Improving Emotion Recognition Accuracy with Combination of Bidirectional and Long Short-Term Memory Models

Erna Haerani[1], Alam Rahmatulloh[2]*, Randi Rizal[2,3]

[1]Department of Information System, Universitas Siliwangi, Tasikmalaya (Indonesia)
[2]Department of Informatics, Universitas Siliwangi, Tasikmalaya (Indonesia)
[3]Faculty of Artificial Intelligence and Cyber Security, Universiti Teknikal Malaysia Melaka (Malaysia)

* Corresponding author: alam@unsil.ac.id

International Jornal of Informatics and Computing

Received 16 September 2025 | Accepted 21 October 2025 | Early Access 28 November 2025

Abstract

Emotions play a vital role in shaping human behavior and mental health, making accurate emotion recognition essential for mitigating potential negative impacts. This study explores the application of Bidirectional Long Short-Term Memory (Bi-LSTM) for recognizing emotions from text-based data. Bi-LSTM extends the standard LSTM by enabling the model to process input sequences in both forward and backward directions, thereby capturing contextual dependencies more effectively. The research methodology consists of data collection, manual emotion labeling, and pre-processing techniques, including stemming, tokenization, and one-hot encoding. Visualization of the dataset and the distribution of labeled emotions was conducted to gain deeper insights into the data. The Bi-LSTM model was trained for 25 epochs, achieving a training accuracy of 0.9954 and validation accuracy of 0.8790, along with a training loss of 0.0133 and validation loss of 0.658. A confusion matrix was used to further evaluate model performance and classification accuracy across various emotion categories. The experimental results confirm that the Bi-LSTM model is highly effective in recognizing emotions from textual input. Its ability to capture long-term dependencies in both directions contribute to improved learning and prediction. However, opportunities for enhancement remain, particularly in refining the model architecture, expanding the dataset, and exploring additional feature extraction techniques. This research demonstrates the potential of Bi-LSTM in building intelligent emotion-aware systems for applications in mental health monitoring, customer feedback analysis, and human-computer interaction.

Keywords

Emotion Recognition,
Bi-LSTM,
Text Classification,
Deep Learning,
Natural Language Processing

I. Introduction

Emotion constitutes a fundamental dimension of human life, shaping cognition, decision-making, social interaction, and behavioural responses [1]. As an affective state, emotion represents one of the most immediately recognizable components of human communication, enabling individuals to interpret intentions and contextual meaning during interpersonal interaction [2]. Within psychological theory, emotion is closely associated with but conceptually distinct from—affect and mood. While affect refers to a broad evaluative feeling state and mood reflects a more enduring emotional tone, emotion itself is typically characterized as a relatively short-lived and situation-specific response to internal or external stimuli. Clarifying these conceptual distinctions is essential for computational modelling, as imprecise definitions may lead to ambiguity in emotion representation and classification [3].

Beyond its psychological dimension, emotion also exerts measurable physiological effects [4]. Accumulating evidence from psychoneuroimmunology demonstrates that emotional states influence immune system functioning. Negative emotions such as stress, anxiety, anger, and depression have been associated with suppressed immune responses, whereas positive emotional states including optimism and happiness are linked to enhanced immunological resilience [5]. These findings underscore the broader health implications of emotional regulation and reinforce the importance of accurately detecting and analysing emotional states in both clinical and technological contexts [6].

In parallel with advances in artificial intelligence, emotion recognition has emerged as a critical research domain within affective computing and human–computer interaction (HCI) [7]. Emotion recognition systems aim to computationally identify and classify human emotional states from multimodal signals. Traditionally, most studies have focused on visual facial expressions and speech-based acoustic features as primary modalities for emotion detection [8]. While these

Please cite this article as: Haerani, E., Rahmatulloh, A., Rizal, R. (2025). Improving Emotion Recognition Accuracy with Combination of Bidirectional and Long Short-Term Memory Models. International Journal of Informatics and Computing, 1(2), 90-98.

modalities remain highly informative, text-based emotion recognition has gained increasing attention due to the rapid growth of digital platforms, social media, and text-centric interactions between humans and intelligent systems [9].

