JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7321

THE EMOTIONAL ANALYSIS OF SONG LYRICS AND VIDEO COMMENTS ON YOUTUBE USING DEEP LEARNING

Nova Noor Kamala Sari[1*]; Widiatry[1]; Inda Fitria Maharisty[1]

Informatics Engineering Department, Faculty of Engineering[1]
University of Palangka Raya, Palangka Raya, Indonesia[1]
https://upr.ac.id/[1]
novanoorks@it.upr.ac.id*, widiatry@it.upr.ac.id, indafitria@mhs.eng.upr.ac.id

(*) Corresponding Author
(Responsible for the Quality of Paper Content)

Abstract— *Digital media platforms shape public perception of music through song lyrics and audience comments. This study analyzes emotions expressed in the lyrics and YouTube comments of Taylor Swift's "Fortnight" using deep learning models. The dataset consists of 42 lyric lines and 13,406 user comments collected from April to December 2024. Emotion labeling was manually performed based on Plutchik's eight basic emotions with an additional neutral category. This research applies two models: Long Short-Term Memory (LSTM) and DistilRoBERTa, with random oversampling to address class imbalance. Performance was evaluated using accuracy, precision, recall, F1-score, and confusion matrices. As a single-song case study, this research provides a focused comparison of sequential and transformer-based architectures for simultaneous emotion analysis of lyrics and audience responses. The results show that DistilRoBERTa achieved higher accuracy (94.07%) than LSTM (90.75%), indicating the advantage of contextual transformer models in capturing nuanced emotional expressions within this dataset. However, the findings are limited to the thematic characteristics of this single-song dataset and should be interpreted within this contextual scope.*

Keywords: *DistilRoBERTa, Emotion Classification, LSTM, Lyrics, YouTube Comments*

Intisari— *Platform media digital membentuk persepsi publik terhadap musik melalui lirik lagu dan komentar audiens. Penelitian ini menganalisis emosi yang diekspresikan dalam lirik dan komentar YouTube dari lagu Taylor Swift "Fortnight" menggunakan model deep learning. Dataset terdiri atas 42 baris lirik dan 13.406 komentar pengguna yang dikumpulkan dari April hingga Desember 2024. Pelabelan emosi dilakukan secara manual berdasarkan delapan emosi dasar Plutchik dengan tambahan satu kategori netral. Penelitian ini menerapkan dua model, yaitu Long Short-Term Memory (LSTM) dan DistilRoBERTa, dengan teknik random oversampling untuk mengatasi ketidakseimbangan kelas. Kinerja model dievaluasi menggunakan metrik akurasi, presisi, recall, F1-score, serta confusion matrix. Sebagai studi kasus pada satu lagu, penelitian ini memberikan perbandingan terfokus antara arsitektur sekuensial dan berbasis transformer dalam analisis emosi secara simultan pada lirik dan respons audiens. Hasil penelitian menunjukkan bahwa DistilRoBERTa mencapai akurasi lebih tinggi (94,07%) dibandingkan LSTM (90,75%), yang mengindikasikan keunggulan model transformer dalam menangkap ekspresi emosional yang kontekstual dan bernuansa pada dataset ini. Namun, temuan ini terbatas pada karakteristik tematik dari dataset satu lagu tersebut dan perlu diinterpretasikan dalam konteks yang spesifik.*

Kata Kunci: *Distilroberta, Klasifikasi Emosi, LSTM, Lirik, Komentar Youtube.*

INTRODUCTION

The rapid development of digital technology has transformed how audiences experience and respond to music. Beyond listening, music now functions as a medium of textual interaction, particularly through song lyrics and user-generated comments on digital platforms. Song lyrics reflect the emotional intentions and inner experiences of creators, while audience comments represent interpretative and affective responses after engaging with the song. Therefore, both lyrics and comments constitute complementary textual sources for examining emotional expression from the perspectives of creators and listeners [1], [2].

