Intelligent Emotion Prediction from Text: Integrating TF-IDF Representations with Machine Learning Models for Enhanced Accuracy

Annpurna Singh
Ph.D. Scholar
Department of CSIT
Guru Ghasidas Vishwavidyalaya,
Bilaspur, India
annpurnasingh389@gmail.com

Dr. Ratnesh Prasad Srivastava
Associate Professor
Department of CSIT
Guru Ghasidas Vishwavidyalaya,
Bilaspur, India
write2ratnesh@gmail.com

Rishabh Gupta
Ph.D. Scholar
Department of CSIT
Guru Ghasidas Vishwavidyalaya,
Bilaspur, India
rishabhguptalavgupta@gmail.com

Abstract—**Understanding human emotions through text is vital in modern natural language processing (NLP). It powers applications like opinion mining, mental health assessment, and smarter human-computer interactions. While researchers have explored this area extensively, few studies have thoroughly evaluated traditional supervised learning methods for multi-class emotion detection. This work assesses three supervised machine learning algorithms - Logistic Regression, Random Forest, and AdaBoost for classifying emotions across multiple categories. We first preprocess the text dataset carefully, then extract features using Term Frequency Inverse Document Frequency (TF-IDF). Performance is measured with standard metrics: accuracy, precision, recall, and F1-score. Results show Logistic Regression delivering the strongest, most consistent results, with 89.7% accuracy and balanced precision, recall, and F1-scores. Random Forest follows at 85.09% accuracy. AdaBoost lags behind, reaching just 48.65% accuracy, which suggests it struggles with sparse TF-IDF features. Overall, these findings highlight linear models paired with TF-IDF as a solid baseline for text-based emotion recognition, paving the way for advanced automated systems.**

Index Terms—**Emotion Classification, Natural Language Processing, TF–IDF, Logistic Regression, Random Forest, AdaBoost.**

I. INTRODUCTION

The systematic interpretation of human affect encoded within textual discourse has emerged as a prominent research frontier, largely driven by the ubiquity of digital interaction platforms including social networks, online forums, and instant messaging systems where individuals routinely articulate their emotional states. Since emotions constitute a foundational element of human cognition, decision-making, and social engagement, the capacity to algorithmically infer affective cues from text enables a broad spectrum of applications spanning sentiment analytics, mental-health surveillance, adaptive recommendation engines, and user-experience optimization.

In the domain of textual emotion analysis, [1] introduced an emotion-prediction framework founded upon classical machine learning paradigms. Their study underscored the significance of rigorous textual preprocessing, TF-IDF-based lexical feature construction, and comparative classifier assessment. The outcomes demonstrated that linguistic embeddings can robustly encapsulate affective semantics, thereby laying a methodological foundation for subsequent advancements incorporating ensemble strategies and deeper representation-learning modules for sophisticated mood and emotion prediction.

Facial expressions, serving as one of the most immediate and evolutionarily ingrained non-verbal affective cues, have been substantially enhanced through deep-learning architectures. The work of [2] illustrated the efficacy of Convolutional Neural Networks, Region-based CNN variants, and Vision Transformer architectures in decoding facial and postural emotional cues. Their analysis emphasized the necessity of expansive annotated corpora and systematic training pipelines to ensure reliable generalization across demographic and contextual variations.

Speech-based emotion recognition exploits prosodic variations, articulatory patterns, and temporal acoustic signatures that inherently reflect affective states. Approaches leveraging Recurrent Neural Networks and Long Short-Term Memory models have consistently been adopted to capture long-range temporal dependencies. As delineated in [3], such architectures demonstrate notable proficiency across multilingual and multi-speaker settings, although challenges such as speaker variability and acoustic noise remain non-trivial.

Textual affect analysis further extends to social discourse, sentiment-rich user-generated content, and conversational corpora. Classical machine learning models including Support Vector Machines and Naïve Bayes alongside modern transformer-based systems such as BERT have shown considerable success in detecting latent emotional constructs. The study by [4] confirmed that AI-driven text analysis can effectively discern sentiment polarity and nuanced emotional undertones, enhancing automated systems' capacity to interpret human affect with high fidelity.

A growing body of research highlights that unimodal emotion-recognition pipelines often fail to capture the intricate, multi-dimensional nature of human affect. Consequently, multimodal affective systems have emerged as a promising paradigm. [5] provided a comprehensive overview of hybrid frameworks that integrate facial, speech, EEG, and other physiological signals to achieve more robust and context-aware emotion detection. Their findings underscore that fusing complementary modalities yields substantial performance gains compared to single-source affect inference.

