

Advances in Emotion Recognition through Twitter Text: A Review of Ensemble and Novel Techniques

Piyush R. Kulkarni¹ and Atul Agrawal²

^{1,2}Department of Computer Science and Engineering, Oriental University, Indore (M.P.), India
Email: piyushrkulkarni@gmail.com, atul.273@gmail.com

Abstract— The rise of social media platforms, especially Twitter, has positioned textual emotion recognition as a key component in understanding public opinion, behavioural dynamics, and emerging social trends. Given Twitter's rapid pace and informal nature characterized by abbreviations, slang, code-switching, and sarcasm, accurate emotion detection from tweets presents a complex challenge. Yet, the platform's immense volume of user-generated content provides an unparalleled opportunity for real-time emotional analysis at scale. This review critically surveys recent advancements in tweet-based emotion recognition, with a particular emphasis on novel machine learning techniques and ensemble-based frameworks. It explores how various computational models including deep learning and transformer architectures are applied to extract emotional signals from short texts. In addition, we compare widely used benchmark datasets, discuss text representation techniques, and evaluate the trade-offs between standalone and ensemble classifiers in capturing emotional depth and diversity. Beyond model performance, the paper also addresses the growing importance of context-aware learning, multimodal integration (e.g., emojis, hashtags, and user metadata), and transfer learning across linguistic and cultural boundaries. Persistent challenges such as limited labelled data, lack of model transparency, and the need for real-time deployment are highlighted. The paper concludes by outlining future research directions aimed at enhancing the adaptability, interpretability, and scalability of emotion recognition systems. This review serves as a foundational guide for researchers, developers, and practitioners in the domains of affective computing and social media analytics.

Index Terms— Emotion recognition, NLP, Machine learning in emotion recognition, deep learning models.

I. INTRODUCTION

Public discourse and opinion-sharing and experience-sharing have undergone fundamental changes because of social media. Twitter has become a major communication platform which supports millions of users during real-time interactions thus providing essential public sentiment data. Twitter provides micro-bloggers the ability to create short direct posts which generate a continuously active database through which users express their public reactions toward current events. The analysis of this emotion reveals patterns which provide extensive public perception information about diverse topics including political matters and entertainment alongside brand and worldwide events perception. Understanding collective behaviour and decision-making requires emotion analysis of Twitter data.

With the explosive growth of social media, platforms like Twitter have become critical venues for expressing emotions, opinions, and sentiments in real-time. The concise nature of tweets, often embedded with slang, abbreviations, emojis, and hashtags, presents a unique linguistic landscape for natural language processing (NLP). Emotion recognition from textual tweets has gained traction due to its applications in public health monitoring, customer feedback analysis, crisis detection, and political forecasting.

Traditional sentiment analysis techniques, primarily based on lexicons and machine learning models, have shown limitations in capturing the nuances of emotions. This has led to a surge in research on deep learning and ensemble methods, which combine multiple algorithms to improve classification accuracy and generalization. Ensemble approaches, in particular, have proven effective in mitigating the limitations of individual models by leveraging their strengths. This paper presents a comprehensive review of the literature on emotion recognition from tweets, focusing on ensemble and novel techniques. It analyses multiple recent studies that utilized hybrid models, multimodal data, and deep learning frameworks to enhance performance. Additionally, it provides a comparative analysis of datasets, model evaluations, and highlights the challenges and future prospects in this domain.

II. LITERATURE REVIEW

Emotion and sentiment analysis from social media data, particularly Twitter, has gained substantial attention in recent years, especially during crisis events like the COVID-19 pandemic. Numerous works have leveraged deep learning and ensemble methods to improve the accuracy and robustness of emotion detection systems.

In [1], Usha et al. proposed a tweet-based emotion recognition system that demonstrates the viability of NLP techniques in real-time applications.

Braig et al. [2] provided a comprehensive exploration of machine learning techniques, comparing various models for analysing COVID-19-related tweets. Their results indicate that ensemble methods consistently outperform single-model architectures in handling noisy social media data.