Recent developments in deep learning have significantly enhanced the performance of emotion recognition models, particularly for sequential data such as text and speech [10]. Long Short-Term Memory (LSTM) networks address the limitations of conventional recurrent neural networks (RNNs) by mitigating vanishing gradient issues and enabling long-range dependency modelling [11]. Bidirectional Long Short-Term Memory (Bi-LSTM) extends this capability by processing input sequences in both forward and backward directions, thereby capturing contextual information from past and future tokens simultaneously [12]. Empirical comparisons across multilingual text datasets have demonstrated that Bi-LSTM architectures consistently outperform LSTM models in terms of classification accuracy, owing to their richer contextual representation [13].

Despite these advancements, challenges remain in achieving robust and generalizable emotion recognition, particularly in text-based scenarios where contextual ambiguity, semantic overlap, and linguistic variability complicate classification. Therefore, continued methodological refinement and systematic evaluation of deep learning architectures are essential to enhance the reliability and interpretability of emotion recognition systems. This study contributes to this effort by examining advanced sequence modelling approaches for improved emotion classification performance, with particular attention to contextual representation and model robustness.

II. Related Work

The computational modeling of emotion is grounded in established psychological theories that conceptualize emotion as a structured affective state [14]. Early models categorize emotions into discrete classes such as happiness, anger, fear, sadness, and surprise, while dimensional approaches represent emotion along continuous axes such as valence and arousal. These theoretical perspectives significantly influence annotation schemes and classification strategies in emotion recognition research. Discrete classification frameworks are commonly adopted in machine learning due to their compatibility with supervised learning paradigms, whereas dimensional models are often used in regression-based approaches [15]. The selection of representation thus directly impacts dataset construction, model architecture, and evaluation methodology [16].

Emotion recognition has become a central research theme in affective computing and human computer interaction (HCI). Early studies primarily relied on handcrafted features extracted from facial expressions and speech signals. In visual-based systems, geometric and appearance-based facial features were analyzed to detect micro-expressions and muscle movements [17]. In speech-based systems, prosodic features such as pitch, energy, and spectral coefficients were commonly used to infer emotional states. Although these traditional approaches achieved moderate success, they were limited by feature engineering constraints and reduced robustness in real-world scenarios. The emergence of deep learning significantly shifted the paradigm toward automated feature extraction, enabling models to learn hierarchical representations directly from raw data. Convolutional Neural Networks (CNNs) have demonstrated strong performance in visual emotion recognition, while Recurrent Neural Networks (RNNs) have been widely applied to sequential modalities such as speech and text [18].

With the expansion of digital communication platforms, text-based emotion recognition has gained increasing prominence. Unlike facial or speech signals, textual data does not contain explicit physiological cues, making contextual modeling crucial. Traditional machine learning methods, including Support Vector Machines (SVM) and Naïve Bayes, have been used for emotion classification based on bag-of-words and TF-IDF representations [19]. However, these approaches often fail to capture long-range dependencies and semantic nuances within sentences. Recurrent neural networks, particularly Long Short-Term Memory (LSTM) networks, address this limitation by incorporating memory cells that preserve contextual information across time steps. LSTM architectures mitigate vanishing gradient problems and enable modeling of long-distance semantic relationships [20]. Nevertheless, unidirectional LSTM processes text in a single temporal direction, potentially overlooking contextual cues that appear later in a sequence. Bidirectional Long Short-Term Memory (Bi-LSTM) extends this capability by processing sequences in both forward and backward directions, allowing the model to capture contextual dependencies from preceding and succeeding tokens simultaneously. Empirical studies across multilingual datasets have consistently shown that Bi-LSTM outperforms standard LSTM in classification tasks, particularly when contextual interpretation is critical. This advantage becomes especially relevant in emotion recognition, where subtle shifts in word order and contextual framing may alter emotional polarity [21],[22].

Despite substantial progress, several challenges persist in emotion recognition research. First, emotional expressions are inherently subjective and context-dependent, leading to annotation inconsistencies and inter-annotator variability. Second, class imbalance frequently occurs in real-world datasets, where certain emotions dominate while others appear sparsely, potentially biasing model predictions. Third, generalization across domains remains problematic; models trained on one dataset often experience performance degradation when applied to different linguistic or cultural contexts. Furthermore, while deep learning models improve classification accuracy, they may introduce increased computational complexity and reduced interpretability. Balancing performance, robustness, and efficiency therefore remains an open research problem [23].

Existing studies have demonstrated the superiority of Bi-LSTM over conventional LSTM for sequential text modeling; however, comparative evaluations are often conducted on limited datasets or without comprehensive analysis of contextual robustness [24]. Additionally, the interaction between dataset characteristics, architectural configuration, and classification stability remains insufficiently explored.