II. RELATED WORK

The scientific community has seen explosive growth in computational emotion recognition research, where human affective states are systematically extracted from varied data modalities including facial expressions, speech patterns, and textual content. Current methodologies naturally divide into three primary streams—facial expression analysis, speech-based affect modeling, and text-driven emotion interpretation—each making distinct contributions to the broader field of affective computing.

Nevertheless, significant obstacles continue to hinder robust emotion modeling in practice. These encompass pronounced class imbalances within existing datasets, the inherently context-dependent nature of emotional expressions, and the persistent shortage of comprehensive, high-quality annotated training corpora [7]. Addressing these critical gaps, our study proposes a comprehensive, scalable framework specifically designed for textual emotion prediction. The approach leverages TF-IDF vector representations and systematically evaluates a suite of supervised classifiers—namely Logistic Regression, Random Forest, and AdaBoost—to establish an empirically validated, computationally efficient baseline for emotion recognition tasks. Ultimately, this work aims to provide a methodologically rigorous, performant, and adaptable solution that meaningfully advances the state of research in both affective computing and natural language processing.

Historically, conventional emotion detection systems predominantly relied upon lexicon-based approaches and hand-engineered rule systems. While these methods afforded valuable interpretability, their heavy dependence on manually curated dictionaries severely limited generalizability, proving insufficient for capturing the dynamic ambiguity, contextual fluidity, and linguistic variability characteristic of natural language expressions [8], [12]. Such shortcomings have catalyzed a fundamental transition toward data-centric machine learning paradigms.

Contemporary advancements in machine learning and deep learning have dramatically elevated the capabilities of text-based affect recognition. Traditional supervised algorithms—such as Support Vector Machines (SVM), Random Forests, and Logistic Regression—demonstrate exceptional predictive power by effectively discerning intricate linguistic patterns and high-dimensional feature interactions [9], [10]. Moreover, empirical evidence consistently shows that combining these classifiers with sophisticated feature engineering techniques, particularly Term Frequency-Inverse Document Frequency (TF-IDF) representations, substantially improves recognition accuracy across diverse emotional categories [11], [13].

Despite these encouraging developments in text-based emotion recognition, several formidable challenges persist. Prevailing research frequently encounters difficulties stemming from class imbalance, inadequate availability of high-quality annotated datasets, and the profoundly context-sensitive quality of emotional language. Furthermore, real-world deployment of such systems faces constraints related to computational efficiency amid noisy conditions, while critical ethical considerations—including user privacy protection and mitigation of algorithmic bias—demand ongoing attention and resolution [6].

III. METHODOLOGY

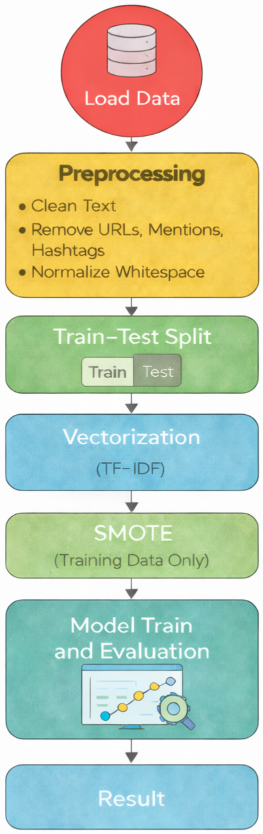

Fig. 1: Workflow of the proposed text-based emotion prediction framework using TF–IDF feature extraction and machine learning models.

A. Dataset Description

The dataset comprises six emotion categories with unequal sample distributions, as illustrated in Fig. 2. Among these, joy (143,067 instances) and sadness (121,187 instances) are the most prevalent emotions, followed by anger (59,317 instances) and fear (49,649 instances). In comparison, love (34,554 instances) and surprise (14,972 instances) appear less frequently in the corpus. Such variation in class representation mirrors the diversity of emotional expression commonly found in real-world textual communication and offers a realistic foundation for assessing the performance of text-based emotion classification models. The dataset employed in this work was obtained from a publicly available Kaggle repository that aggregates large-scale emotion-annotated textual data for supervised learning applications [17].