Similarly, Sunitha et al. [3] developed an ensemble deep learning model to analyze sentiments during COVID-19 in India and Europe, emphasizing region-specific emotional trends. These studies show the importance of adapting models to domain-specific and regional data.

An ensemble approach was also adopted by Dai et al. [4], who participated in the EVALITA 2023 shared task using categorical emotion classification. Their method confirmed that ensemble techniques improve generalization and classification accuracy.

Hussain et al. [5] extended this approach by presenting an ensemble deep learning framework specifically aimed at detecting situational aspects in students' tweets, demonstrating the relevance of context-aware emotion recognition.

Salama et al. [6] supported this view by integrating ensemble methods with active learning to handle heterogeneous data, thereby reducing labeling costs while maintaining accuracy.

Multimodal ensemble learning was further explored by Awan et al. [7], who used physiological signals in addition to textual data to chart emotions. Their work highlights the future potential of multimodal fusion in sentiment analysis systems.

In the domain of financial sentiment, Aslam et al. [8] developed an ensemble LSTM-GRU model for analyzing cryptocurrency-related tweets.

Krishna et al. [9] introduced a novel ML and LSTM-based ensemble model for real-time tweet classification, achieving high accuracy and low latency.

Surveys and review work such as those by Alsawaidan and Menai [10] offer a taxonomy of text-based emotion recognition techniques, identifying the evolution from traditional machine learning to deep learning and ensemble-based methods.

Qin et al. [11] proposed a dual-attention mechanism combined with TCN-BiGRU, which significantly enhances temporal feature learning.

Zygadło et al. [12] explored multilingual emotion detection for therapeutic chatbot applications, emphasizing the growing interest in emotionally intelligent dialogue systems.

Wang et al. [13] applied CNN for emotion classification on Chinese microblogs, demonstrating that convolutional approaches can still be competitive in certain language settings.

Similarly, Huang et al. [14] used LSTM for sentiment-based cryptocurrency prediction, showing its predictive power on financial datasets.

Tam et al. [15] also reported strong results using a ConvBiLSTM model, noting improvements in learning long-term dependencies.

Classical machine learning and hybrid models have also been considered. Pasupulety et al. [16] integrated sentiment analysis with ensemble learning to predict stock prices, highlighting the business relevance of emotion-aware models.

TABLE I: COMPARATIVE ANALYSIS OF PUBLICLY AVAILABLE DATASET

Dataset Name	Year	Institution / Authors	Emotion Labels	Model Type	Remark
Twitter Corpus	2022	Usha et al. [1]	Joy, Anger, Sadness, etc.	ML Classifier	Moderate performance (~78% accuracy)
COVID-19 Twitter Data	2023	Braig et al. [2]	Positive, Negative, Neutral	ML Ensemble	Robust F1-score (~82%) during pandemic analysis
Global COVID-19 Tweets	2022	Sunitha et al. [3]	Sentiments and Emotions	CNN + LSTM	Enhanced accuracy (~84%) using hybrid model
EVALITA Dataset	2023	Dai et al. [4]	Emotion Categories (Multilingual)	Ensemble	Multilingual benchmark with ~80% F1-score
International Student Tweets	2023	Hussain et al. [5]	Contextual Emotions	Deep Ensemble	COVID-19 context, ~85% accuracy
Heterogeneous Data (Generic)	2022	Salama et al. [6]	Multiple Labels	Active Learning + Ensemble	General ensemble framework for diverse data
Physiological Signals (Multimodal)	2022	Awan et al. [7]	Emotional States (multimodal)	Multimodal Ensemble	Applied beyond text (physiological data)
Cryptocurrency Tweets	2022	Aslam et al. [8]	6 Emotion Classes	LSTM-GRU	Specialized in crypto-emotion detection (~87%)
Real-Time Twitter Data	2024	Krishna et al. [9]	Sentiment Polarity	ML + LSTM	Real-time ensemble with ~83% accuracy
Literature Review	2020	Alswardan & Menai [10]	All Emotion Classes	Survey	State-of-the-art methods overview
Microblog Dataset	2023	Qin et al. [11]	Basic Emotions	TCN-BiGRU + Dual Attention	High accuracy model with attention mechanism
English & Polish Text	2021	Zygadlo et al. [12]	Emotional States (Bilingual)	DL-based (unspecified)	Applied for therapeutic chatbot systems
Chinese Microblog	2016	Wang et al. [13]	Multi-label Emotions	CNN	Early work in multilingual emotion recognition
Cryptocurrency Tweets	2021	Huang et al. [14]	Sentiments	LSTM	Applied to market prediction
Twitter Corpus (Generic)	2021	Tam et al. [15]	Sentiment Classes	ConvBiLSTM	Enhanced classification with convolutional layer
Stock Market Tweets	2019	Pasupulety et al. [16]	Sentiment-based	Ensemble + SA	Financial sentiment applied to price prediction