Emotion analysis is a branch of natural language processing that focuses on identifying and classifying emotional states expressed in text, such as joy, sadness, anger, or fear [3]. Various approaches have been applied in this field, including lexicon-based techniques, traditional machine learning algorithms, and deep learning methods [4]. Recent studies have highlighted the growing importance of deep learning techniques in improving the effectiveness of emotion detection across various textual domains [5]. In particular, deep learning models have demonstrated superior performance due to their capacity to capture contextual and semantic relationships within textual data [6]

Previous studies have confirmed the effectiveness of sequential models such as Long Short-Term Memory (LSTM) and transformer-based architectures such as BERT in text classification tasks. LSTM-based architectures have demonstrated strong performance in sentiment and text classification due to their ability to capture long-range contextual dependencies within sequential textual data [7]. Similarly, Alhadlaq and Altheneyan reported that DistilRoBERTa achieved strong performance in sentiment analysis by effectively modeling linguistic context [8]. While LSTM is designed to process sequential information and preserve long-term dependencies, BERT and its variants employ a bidirectional transformer architecture that enables deeper contextual understanding by considering relationships between words in both directions. This capability is particularly important in natural language processing tasks where semantic meaning is highly dependent on contextual information.

Although these studies highlight the strengths of both architectures, most prior research has focused on product reviews, service evaluations, or general social media content. Several recent studies have also explored emotion classification in social media platforms using deep learning and transformer-based approaches; particularly in short-form public discourse such as tweets and online discussions [9]. However, limited research has simultaneously examined emotional expressions in artistic textual content, such as song lyrics, and audience responses in YouTube comments within a unified analytical framework [10], [11]. Most existing studies analyze either creator-generated text or user-generated content separately, without integrating both perspectives. This gap is significant because lyrics represent intended emotional expression, while comments reflect audience interpretation, forming a dynamic emotional interaction in digital music environments.

Taylor Swift's song "Fortnight" was selected as the case study due to its high popularity and extensive audience engagement. The song was released on April 19, 2024, as part of the album The Tortured Poets Department. The song, a collaboration with Post Malone, reached number one on the Billboard Hot 100 and achieved record-breaking streaming performance on Spotify. In addition, the music video has attracted significant viewership on YouTube, reflecting strong audience interaction. These factors make "Fortnight" a relevant subject for emotion analysis based on both its lyrical content and user-generated comments [12]. This study aims to examine the comparative performance of LSTM and DistilRoBERTa in classifying emotions in song lyrics and YouTube comments. It further seeks to identify differences in emotion distribution between creator-generated text (lyrics) and audience-generated text (comments), as well as to determine which emotion categories present greater classification challenges within the scope of a single-song case study.

To achieve these objectives, this research applies LSTM and DistilRoBERTa (j-hartmann/emotion-english-distilroberta-base) for emotion classification based on Plutchik's eight basic emotions—anger, fear, disgust, sadness, anticipation, joy, surprise, and trust—with an additional neutral category. Emotion labeling is performed manually by human annotators. By comparatively evaluating sequential and transformer-based architectures on both lyrics and comments, this study provides an empirical assessment of their effectiveness within a focused dataset.

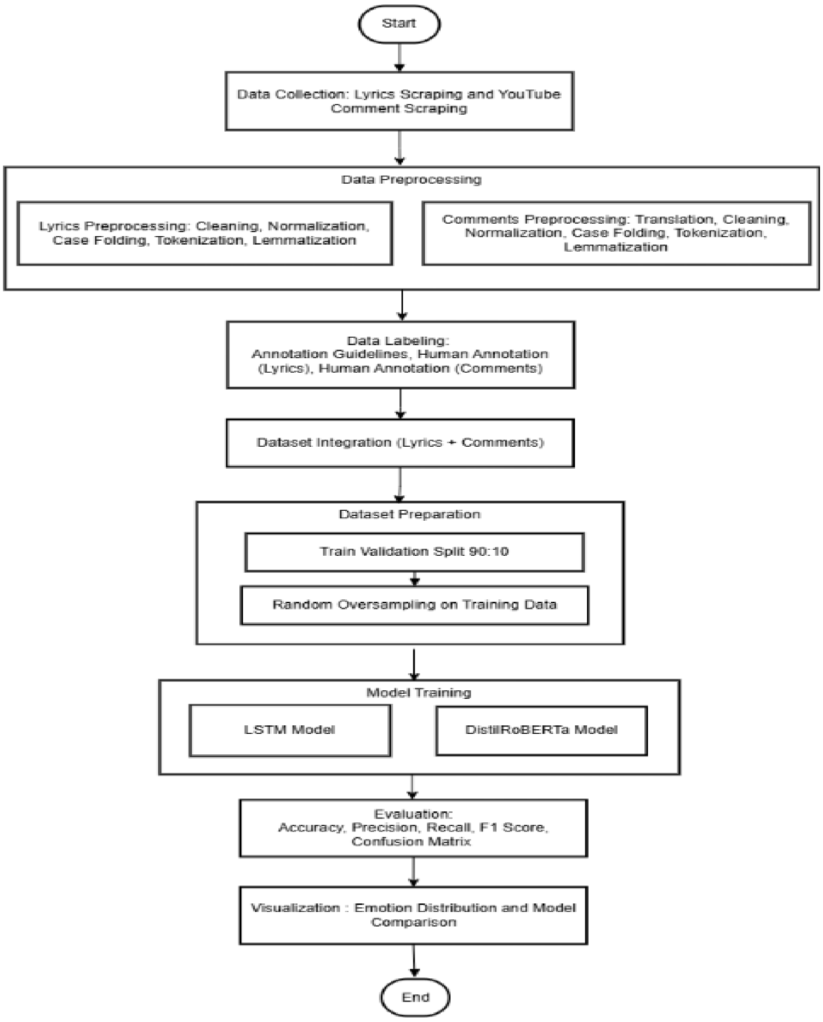
The remainder of this paper is organized as follows. Section 2 describes the materials and methods, including data collection, preprocessing, labeling, dataset preparation, and modeling. Section 3 presents the results and discussion. Section 4 concludes the study and outlines its limitations.