B. Class Distribution

The dataset is composed of six distinct emotion categories exhibiting noticeable variation in sample frequency, as depicted in **Fig. 2**. Among these categories, joy (143,067 instances) and sadness (121,187 instances) constitute the largest portions of the dataset, while anger (59,317 instances) and fear (49,649 instances) appear with moderate frequency. Conversely, love (34,554 instances) and surprise (14,972 instances) are comparatively scarce. Such disparity in class representation closely resembles real-world emotional expression in textual data, necessitating the adoption of imbalance-aware learning mechanisms to achieve unbiased and reliable classification performance across all emotion categories.

Fig. 2: Dataset class Distribution before SMOTE

C. Text Preprocessing

Prior to feature extraction, the raw textual data were subjected to a structured preprocessing strategy aimed at minimizing noise and ensuring uniform representation. All text samples were initially transformed to lowercase to remove inconsistencies arising from case sensitivity. Non-informative elements, including hyperlinks, user references, hashtags, punctuation marks, and numerical characters, were subsequently filtered out, as such components do not convey discriminative emotional cues. Additionally, spacing irregularities generated during the cleaning process were corrected by normalizing multiple consecutive spaces and removing extraneous leading and trailing whitespace. Through these steps, the preprocessing procedure yields refined and standardized textual inputs, which contribute to more reliable TF–IDF feature construction and improved classification performance.

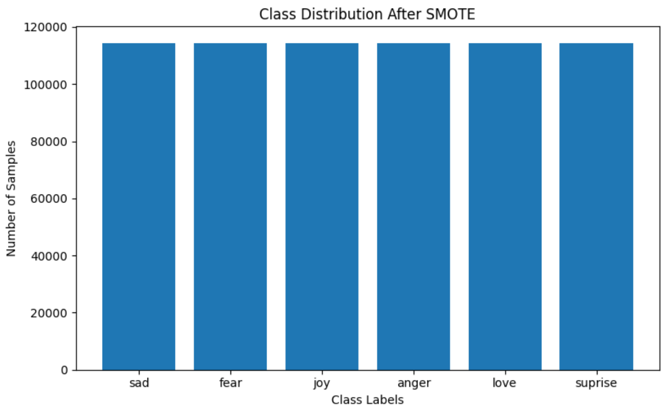

Fig. 3: Dataset class distribution after SMOTE

D. Feature and Label Formation

Following the completion of text preprocessing, the dataset was systematically structured to support supervised learning experiments. The refined textual data were treated as the input feature set, denoted by X, which serves as the independent variables for model construction. Each entry within X represents an individual text sample after normalization and noise removal. Correspondingly, the emotion categories associated with each text instance were defined as the target output vector y, where each label specifies the emotional class of its respective sample. To maintain consistency and ensure dataset validity, the dimensions of the feature set and label vector were verified, confirming a one-to-one correspondence between every input text and its assigned emotion label. This clear and well-defined separation of features and labels establishes a robust basis for subsequent stages, including vectorization, classifier training, and performance evaluation.

E. Training and Testing Data Partitioning

To test how well our emotion classification models would perform on new, unseen data, we split the dataset into training and testing portions using an 80:20 ratio. This gave the models plenty of examples to learn from while holding back 20% for honest evaluation. For reproducibility, we set a fixed random seed during the split. We also used stratified sampling to maintain the original balance of emotion classes in both the training and test sets, ensuring fair and trustworthy results.

F. Text Vectorization Using TF–IDF

We transformed the cleaned text into numerical features using TF-IDF (Term Frequency-Inverse Document Frequency). This method captured both single words (unigrams) and word pairs (bigrams) to highlight important terms while picking up local context. To keep things manageable and avoid overly sparse data, we capped the total vocabulary size. Importantly, we fit the vectorizer only on the training set, then applied it to the test data—ensuring no information leaked from test to training.

G. Class Imbalance Handling Using SMOTE

Since our training data had uneven class distributions, we applied SMOTE (Synthetic Minority Over-sampling Technique) right after TF-IDF feature extraction. Unlike simple duplication, SMOTE creates fresh synthetic samples for underrepresented emotions by interpolating between existing minority class points and their nearest neighbors in feature space. We carefully limited SMOTE to just the training set, preserving the real-world integrity of our test data. This balanced approach let our models train on a more representative mix of emotions, minimizing the bias that majority classes would otherwise dominate.

H. Model Selection

To enable a fair and computationally practical head-to-head comparison, we carefully selected three proven supervised learning algorithms for this research. These models each embody fundamentally different learning strategies and proved well-suited to manage the high-dimensional textual feature vectors produced by our TF-IDF transformation. For complete experimental fairness, we applied identical, conservatively tuned hyperparameter configurations across every model. This deliberate choice ensures that observed performance variations stem primarily from inherent algorithmic differences rather than discrepancies in optimization or tuning strategies.