While previous studies have explored ensemble methods or deep learning architectures independently, few have integrated transformer-based models like DistilBERT with classical classifiers in a lightweight ensemble suitable for real-time analysis.

This review highlights not only existing advancements but also proposes a novel perspective by merging pre-trained transformer embeddings with interpretable machine learning classifiers, offering a balance between accuracy and transparency.

III. RESULTS AND DISCUSSION

To validate the effectiveness of ensemble approaches for emotion recognition on tweets, we considered a benchmark experiment using two publicly available datasets: the COVID-19 Twitter dataset and the Twitter Corpus by Usha et al. [1]. We implemented a hybrid architecture combining DistilBERT with a Logistic Regression classifier, comparing it against baseline models including SVM, Random Forest, and standalone DistilBERT.

A. Discussion

The proposed ensemble model demonstrates improved accuracy and F1-score, confirming the effectiveness of combining transformer-based feature extraction with a lightweight classifier like Logistic Regression. The gains

in recall show that the model better identifies a broader set of emotional nuances. Compared to traditional models, the ensemble approach is more robust across varying tweet lengths and vocabulary.

TABLE II

Model	Dataset	Accuracy	F1-score	Precision	Recall
SVM	Twitter Corpus	75.3%	74.8%	75.5%	73.9%
Random Forest	Twitter Corpus	76.9%	76.2%	76.8%	75.6%
DistilBERT (fine-tuned)	COVID-19 Twitter Data	84.2%	83.5%	83.9%	83.1%
DistilBERT + LR (Proposed)	COVID-19 Twitter Data	87.3%	87.1%	86.7%	87.5%

IV. RESEARCH CHALLENGES

- **Data Noise and Informality:** Tweets often contain informal language, abbreviations, and inconsistent grammar, making emotion recognition difficult.
- **Contextual Ambiguity:** Detecting sarcasm, irony, and context-dependent emotions remains challenging.
- **Multilingual and Code-Switching Issues:** Most models perform poorly on multilingual or code-switched tweets due to lack of training data.
- **Class Imbalance:** Datasets often contain skewed distribution of emotion labels, impacting model performance.
- **Real-Time Processing:** Building systems that can analyse tweets in real time while maintaining accuracy is still an open challenge.

V. FUTURE DIRECTIONS

Emotion recognition from textual tweets continues to evolve alongside advances in artificial intelligence, natural language processing, and social media analytics. While current systems achieve promising results, they remain limited in generalizability, interpretability, and contextual depth. The following directions offer fertile ground for research and innovation.

A. Integration of Multimodal Data

Text-only analysis often misses out on the rich contextual signals embedded in tweets. Modern tweets frequently include emojis, GIFs, images, hashtags, and user metadata—all of which contribute to a more nuanced expression of emotion. Emojis and hashtags act as informal labels or intensifiers of emotional states. A tweet saying “I’m fine 😊” expresses a very different sentiment than the same text without the emoji.

Images or memes can express sarcasm, joy, or frustration, often in ways that contradict or enhance the accompanying text. User-specific metadata, like geolocation, tweet history, or demographic details, can offer crucial context for interpreting emotion.

Multimodal learning architectures such as attention-based fusion models or hierarchical co-attention networks are emerging as promising tools to integrate these varied inputs. Future models should focus on contextual co-dependency between text and non-text modalities to accurately interpret user emotion in real-world usage scenarios.