MATERIALS AND METHODS

The research methodology follows a structured workflow consisting of data collection, preprocessing, data labeling, dataset integration, dataset preparation, model training, evaluation, and visualization. The overall research pipeline is illustrated in Figure 1.

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7321



Source: (Research results, 2026)

Figure 1. Research methodology

1. Data Collection

Data were collected from two primary sources representing different forms of textual emotional expression. Song lyrics from Taylor Swift's "Fortnight" were obtained through web scraping from the Genius.com platform. Audience responses were collected from the official "Fortnight" music video on Taylor Swift's YouTube channel using the YouTube Data API v3. In this study, the term video comments refers specifically to comments posted on the YouTube music video. The comment collection process was conducted from April 20 to December 5, 2024. In total, 42 lyric lines and 14,404 YouTube comments were initially obtained. These two data sources capture complementary perspectives of emotional communication, where song lyrics reflect the emotional intention of the creator while audience comments represent emotional responses from listeners [13], [14].

2. Data Pre-processing

Text preprocessing was performed to clean and normalize textual data before model training. Several standard natural language processing techniques were applied [15]. First, all user comments were translated into English to ensure language consistency within the dataset. Noise removal was then performed by eliminating emojis, URLs, mentions, hashtags, numbers, punctuation marks, and other non-alphabetic characters [16], [17]. Next, normalization was applied to convert informal expressions, abbreviations, and slang into standard word forms . Case folding was performed to convert all characters to lowercase [18]. Tokenization was then applied to split sentences into individual word tokens. Finally, lemmatization was conducted to reduce words to their base forms while preserving semantic meaning [19].

After preprocessing, the number of usable comments decreased from 14,404 to 13,406 due to the removal of empty or unprocessable entries. The lyric dataset remained unchanged with 42 lines.

3. Data Labelling

Emotion labeling was performed manually by human annotators. Manual annotation remains important in emotion classification tasks because emotional interpretation is often subjective and context-dependent [20]. Each text was assigned to one of eight basic emotion categories: anger, fear, disgust, sadness, anticipation, joy, surprise, and trust. An additional neutral category was used when no dominant emotion was identified. The labeling scheme was based on Plutchik's theory of basic emotions, which categorizes human emotions into eight primary emotional dimensions.

To ensure labeling consistency, annotators followed predefined annotation guidelines that included definitions and contextual examples for each emotion category. The labeling process was conducted separately for song lyrics and YouTube comments to preserve contextual interpretation.

Inter-Annotator Agreement (IAA) was not formally calculated in this study. Therefore, the labeling process may contain subjective interpretations depending on annotators' understanding of emotional context. This limitation is acknowledged in the interpretation of the results [21]

4. Data Integration

After the labeling process, the annotated lyric and comment datasets were combined into a single dataset for model training and evaluation. Integrating these labeled datasets enables the model to learn emotional patterns from both creator-generated and audience-generated text. This integration also facilitates comparative analysis between emotional expressions found in song lyrics and those expressed by audiences through YouTube comments .