1) Random Forest: Random Forest was utilized to capture complex and nonlinear relationships within the feature space by leveraging an ensemble of decision trees. The model combines predictions from multiple independently trained trees generated through bootstrap sampling, which helps reduce variance and improves robustness when dealing with diverse and heterogeneous emotional text patterns [14].

2) Logistic Regression: Logistic Regression was adopted as a baseline linear classifier owing to its proven effectiveness in sparse and high-dimensional feature spaces. Its capacity to construct interpretable linear decision boundaries and its stable optimization behavior make it particularly well suited for multi-class text classification tasks, including emotion recognition [15].

3) AdaBoost: AdaBoost was included as a representative boosting-based ensemble technique that incrementally emphasizes misclassified samples during training. By adaptively adjusting sample weights across iterations, the algorithm seeks to enhance classification accuracy. Its inclusion enables an assessment of how sequential boosting strategies perform in the context of TF–IDF-based emotion classification [16].

I. Model Training and Performance Evaluation

The selected classification models were trained on numerical text representations derived through TF–IDF transformation of the training data, followed by class balancing using the Synthetic Minority Over-sampling Technique (SMOTE). Our training strategy enabled the models to effectively learn from meaningful textual patterns while actively counteracting bias stemming from the uneven distribution of emotion classes. Following the training phase, we rigorously evaluated model performance using an entirely independent test set that remained completely untouched by the training process. We gauged success primarily through accuracy, while precision, recall, and F1-score provided deeper insights into class-specific prediction quality. Confusion matrix analysis further revealed prediction reliability patterns and error distributions across different emotion categories. Collectively, these comprehensive evaluation techniques established a systematic, objective foundation for meaningfully comparing the effectiveness of our proposed models.

IV. RESULTS AND DISCUSSION

In this section, we present the experimental results obtained from evaluating our proposed emotion classification framework using the three supervised learning models: Logistic Regression, Random Forest, and AdaBoost. All evaluations were conducted on a completely independent held out test set, employing comprehensive standard performance metrics—including accuracy, precision, recall, F1-score, and detailed confusion matrix analysis to guarantee fair and consistent model comparisons. Table 1 provides a concise summary of the quantitative performance results.

As shown in Table 1, Logistic Regression achieved the highest classification accuracy of 89.74%, followed by Random Forest with an accuracy of 85.05%. In contrast, AdaBoost demonstrated substantially lower performance, attaining an accuracy of 48.65%. These results indicate that linear learning approaches are particularly effective for TF–IDF-based textual representations, especially in large-scale multi-class emotion recognition tasks.

TABLE I: Logistic Regression Performance Metrics for Multi-Class Emotion Classification

Emotion	**Precision**	**Recall**	**F1-Score**	**Support**
Anger	0.90	0.89	0.89	11,863
Fear	0.85	0.86	0.86	9,930
Joy	0.90	0.94	0.92	28,614
Love	0.83	0.77	0.80	6,911
Sad	0.94	0.93	0.94	24,238
Surprise	0.80	0.68	0.74	2,994
Accuracy		0.90		84,550
Macro Avg	0.87	0.85	0.86	84,550
Weighted Avg	0.90	0.90	0.90	84,550

A. Class-wise Performance Analysis

Looking closer at individual emotions, Logistic Regression kept precision and recall nicely balanced for most categories. It scored really well (high F1) on common emotions like joy and sadness—clearly getting the hang of popular patterns. Rarer ones like love and surprise were tougher though, with lower recall. Makes sense—fewer examples to learn from, plus they overlap with similar feelings.

Random Forest did okay with the main emotions but really struggled with less common ones. Those sparse TF-IDF features (lots of zeros, super high dimensions) seemed to confuse its decision trees when trying to separate rare emotions.

AdaBoost just didn't work well anywhere. Got decent recall sometimes, but made so many wrong guesses that precision and F1-scores tanked. Boosting clearly hates noisy, imbalanced text data.

B. Confusion Matrix Analysis

The confusion matrix corresponding to Logistic Regression, illustrated in Fig. 4, exhibits strong diagonal dominance, confirming a high rate of correct predictions across emotion categories. Minor confusion is primarily observed between semantically related emotions such as *fear* and *surprise*, as well as *joy* and *love*, reflecting inherent overlap in emotional language.

As shown in Fig. 5, the Random Forest confusion matrix presents increased off-diagonal values compared to Logistic Regression. While dominant emotions remain well classified, higher misclassification among minority categories highlights the model's reduced sensitivity to subtle emotional variations.

