important areas. Smart assistants like Alexa, Siri and chatbots will become more intelligent and will be able to understand not only words but also voice tone and facial expressions. This will make communication between humans and machines more natural. Besides this, emotionally intelligent machines will be developed that can understand human feelings and respond in a better way. These machines can be used in customer service, healthcare and even in supporting elderly people.

In the field of education, multimodal sentiment analysis will help teachers understand students' interest, confusion or stress during online classes, which will improve the learning experience. In healthcare, especially mental health, this technology will help in identifying stress, anxiety and emotional changes at an early stage. It will support doctors in providing better treatment. Moreover, in smart cities, this technology can be used to improve public safety, monitor behavior and detect unusual activities.

Finally, multimodal sentiment analysis has a bright future. As technology continues to grow, this system will become more accurate, intelligent and widely used. It will play an important role in improving human–computer interaction and making modern systems smarter and more helpful in daily life.

REFERENCES

1. Saravanan Arumugam, "Deep Learning-Based Smart Invigilation System for Enhanced Exam Integrity," *Proc. Eng. Technol. Innov.*, vol. 29, pp. 99–115, Feb. 2025, doi: 10.46604/peti.2024.14105.

© Copyright @ Authors

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

2. C. Joseph and P. U. Maheswari, “Facial emotion based smartphone addiction detection and prevention using deep learning and video based learning,” *Sci. Rep.*, vol. 15, no. 1, p. 18025, May 2025, doi: 10.1038/s41598-025-99681-7.
3. N. Shivsharan, V. Kambli, S. Dabholkar, A. Dalvi, and T. Sukali, “Evaluating Machine Learning Models for Sentiment Analysis of YouTube Comments and Creating an Accessible Web Application for Comment Analysis,” *Procedia Comput. Sci.*, vol. 258, pp. 3165–3174, 2025, doi: 10.1016/j.procs.2025.04.574.
4. S. Taware and H. Pawar, “Video Sentiment Analysis Using Deep Learning,” vol. 12, no. 2, 2025.
5. H. Hadjar, B. Vu, and M. Hemmje, “TheraSense: Deep Learning for Facial Emotion Analysis in Mental Health Teleconsultation,” *Electronics*, vol. 14, no. 3, p. 422, Jan. 2025, doi: 10.3390/electronics14030422.
6. M. Kumar, L. Khan, and H.-T. Chang, “Evolving techniques in sentiment analysis: a comprehensive review,” *PeerJ Comput. Sci.*, vol. 11, p. e2592, Jan. 2025, doi: 10.7717/peerj-cs.2592.
7. M. Aly, “Revolutionizing online education: Advanced facial expression recognition for real-time student progress tracking via deep learning model,” *Multimed. Tools Appl.*, vol. 84, no. 13, pp. 12575–12614, Jun. 2024, doi: 10.1007/s11042-024-19392-5.
8. A. Fernández and S. Awinat, “Multimodal Sentiment Analysis based on Video and Audio Inputs,” *Procedia Comput. Sci.*, vol. 251, pp. 41–48, 2024, doi: 10.1016/j.procs.2024.11.082.
9. P. - and P. -, “DeepEmoVision: Unveiling Emotion Dynamics in Video Through Deep Learning Algorithms,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 3, 2024, doi: 10.14569/IJACSA.2024.0150388.
10. Md. S. Islam *et al.*, “Challenges and future in deep learning for sentiment analysis: a comprehensive review and a proposed novel hybrid approach,” *Artif. Intell. Rev.*, vol. 57, no. 3, p. 62, Mar. 2024, doi: 10.1007/s10462-023-10651-9.
11. H. Shi, “A short video sentiment analysis model based on multimodal feature fusion,” *Syst. Soft Comput.*, vol. 6, p. 200148, Dec. 2024, doi: 10.1016/j.sasc.2024.200148.
12. M. S. Akhtar, D. Chauhan, D. Ghosal, S. Poria, A. Ekbal, and P. Bhattacharyya, “Multi-task Learning for Multi-modal Emotion Recognition and Sentiment Analysis,” in *Proceedings of the 2019 Conference of the North*, Minneapolis, Minnesota: Association for Computational Linguistics, 2019, pp. 370–379. doi: 10.18653/v1/N19-1034.
13. G. Meena, K. K. Mohbey, and S. Kumar, “Sentiment analysis on images using convolutional neural networks based Inception-V3 transfer learning approach,” *Int. J. Inf. Manag. Data Insights*, vol. 3, no. 1, p. 100174, Apr. 2023, doi: 10.1016/j.jjimei.2023.100174.
14. Sharma Vishalkumar Sureshbhai and Dr. Tulsidas Nakrani, “A Literature Review : Enhancing Sentiment Analysis of Deep Learning Techniques Using Generative AI Model,” *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.*, vol. 10, no. 3, pp. 530–540, Jun. 2024, doi: 10.32628/CSEIT24103204.