5. Dataset Preparation

After dataset integration, the labeled dataset was prepared for model training. The dataset was divided into training and validation sets using a 90:10 split. To address class imbalance, random oversampling was applied only to the training dataset. This technique duplicates samples from minority classes until class distributions become more balanced [22]. Applying oversampling exclusively to the training data prevents data leakage and ensures unbiased validation results.

6. Model Training

Emotion classification was conducted using two deep learning models: Long Short-Term Memory (LSTM) and DistilRoBERTa. Both models were trained using the prepared dataset and employed random oversampling to address class imbalance.

a. LSTM Model

For the LSTM model, labeled text was converted into numerical sequences using space-based tokenization with a maximum length of 64 tokens per input. A vocabulary dictionary was constructed, and padding was applied to ensure uniform sequence length. Random oversampling was used to balance the emotion classes by duplicating samples from minority classes.

The LSTM architecture consisted of an embedding layer, three LSTM layers, and a linear output layer. Model training employed a weighted CrossEntropyLoss function and the AdamW optimizer with a learning rate of 0.001. Key parameters included an embedding and hidden dimension of 128, a dropout rate of 0.3, a batch size of 64, and a maximum of 10 epochs. Early stopping was applied with a patience of three epochs to mitigate overfitting.

b. DistilRoBERTa Model

The DistilRoBERTa-based model utilized the *j-hartmann/emotion-english-distilroberta-base* architecture. Emotion labels, which were originally in textual form, are converted to numerical format according to a predefined mapping. To address the class imbalance problem, a random oversampling technique is employed, increasing the number of samples from the minority class, which ensures a more balanced label distribution by randomly duplicating samples so that the number of samples in the minority class matches the number of samples in the majority class [35]. The text is then processed using the built-in DistilRoBERTa tokenizer, which has a maximum length of 128 tokens and auto-completion to ensure uniform input length.

The model is adjusted to generate nine output labels, corresponding to the number of emotion classes. Training is performed using the AdamW optimizer with a learning rate of 0.00002 and a weight distribution for regularization. To maintain training stability, gradient accumulation is applied every four batches, along with a linear learning rate schedule with a warm-up of 10% of the total training steps. The training process lasts a maximum of 10 epochs, and the batch size is 64.

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7321

7. Evaluation

Model performance was evaluated using validation data to assess generalization capability. Predictions generated by both models were compared with ground truth labels using accuracy, precision, recall, and F1-score metrics. These evaluation metrics are widely used in sentiment and emotion classification studies to assess both overall classification effectiveness and class-specific predictive performance [23]. In addition, confusion matrices were employed to visualize classification performance across emotion categories and to identify misclassification patterns. This evaluation enabled a direct comparison between the LSTM and DistilRoBERTa models.

8. Visualization

Visualization was conducted to illustrate the distribution of emotions in song lyrics and YouTube video comments. Emotion frequencies were presented using stacked bar charts, representing the intensity and prevalence of each emotion category. The visualization facilitated comparison of emotional patterns between lyrical content and audience responses, highlighting dominant emotions in both data sources.

RESULTS AND DISCUSSION

Results

This section presents the performance results of the LSTM and DistilRoBERTa models in classifying emotions from song lyrics and YouTube video comments. The evaluation was conducted using a validation dataset consisting of 4,217 entries derived after the random oversampling process.
Prior to oversampling, the integrated dataset contained 13,448 labeled texts consisting of song lyric lines and YouTube comments. To address class imbalance during model training, random oversampling was applied to the training data, increasing the dataset size to 42,165 entries. Of this total, 37,948 entries were used for training, while 4,217 entries were reserved for validation and model evaluation.

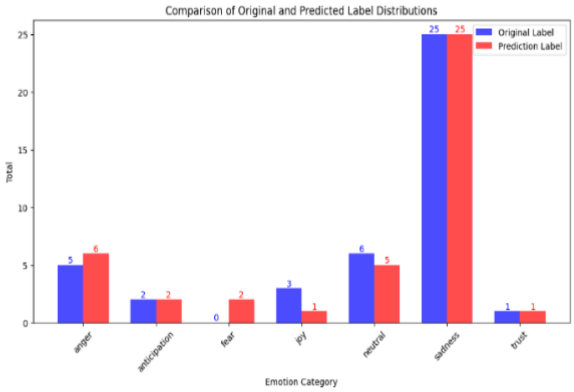
Both models demonstrated strong performance in the emotion classification task. However, the transformer-based DistilRoBERTa model achieved higher predictive performance than the LSTM model. Specifically, DistilRoBERTa obtained an accuracy of 94.07%, while the LSTM model achieved 90.75% accuracy. Detailed evaluation metrics including accuracy, precision, recall, and F1-score are presented in Table 1.