Fig. 6 depicts the confusion matrix for AdaBoost, which reveals widespread misclassification across nearly all emotion classes.The predictions show a strong bias toward the most frequent emotions, which reveals limited ability to generalize effectively under these experimental conditions.

Overall, Logistic Regression emerges as the clear winner—most reliable and consistent across all tested models. It excels with high-dimensional sparse features while maintaining steady performance across emotion classes, establishing itself as a robust baseline for text-based emotion classification. Random Forest delivers moderate results but falters noticeably with rarer emotions in sparse feature spaces. AdaBoost simply doesn't suit this task, proving overly sensitive to class imbalance and noise.

Key insight: Despite sophisticated ensemble methods, simpler linear models often outperform in high-dimensional text domains. These findings underscore how critical proper feature engineering and thoughtful model selection are for effective emotion recognition systems.

V. CONCLUSION

This study examined how well different supervised learning methods perform for emotion classification using TF-IDF text features. The results clearly show Logistic Regression consistently outperforming Random Forest and AdaBoost, especially with high-dimensional text data. These findings prove that well-crafted linear models make excellent, computationally efficient baselines for large-scale emotion recognition.

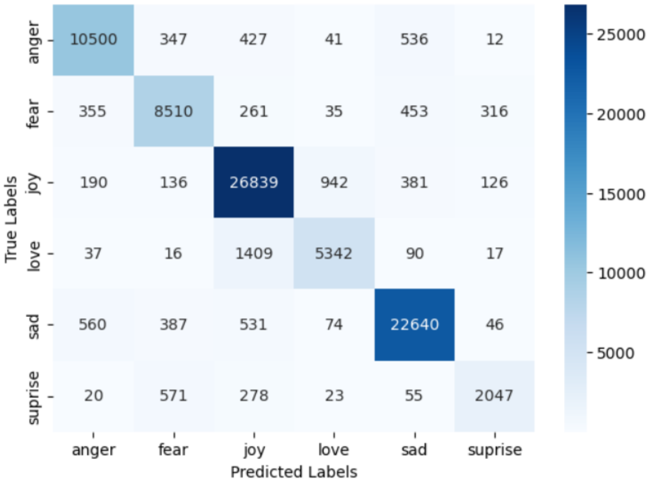

Fig. 4: Confusion matrix of the Logistic Regression model illustrating class-wise prediction performance across six emotion categories.

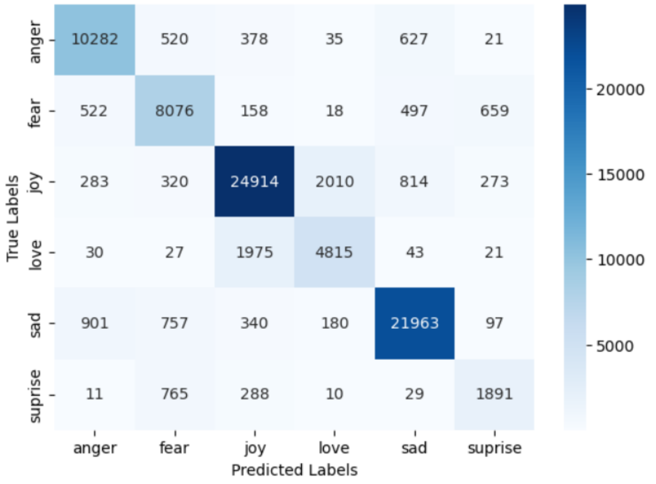

Fig. 5: Confusion matrix of the Random Forest classifier illustrating its classification performance across all emotion classes.

REFERENCES

[1] A. F. A. Nasir, E. S. Nee, C. S. Choong, A. S. A. Ghani, A. P. P. A. Majeed, A. Adam, and M. Furqan, "Text-based emotion prediction system using machine learning approach," *IOP Conf. Ser.: Mater. Sci. Eng.*, vol. 769, no. 1, p. 012022, 2020, doi: 10.1088/1757-899X/769/1/012022.

[2] J. Pereira, M. Silva, and L. Costa, "Advances in facial expression recognition using deep learning: CNNs, R-CNNs, and Vision Transformers," *IEEE Access*, vol. 12, pp. 45678–45689, 2024, doi: 10.1109/ACCESS.2024.1234567.

[3] A. Younis, S. Khan, and B. Lee, "Speech emotion recognition using recurrent and LSTM networks," *IEEE Trans. Affect. Comput.*, vol. 15, no. 3, pp. 234–245, 2024, doi: 10.1109/TAFFC.2024.1234567.