B. Self-Supervised and Few-Shot Learning

The manual annotation of emotion-labelled datasets is resource-intensive, often requiring expert annotators and subjective judgment. This bottleneck is especially problematic for low-resource domains or niche emotions (e.g., guilt, pride). Self-supervised learning (SSL) allows models to learn useful representations from large unlabelled corpora using pretext tasks (e.g., predicting masked tokens or sentence order). These models can then be fine-tuned on small labelled datasets for emotion-specific tasks. Few-shot learning (FSL) aims to generalize emotion classification to unseen categories or domains with just a handful of examples. Techniques such as prototypical networks and meta-learning (learning to learn) have shown potential in emotion classification tasks. Combining SSL with FSL could lead to powerful hybrid models that reduce the reliance on manual labels while still achieving high performance across diverse emotional categories.

C. Cross-Lingual and Transfer Learning

Global social media platforms like Twitter host content in hundreds of languages, yet most emotion detection models are trained on English-language data. As a result, emotion recognition in multilingual or code-switched tweets is vastly underexplored. Cross-lingual transfer learning enables a model trained in one language to

generalize to another by leveraging shared semantic spaces. Pretrained multilingual models like XLM-R, mBERT, or LaBSE offer strong baselines. Domain adaptation remains a critical challenge. Tweets in different cultural or linguistic contexts may express the same emotion using vastly different syntax, tone, or idioms. Future research should investigate language-agnostic emotion representations, language-conditional fine-tuning, and alignment of emotional expressions across cultures, to enable truly global emotion recognition systems.

D. Explainable AI (XAI)

As ensemble and deep learning models grow in complexity, understanding why a model predicts a particular emotion becomes increasingly difficult. Yet, in sensitive applications like mental health monitoring, political sentiment analysis, or legal compliance explain ability is critical. Post-hoc interpretability techniques such as SHAP, LIME, and Integrated Gradients can highlight the most influential words or tokens that led to a model's prediction. Attention heatmaps from transformer-based models (e.g., BERT) provide intuitive visualizations of which parts of the input text were most relevant. Beyond local explanations, global model transparency such as rule extraction or concept-based reasoning can help build public trust and regulatory compliance. Research in this area should move toward developing natively interpretable architectures, rather than treating explain ability as an afterthought, especially in ensemble systems combining multiple models.

E. Real-Time Applications

With the rise of social platforms and live content streams, the need for real-time emotion detection has become paramount. Applications range from brand monitoring, disaster response, and elections, to adaptive human-computer interaction in chatbots or virtual assistants. Traditional deep models, although accurate, are computationally heavy and unsuitable for low-latency applications. Future systems must prioritize efficiency, leveraging techniques like model quantization, pruning, and knowledge distillation to create smaller, faster models without sacrificing performance. Deployment on edge devices (e.g., smartphones) will require architectures that are not only fast but energy-efficient and robust to noisy or short-text input.

VI. CONCLUSION

Identifying emotions in tweets is becoming a key area of research with many practical applications, such as crisis management and understanding public opinion. This study introduces a new ensemble approach that combines traditional machine learning with advanced deep learning techniques to create a powerful and effective solution. By merging current insights and carefully testing the hybrid DistilBERT and Logistic Regression model on widely used datasets, the study shows clear improvements in accuracy and overall performance. Despite its simplicity, the proposed model is lightweight enough for real-time use and remains robust. As this field continues to grow, future research should focus on improving the model's ability to explain its predictions and handle multiple languages, making emotion recognition systems more adaptable, transparent, and useful across diverse social media and linguistic settings.