Table 1. Model evaluation results

Model	Accuracy	Precision	Recall	F1-Score
LSTM	90.75%	90.59%	90.84%	90.57%
DistilRoBERTa	94.07%	93.99%	94.07%	93.91%

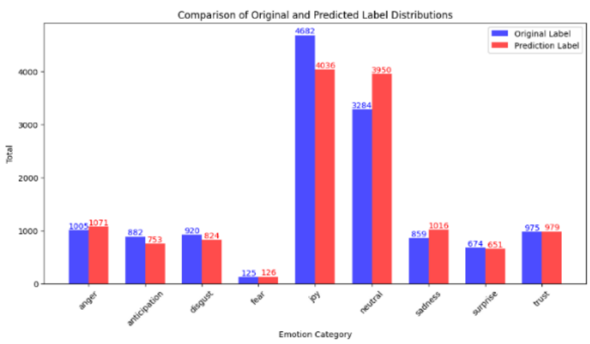
Source: (Research results, 2026)

The distribution of emotions in song lyrics predicted by the LSTM model and annotated by human annotators is illustrated in **Figure 2**. The figure shows that sadness is the dominant emotion in the lyrical content. Minor differences between the human annotations and model predictions are observed, such as the detection of fear by the model that was not originally identified by the annotators.



Source: (Research results, 2026)
Figure 2. Comparison of Emotion Distribution Between Human Annotation and LSTM Model Prediction on Song Lyrics

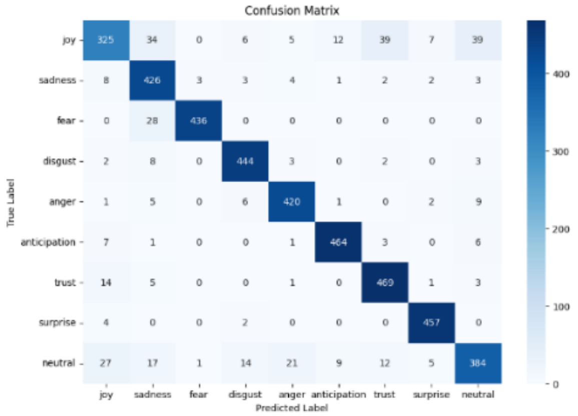
The emotion distribution for YouTube comments based on human annotations and LSTM predictions is presented in Figure 3. Joy appears as the dominant emotion in audience responses, reflecting positive engagement from listeners toward the song.



Source: (Research results, 2026)
Figure 3. Comparison of Emotion Distribution Between Human Annotation and LSTM Model Prediction on YouTube Video Comments

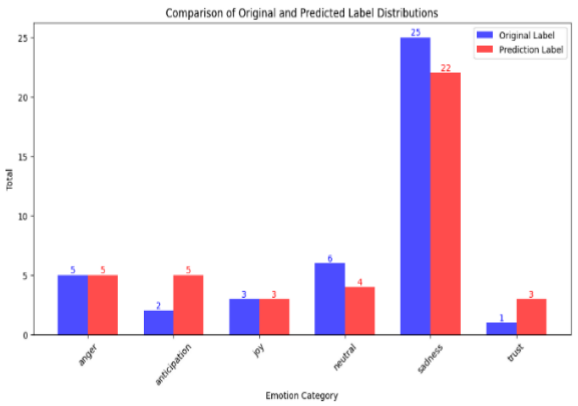
VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7321

To further examine the classification performance of the LSTM model, the confusion matrix is presented in Figure 4. Most predictions are concentrated along the main diagonal, indicating that the model correctly classified the majority of samples across emotion categories. However, several misclassifications are observed, particularly in the neutral category.



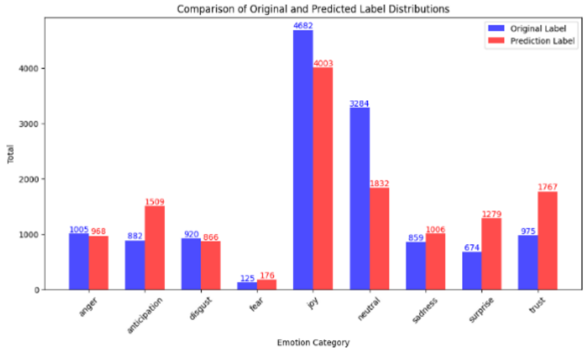
Source: (Research results, 2026)
Figure 4. Confusion Matrix of the LSTM Model

For the transformer-based model, the emotion distribution comparison between human annotations and model predictions for song lyrics is presented in Figure 5. The figure shows that the predicted emotion distribution closely aligns with the annotations provided by human annotators. Minor differences can still be observed in several emotion categories, indicating slight variations in how the model interprets contextual emotional cues in lyrical expressions.