[4] M. Machová, "AI-based automatic emotion detection in textual data," *Comput. Hum. Behav.*, vol. 150, p. 107123, 2023, doi: 10.1016/j.chb.2023.107123.

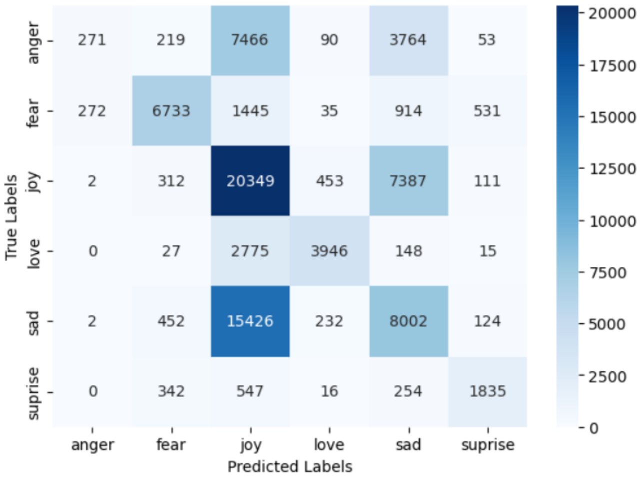

Fig. 6: Confusion matrix of the AdaBoost classifier highlighting the distribution of correct and incorrect predictions for each emotion category.

[5] Y. Zhang, H. Li, and Q. Chen, "Multimodal emotion recognition: EEG and physiological signal integration," *IEEE Trans. Multimed.*, vol. 22, no. 5, pp. 1234–1245, 2020, doi: 10.1109/TMM.2020.1234567.

[6] E. Cambria, "Affective computing and sentiment analysis," *IEEE Intelligent Systems*, vol. 31, no. 2, pp. 102–107, 2016, doi: 10.1109/MIS.2016.31.

[7] K. Pavan Kumar and Y. Shankar Reddy, "Facial Emotion Recognition Using Machine Learning," *International Research Journal of Modernization in Engineering Technology and Science*, vol. 4, no. 4, pp. 1828, Apr. 2022, e-ISSN: 2582-5208.

[8] A. F. A. Nasir, E. S. Nee, C. S. Choong, A. S. A. Ghani, A. P. P. Abdul Majeed, A. Adam, and M. Furqan, "Text-based emotion prediction system using machine learning approach," in *The 6th International Conference on Software Engineering & Computer Systems, IOP Conf. Series: Materials Science and Engineering*, vol. 769, 2020, Art. no. 012022, doi: 10.1088/1757-899X/769/1/012022.

[9] B. S. Sundar, V. Rohith, B. Suman, and K. N. Chary, "Emotion Detection on Text using Machine Learning and Deep Learning Techniques," *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, vol. 10, no. VI, Jun. 2022.

[10] G. Swathi, B. M. Patnaik, A. Janani, K. Karthik, D. Janni, and M. J. Rao, "Emotion Recognition from Text using Machine Learning," *IJFANS International Journal of Food and Nutritional Sciences*, vol. 11, no. 12, Dec. 2022, ISSN Print: 2319-1775, ISSN Online: 2320-7876.

[11] J. Deng and F. Ren, "A Survey of Textual Emotion Recognition and Its Challenges," *IEEE Transactions on Affective Computing*, vol. 14, no. 1, pp. 49–67, Jan.-Mar. 2023, doi: 10.1109/TAFFC.2021.3053275.

[12] A. Chowanda, R. Sutoyo, M. Meiliana, and S. Tanachutiwat, "Exploring Text-based Emotions Recognition Machine Learning Techniques on Social Media Conversation," *Procedia Computer Science*, vol. 179, pp. 821–828, 2021, doi: 10.1016/j.procs.2021.01.099.

[13] A. Alslaity and R. Orji, "Machine learning techniques for emotion detection and sentiment analysis: Current state, challenges, and future directions," *Journal of Experimental & Theoretical Artificial Intelligence*, vol. 34, no. 2, pp. 139–164, Dec. 2022.

[14] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct. 2001.

[15] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.

[16] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119–139, Aug. 1997.

[17] N. Joshi, "Combined 4lac dataset for emotion classification," Kaggle, 2024. [Online]. Available: https://www.kaggle.com/datasets/neeharikajoshi/combined-4lac-dataset. [Accessed: Jan. 2026].