© Copyright @ Authors

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

15. A. S. Hashim *et al.*, "Leveraging Social Media Sentiment Analysis for Enhanced Disaster Management: A Systematic Review and Future Research Agenda," vol. 15, no. 4, 2025.
16. K. I. Roumeliotis, N. D. Tselikas, and D. K. Nasiopoulos, "LLMs and NLP Models in Cryptocurrency Sentiment Analysis: A Comparative Classification Study," *Big Data Cogn. Comput.*, vol. 8, no. 6, p. 63, Jun. 2024, doi: 10.3390/bdcc8060063.
17. R. Das and T. D. Singh, "Multimodal Sentiment Analysis: A Survey of Methods, Trends, and Challenges," *ACM Comput. Surv.*, vol. 55, no. 13s, pp. 1–38, Dec. 2023, doi: 10.1145/3586075.
18. G. Meena, K. K. Mohbey, A. Indian, and S. Kumar, "Sentiment Analysis from Images using VGG19 based Transfer Learning Approach," *Procedia Comput. Sci.*, vol. 204, pp. 411–418, 2022, doi: 10.1016/j.procs.2022.08.050.
19. S. Çoban, M. Uçar, N. Ulusay, and N. Kaya, "Investigating the Impact of Social Media-Driven Sentiment on the Electric Vehicle Stock Market Using Quantile Regression," *Int. J. Econ. Financ. Issues*, vol. 15, no. 4, pp. 299–316, Jun. 2025, doi: 10.32479/ijefi.19430.
20. S. T. Eberhardt *et al.*, "Decoding emotions: Exploring the validity of sentiment analysis in psychotherapy," *Psychother. Res.*, vol. 35, no. 2, pp. 174–189, Feb. 2025, doi: 10.1080/10503307.2024.2322522.
21. W. Hanmei, "Deep Learning-Based Analysis of Emotional and Content Relevance Between Bullet Screens and Subtitles as Movie Narrative Medium," *Sage Open*, vol. 14, no. 3, p. 21582440241280840, Jul. 2024, doi: 10.1177/21582440241280840.
22. X. Liu *et al.*, "MTR-SAM: Visual Multimodal Text Recognition and Sentiment Analysis in Public Opinion Analysis on the Internet," *Appl. Sci.*, vol. 13, no. 12, p. 7307, Jun. 2023, doi: 10.3390/app13127307.
23. T. S. G and A. D, "Multimodal Deep Learning Approach for Real-Time Sentiment Analysis in Video Streaming," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 8, 2023, doi: 10.14569/IJACSA.2023.0140881.
24. T. Mazhar *et al.*, "Movie Reviews Classification through Facial Image Recognition and Emotion Detection Using Machine Learning Methods," *Symmetry*, vol. 14, no. 12, p. 2607, Dec. 2022, doi: 10.3390/sym14122607.
25. Akhtar, M. S., Chauhan, D., Ghosal, D., Poria, S., Ekbal, A., & Bhattacharyya, P. (2019). Multi-task learning for multi-modal emotion recognition and sentiment analysis. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics*, 370–379. https://doi.org/10.18653/v1/N19-1034
26. Al Amrani, Y., Lazaar, M., & El Kadiri, K. E. (2018). Random forest and support vector machine based hybrid approach to sentiment analysis. *Procedia Computer Science, 127*, 511–520. https://doi.org/10.1016/j.procs.2018.01.150
27. Ahmed, T., Pai, K. S., Devanbu, P., & Barr, E. (2024). Automatic semantic augmentation of language model prompts (for code summarization). *Proceedings of the IEEE/ACM 46th International Conference on Software Engineering*, 1–13. https://doi.org/10.1145/3597503.3639183

© Copyright @ Authors

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

28. Alsokkar, A., Otair, M., Alfar, H. E., Nasereddin, A. Y., Aldiabat, K., &Abualigah, L. (2024). Sentiment analysis for Arabic call center notes using machine learning techniques. *Journal of Autonomous Intelligence, 7*(3). https://doi.org/10.32629/jai.v7i3.940