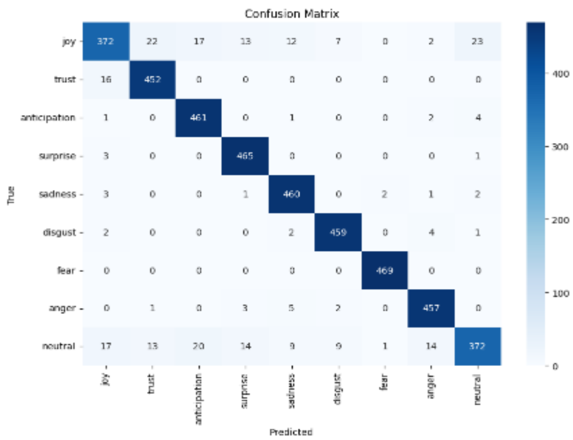
Source: (Research results, 2026)
Figure 5. Comparison of Emotion Distribution Between Human Annotation and DistilRoBERTa Model Prediction on Song Lyrics

The comparison of emotion distribution between human annotations and model predictions for YouTube comments is illustrated in Figure 6. Similar to the lyric analysis, the predicted distribution generally follows the pattern identified through manual annotation. However, small differences in frequency appear in certain emotion categories, which may reflect the variability of emotional expressions in user-generated comments.



Source: (Research results, 2026)
Figure 6. Comparison of Emotion Distribution Between Human Annotation and DistilRoBERTa Model Prediction on YouTube Video Comments

The classification performance of the DistilRoBERTa model is further illustrated through the confusion matrix shown in Figure 7. Similar to the LSTM model, most predictions lie along the diagonal line, indicating accurate classification results across multiple emotion categories.



Source: (Research results, 2026)
Figure 7. Confusion Matrix of the DistilRoBERTa Model

Discussion

The experimental results demonstrate that the DistilRoBERTa model achieved higher performance compared to the LSTM model in classifying emotions from song lyrics and YouTube comments. This result suggests that transformer-based architectures are more effective in capturing

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7321

contextual relationships within textual data, particularly when dealing with emotionally nuanced language.

Despite the high overall accuracy achieved by both models, several misclassification patterns were observed in the confusion matrices. The neutral category was the most frequently misclassified emotion. This may be caused by semantic overlap between neutral expressions and mildly positive emotions such as joy, trust, or anticipation. In many cases, short user comments contain limited contextual cues, which makes it difficult for the model to distinguish between neutral sentiment and subtle emotional expressions.

Misclassification also occurred between anticipation and surprise, as well as between joy and trust. These overlaps are likely related to linguistic similarities and the contextual nature of emotional expressions in informal online communication. For instance, user comments that express excitement about the song may contain wording that simultaneously conveys anticipation and joy.

Another factor that may influence classification performance is the application of random oversampling during model training. While oversampling helps address class imbalance by increasing the representation of minority emotion categories, it may also introduce duplicated samples that could affect the diversity of training data. Nevertheless, the use of oversampling was limited to the training dataset to prevent data leakage and maintain unbiased evaluation results.

The differences between human annotations and model predictions observed in the emotion distribution visualizations also highlight the inherent subjectivity of emotion interpretation. Emotional expressions in song lyrics are often metaphorical or context-dependent, which may lead to variations in interpretation between annotators and automated models.

Furthermore, the findings of this study should be interpreted within the scope of its single-song case study design. The emotional distribution observed in this dataset is influenced by the thematic characteristics of the song “Fortnight,” which predominantly conveys reflective and melancholic tones. As a result, the model performance reported in this study may vary when applied to songs with different lyrical styles, emotional themes, or audience engagement patterns.

Overall, the results indicate that transformer-based architectures such as DistilRoBERTa provide improved contextual understanding for emotion classification tasks involving complex textual data. However, sequential models like LSTM still demonstrate competitive performance and remain a viable approach for emotion analysis in text-based datasets.