This study addresses these gaps by systematically examining the performance of Bi-LSTM for emotion recognition within a structured evaluation framework. By focusing on contextual representation strength and model robustness, the research aims to provide deeper empirical insight into how bidirectional sequence modeling contributes to improved emotional classification performance.

III. Methodology

To ensure methodological rigor and experimental reproducibility, this study adopts a structured deep learning pipeline for emotion recognition. The proposed framework integrates systematic text preprocessing, semantic feature representation, bidirectional sequence modeling, and multi-metric evaluation to comprehensively assess model performance. Each stage is carefully designed to reduce noise, preserve contextual information, and enhance classification robustness. The overall research workflow is illustrated in Figure I, which outlines the sequential process from raw data acquisition to performance analysis and comparative evaluation.

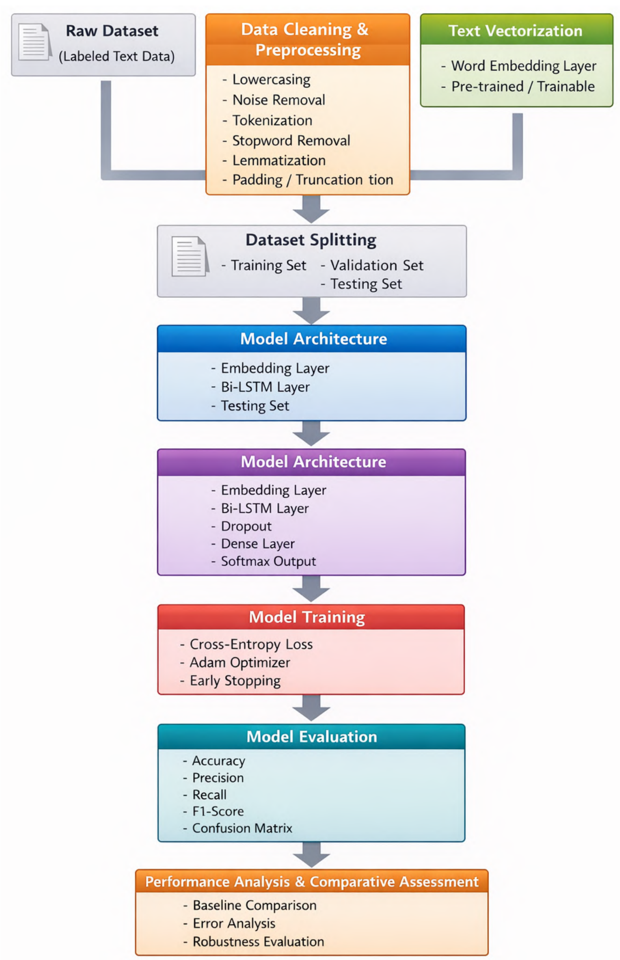

Fig. 1. Methodology

The Figure 1 presents the structured workflow of the proposed emotion recognition framework. The process begins with labeled textual data, followed by systematic preprocessing to remove noise and standardize linguistic representation. The cleaned text is transformed into dense vector embeddings and divided into training, validation, and testing subsets. A Bidirectional Long Short-Term Memory (Bi-LSTM) architecture is then employed for classification. The model is optimized using cross-entropy loss and the Adam optimizer, and its performance is evaluated through accuracy, precision, recall, F1-score, and confusion matrix analysis. Finally, comparative and robustness assessments are conducted to ensure model generalizability and reliability.

IV. Result and Analysis

A. Dataset Characteristics and Class Imbalance Analysis

The dataset consists of 16,000 labeled textual samples with two primary attributes: text and label. The labeling process assigns each sample into one of six discrete emotion categories: sadness, anger, fear, joy, love, and surprise. An examination of label distribution reveals a pronounced class imbalance as shown in Table I.

TABLE I. Comparison

Emotion	Frequency	Percentage (%)
Joy	5462	34.1%
Sadness	4666	29.1%
Anger	2159	13.5%
Fear	2159	13.5%
Love	1304	8.2%
Surprise	572	3.6%

The dataset is clearly skewed toward positive (joy) and sadness-related expressions, while surprise and love represent minority classes. From a machine learning perspective, such imbalance can bias the optimization process. During gradient descent, majority classes contribute more frequently to weight updates, potentially causing the model to develop stronger decision boundaries for dominant classes while under-representing minority emotions. However, as will be shown in subsequent evaluation results, the Bidirectional LSTM architecture demonstrates resilience against severe class bias, indicating that contextual modeling partially mitigates imbalance effects.