29. Bordoloi, M., & Biswas, S. K. (2023). Sentiment analysis: A survey on design framework, applications and future scopes. *Artificial Intelligence Review, 56*(11), 12505–12560. https://doi.org/10.1007/s10462-023-10442-2

30. Chen, X., Xie, H., Tao, X., Wang, F. L., Zhang, D., & Dai, H.-N. (2025). A computational analysis of aspect-based sentiment analysis research through bibliometric mapping and topic modeling. *Journal of Big Data, 12*(1), 40. https://doi.org/10.1186/s40537-025-01068-y

31. Chenchah, F., &Lachiri, Z. (2014). Speech emotion recognition in acted and spontaneous context. *Procedia Computer Science, 39*, 139–145. https://doi.org/10.1016/j.procs.2014.11.020

32. Das, D., Aayushman, Kumar, S., Hussain, M. A., & Reddy, B. R. (2025). Diabetes prediction using ensemble learning techniques. *Procedia Computer Science, 258*, 3155–3164. https://doi.org/10.1016/j.procs.2025.04.573

33. Das, R., & Singh, T. D. (2023). Multimodal sentiment analysis: A survey of methods, trends, and challenges. *ACM Computing Surveys, 55*(13s), 1–38. https://doi.org/10.1145/3586075

34. Du, K., Xing, F., Mao, R., & Cambria, E. (2024). Financial sentiment analysis: Techniques and applications. *ACM Computing Surveys, 56*(9), 1–42. https://doi.org/10.1145/3649451

35. Eberhardt, S. T., Schaffrath, J., Moggia, D., Schwartz, B., Jaehde, M., Rubel, J. A., Baur, T., André, E., & Lutz, W. (2025). Decoding emotions: Exploring the validity of sentiment analysis in psychotherapy. *Psychotherapy Research, 35*(2), 174–189. https://doi.org/10.1080/10503307.2024.2322522

36. Guo, W., & Ye, H. (2023). Emotional analysis and expression in advertising art design based on deep learning algorithms. *Computer-Aided Design and Applications*, 147–162. https://doi.org/10.14733/cadaps.2024.S15.147-162

37. Islam, M. S., Kabir, M. N., Ghani, N. A., Zamli, K. Z., Zulkifli, N. S. A., Rahman, M. M., & Moni, M. A. (2024). Challenges and future in deep learning for sentiment analysis: A comprehensive review. *Artificial Intelligence Review, 57*(3), 62. https://doi.org/10.1007/s10462-023-10651-9

38. Joseph, C., &Maheswari, P. U. (2025). Facial emotion-based smartphone addiction detection using deep learning. *Scientific Reports, 15*(1), 18025. https://doi.org/10.1038/s41598-025-99681-7

39. Kastrati, Z., Dalipi, F., Imran, A. S., PirevaNuci, K., &Wani, M. A. (2021). Sentiment analysis of students' feedback using NLP and deep learning. *Applied Sciences, 11*(9), 3986. https://doi.org/10.3390/app11093986

40. Kumar, M., Khan, L., & Chang, H.-T. (2025). Evolving techniques in sentiment analysis: A comprehensive review. *PeerJ Computer Science, 11*, e2592. https://doi.org/10.7717/peerj-cs.2592

© Copyright @ Authors

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

41. Liu, W., Wu, L., & Zhao, H. (2026). Improving sentiment analysis in tourism through LLM-enhanced irony detection. *Tourism Management, 112*, 105272. https://doi.org/10.1016/j.tourman.2025.105272

42. Mitra, A., &Mohanty, S. (2021). Sentiment analysis using machine learning approaches. In *Emerging Technologies in Data Mining and Information Security* (pp. 63–68). Springer. https://doi.org/10.1007/978-981-33-4367-2_7

43. Zadeh, A., Chen, M., Poria, S., Cambria, E., &Morency, L.-P. (2017). Tensor fusion network for multimodal sentiment analysis. *Proceedings of EMNLP*, 1103–1114. https://doi.org/10.18653/v1/D17-1115

44. Zhang, H., Wang, Y., Yin, G., Liu, K., Liu, Y., & Yu, T. (2023). Learning language-guided adaptive hyper-modality representation for multimodal sentiment analysis. *Proceedings of EMNLP*, 756–767. https://doi.org/10.18653/v1/2023.emnlp-main.49

45. Zhou, J., Liang, Z., Fang, Y., & Zhou, Z. (2024). Exploring public response to ChatGPT with sentiment analysis. *IEEE Access, 12*, 50504–50516. https://doi.org/10.1109/ACCESS.2024.3386362

© Copyright @ Authors