CONCLUSION

This study examined the comparative performance of LSTM and DistilRoBERTa models for emotion classification in song lyrics and YouTube comments using a case study of Taylor Swift’s “Fortnight.” The experimental results indicate that both models are capable of effectively identifying emotional expressions in textual data, with the transformer-based DistilRoBERTa model demonstrating higher classification performance than the LSTM model.

The findings highlight the importance of contextual language understanding in emotion analysis tasks. DistilRoBERTa was able to capture nuanced emotional expressions more effectively, particularly in user-generated comments where emotional cues may be implicit or expressed informally. In contrast, the LSTM model still achieved competitive performance, indicating that sequential architectures remain a viable approach for text-based emotion classification.

This study contributes to the growing body of research on emotion analysis in digital music contexts by providing a comparative evaluation of deep learning architectures applied to both creator-generated text (song lyrics) and audience-generated text (YouTube comments). The results provide insights into how different modeling approaches interpret emotional expressions within a shared digital media environment.

However, several limitations should be acknowledged. First, this study is based on a single-song case study, which may limit the generalizability of the findings to other musical works or genres. Second, emotion labeling was conducted manually without formal measurement of Inter-Annotator Agreement (IAA), which may introduce subjective interpretations in the annotation process.

Future research may extend this work by incorporating larger multi-song datasets, applying automated or multi-annotator labeling strategies, and exploring additional transformer architectures to further improve emotion classification performance in music-related textual data.

REFERENCE

[1] H. Zhao, P. Ru, and H. Jia, "TikTok Users Migration to Xiaohongshu (Rednote): Emotional Dynamics, Platform Governance, and an NCA-SEM Analysis in Cross-Cultural Adaptation," Int. J. Hum. Comput. Interact., vol. 42, no. 4, pp. 2787 – 2812, 2026, doi: 10.1080/10447318.2025.2530088.

[2] C.-J. Wang, X. Zhang, Z. Gou, and Y. Wu, "Yesterday once more: collective storytelling and public engagement with digital cultural products on the music streaming platform," Humanit. Soc. Sci. Commun., vol. 11, no. 1, 2024, doi: 10.1057/s41599-024-03636-8.

[3] P. Nandwani and R. Verma, "A review on sentiment analysis and emotion detection from text," Soc. Netw. Anal. Min., vol. 11, 2021, doi: 10.1007/s13278-021-00776-6.

[4] F. A. Acheampong, H. Nunoo-Mensah, and W. Chen, "Transformer models for text-based emotion detection: a review of BERT-based approaches," Artif. Intell. Rev., vol. 54, no. 8, pp. 5789–5829, Dec. 2021, doi: 10.1007/s10462-021-09958-2.

[5] T. Chutia and N. Baruah, "A review on emotion detection by using deep learning techniques," Artif. Intell. Rev., vol. 57, no. 8, Aug. 2024, doi: 10.1007/s10462-024-10831-1.

[6] S. Minaee, N. Kalchbrenner, E. Cambria, N. Nikzad, M. Chenaghlu, and J. Gao, "Deep Learning–based Text Classification: A Comprehensive Review," ACM Comput. Surv., vol. 54, no. 3, Apr. 2021, doi: 10.1145/3439726.

[7] R. Huang, "Comparative Study of LSTM and Transformer Models for Sentiment Analysis on the IMDB 5000 Dataset," in 2025 4th International Conference on Artificial Intelligence, Internet and Digital Economy (ICAID), 2025, pp. 209–218. doi: 10.1109/ICAID65275.2025.11034482.

[8] A. Alhadlaq and A. Altheneyan, "Distilroberta2gnn: a new hybrid deep learning approach for aspect-based sentiment analysis," PeerJ Comput. Sci., vol. 10, 2024, doi: 10.7717/PEERJ-CS.2267.

[9] R. Olusegun, T. Oladunni, H. Audu, Y. A. O. Houkpati, and S. Bengesi, "Text Mining and Emotion Classification on Monkeypox Twitter Dataset: A Deep Learning-Natural Language Processing (NLP) Approach," IEEE Access, vol. 11, pp. 49882–49894, 2023, doi: 10.1109/ACCESS.2023.3277868.

[10] T. Shaikh and A. Jadhav, "Leveraging Deep Learning and NLP for Data Analysis and Classification," in 2025 International Conference on Applications of Machine Intelligence and Data Analytics, ICAMIDA 2025, 2025. doi: 10.1109/ICAMIDA64673.2025.11209701.