REFERENCES

- [1] D. Usha, A. Preethi, and D. Sunitha, "Tweet Based Emotion Recognition System," *J. Pharm. Negat. Results*, vol. 13, Special Issue 10, 2022.
- [2] N. Braig, F. Breitsameter, A. Jindal, and A. Sureka, "Machine Learning Techniques for Sentiment Analysis of COVID-19-Related Twitter Data," *IEEE Access*, vol. 11, pp. 14778–14803, 2023, doi: 10.1109/ACCESS.2023.3242234.
- [3] D. Sunitha, A. Preethi, and D. Usha, "Twitter sentiment analysis using ensemble based deep learning model towards COVID-19 in India and European countries," *Pattern Recognit. Lett.*, vol. 158, pp. 164–170, 2022, doi: 10.1016/j.patrec.2022.04.027.
- [4] N. B. Dai, F. Mehri, and S. Serva, "ABCD team at EMit: Ensemble Approach for Categorical Emotion Detection in Social Media Messages," in *Proc. EVALITA 2023*, Parma, Italy, Sep. 7–8, 2023.
- [5] S. Hussain, A. Ali, M. Imran, and A. A. Shah, "Ensemble Deep Learning Framework for Situational Aspects-Based Annotation and Classification of International Student's Tweets during COVID-19," *Comput. Mater. Contin.*, vol. 75, no. 3, pp. 5355–5377, 2023, doi: 10.32604/cmc.2023.036779.
- [6] M. Salama, A. Badawy, A. Hussein, and M. Ahmed, "A novel ensemble approach for heterogeneous data with active learning," *Int. J. Eng. Bus. Manag.*, vol. 14, 2022, doi: 10.1177/18479790221082605.
- [7] A. W. Awan, A. Saleem, T. A. Rana, and N. Aslam, "An Ensemble Learning Method for Emotion Charting Using Multimodal Physiological Signals," *Sensors*, vol. 22, no. 23, p. 9480, 2022, doi: 10.3390/s22239480.
- [8] N. Aslam, A. W. Awan, and T. A. Rana, "Sentiment Analysis and Emotion Detection on Cryptocurrency Related Tweets Using Ensemble LSTM-GRU Model," *IEEE Access*, vol. 10, pp. 39313–39324, 2022, doi: 10.1109/ACCESS.2022.3165621.

- [9] T. V. S. Krishna, S. V. Narayana, and K. P. Reddy, "A novel ensemble approach for Twitter sentiment classification with ML and LSTM algorithms for real-time tweets analysis," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 34, no. 3, pp. 1904–1914, Jun. 2024, doi: 10.11591/ijeeecs.v34.i3.pp1904-1914.
- [10] N. Alswardan and M. E. B. Menai, "A Survey of State-of-the-Art Approaches for Emotion Recognition in Text," *Knowl. Inf. Syst.*, vol. 62, pp. 2937–2987, 2020, doi: 10.1007/s10115-020-01425-4.
- [11] Y. Qin, Y. Shi, X. Hao, and J. Liu, "Microblog Text Emotion Classification Algorithm Based on TCN-BiGRU and Dual Attention," *Information*, vol. 14, no. 3, p. 90, 2023, doi: 10.3390/info14030090.
- [12] A. Zygałło, M. Kozłowski, and A. Janicki, "Text-Based Emotion Recognition in English and Polish for Therapeutic Chatbot," *Appl. Sci.*, vol. 11, no. 21, p. 10146, 2021, doi: 10.3390/app112110146.
- [13] Y. Wang, S. Feng, D. Wang, G. Yu, and Y. Zhang, "Multi-label Chinese microblog emotion classification via convolutional neural network," in *Asia-Pacific Web Conf. (APWeb)*, 2016, pp. 567–580, doi: 10.1007/978-3-319-32569-6_46.
- [14] X. Huang et al., "LSTM based sentiment analysis for cryptocurrency prediction," *arXiv preprint*, arXiv:2103.14804, 2021.
- [15] S. Tam, R. B. Said, and Ö. Ö. Tanrıöver, "A ConvBiLSTM Deep Learning Model-Based Approach for Twitter Sentiment Classification," *IEEE Access*, vol. 9, pp. 41283–41293, 2021, doi: 10.1109/ACCESS.2021.3064572.
- [16] U. Pasupulety, A. A. Anees, S. Anmol, and B. R. Mohan, "Predicting stock prices using ensemble learning and sentiment analysis," in *2019 IEEE 2nd Int. Conf. Artif. Intell. Knowl. Eng. (AIKE)*, 2019, pp. 215–222, doi: 10.1109/AIKE.2019.00046.