[11] S. Chaudhary, P. Kapoor, J. Rani, A. K. Kulandasamy, R. Shobana, and T. Jadhav, "NLP-BASED MUSIC LYRIC ANALYSIS IN EDUCATION," ShodhKosh: Journal of Visual and Performing Arts, vol. 6, no. 5s, pp. 239 – 249, 2025, doi: 10.29121/shodhkosh.v6.i5s.2025.6884.

[12] K. Takiguchi, M. Nakahara, and K. Sakamoto, "Influence of Comment Sentiment on YouTube Subscribers and View Counts by Genre," in Proceedings of The 2026 International Conference on Artificial Life and Robotics, 2026.

[13] S. S. Maw, E. C. Lwin, W. Mar, N. S. Paw, M. M. Khaing, and T. T. Aung, "Sentiment Analysis with YouTube Comments Using Deep Learning Approaches," in Proceedings of the 21st IEEE International Conference on Computer Applications 2024, ICCA 2024, 2024, pp. 9 – 15. doi: 10.1109/ICCA62361.2024.10532851.

[14] C. Travanca, M. Cruz, and A. Oliveira, "Emotion in Words: The Role of Ed Sheeran and Sia's Lyrics on the Musical Experience," Computers, vol. 14, no. 11, 2025, doi: 10.3390/computers14110460.

[15] A. Jabbar, S. Iqbal, M. I. Tamimy, A. Rehman, S. A. Bahaj, and T. Saba, "An Analytical Analysis of Text Stemming Methodologies in Information Retrieval and Natural Language Processing Systems," IEEE Access, vol. 11, pp. 133681–133702, 2023, doi: 10.1109/ACCESS.2023.3332710.

[16] V. H. Pranatawijaya, N. N. K. Sari, R. A. Rahman, E. Christian, and S. Geges, "Unveiling User Sentiment: Aspect-Based Analysis and Topic Modeling of Ride-Hailing and Google Play App Reviews," Journal of Information Systems Engineering and Business Intelligence, vol. 10, no. 3, pp. 328–339, Oct. 2024, doi: 10.20473/jisebi.10.3.328-339.

[17] S. Chakraborty, P. Das, S. Mahmud Dipto, M. A. Pramanik, and J. Noor, "An Analytical Review of Preprocessing Techniques in Bengali Natural Language Processing," IEEE Access, vol. 13, pp. 112428–112445, 2025, doi: 10.1109/ACCESS.2025.3574234.

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7321

[18] M. U. Albab, Y. K. P., and M. N. Fawaiq, “Optimization of the Stemming Technique on Text Preprocessing President 3 Periods Topic,” Jurnal Transformatika, vol. 20, no. 2, pp. 1–12, Sep. 2023, doi: 10.26623/transformatika.v20i2.5374.

[19] A. Jabbar, S. Iqbal, M. I. Tamimy, A. Rehman, S. A. Bahaj, and T. Saba, “An Analytical Analysis of Text Stemming Methodologies in Information Retrieval and Natural Language Processing Systems,” IEEE Access, vol. 11, pp. 133681–133702, 2023, doi: 10.1109/ACCESS.2023.3332710.

[20] A. Koufakou and E. Nieves, “Emotion-Annotated Data in NLP: Perspectives on Recent Resources and Practices,” in Eighth Widening NLP Workshop (WiNLP 2024) Phase II, 2024. [Online]. Available: https://openreview.net/forum?id=TTEBCqtfTm

[21] H. Guerdelli, C. Ferrari, J. B. Cardia Neto, S. Berretti, W. Barhoumi, and A. Del Bimbo, “Towards a Better Understanding of Human Emotions: Challenges of Dataset Labeling,” Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 14365 LNCS, pp. 242 – 254, 2024, doi: 10.1007/978-3-031-51023-6_21.

[22] D. Edmonds and J. Sedoc, “Multi-Emotion Classification for Song Lyrics,” in Proceedings of the Eleventh Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, Online: Association for Computational Linguistics, Apr. 2021, pp. 221–235. [Online]. Available: https://aclanthology.org/2021.wassa-1.24/

[23] N. Raghunathan and K. Saravanakumar, “Challenges and Issues in Sentiment Analysis: A Comprehensive Survey,” IEEE Access, vol. 11, pp. 69626–69642, 2023, doi: 10.1109/ACCESS.2023.3293041.