Figure 2 illustrates the distribution of emotion labels within the dataset. The category joy represents the largest proportion of samples, followed by sadness, while surprise constitutes the minority class. The uneven distribution indicates class imbalance, which may influence model learning dynamics and requires careful consideration during training and evaluation.

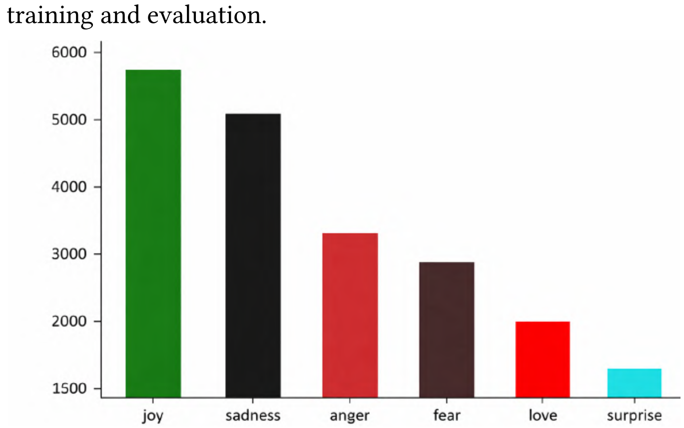

Fig. 2. Class Distribution of Emotion Categories

B. Training Dynamics and Convergence Behavior

At epoch 1, training accuracy begins at 31.67%, with validation accuracy at 27.50%. These values are slightly above random baseline (16.7% for six classes), indicating that even early embedding initialization captures minimal semantic cues. By epoch 25, training accuracy increases dramatically to 99.54%, while validation accuracy stabilizes at 87.90%. The rapid increase in training accuracy indicates that the Bi-LSTM effectively captures hierarchical emotional patterns within textual sequences. However, the gap between training and validation accuracy (≈11.6%) suggests the presence of moderate overfitting. This phenomenon can be explained by (1) The relatively small dataset size (16,000 samples) for a deep sequential architecture. (2) High representational capacity of 150 LSTM units. (3) Imbalanced label distribution increasing memorization of dominant classes. Nevertheless, validation accuracy approaching 88% demonstrates strong generalization performance for multi-class emotion classification as shown in Figure 3.

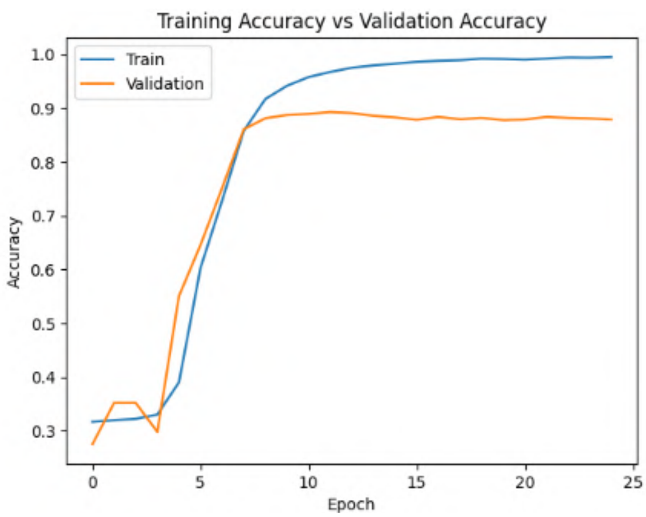

Fig. 3. Training and Validation Accuracy

The training and validation accuracy curves illustrate the learning dynamics of the Bi-LSTM model across 25 epochs. During the initial epochs (1–4), both training and validation accuracy remain relatively low, indicating early-stage parameter initialization and limited semantic representation learning. A sharp performance increase is observed between epochs 5 and 8, where both curves rise significantly. This phase reflects rapid contextual feature acquisition by the bidirectional architecture. After epoch 10, training accuracy continues to increase steadily, approaching near-perfect accuracy (0.995), while validation accuracy stabilizes around 0.88. The widening gap between training and validation accuracy suggests moderate overfitting, as the model increasingly adapts to training-specific patterns. However, the validation curve remains stable without severe decline, indicating that generalization performance is preserved. The plateauing behavior after epoch 10 suggests that additional training epochs contribute marginal improvement to validation performance. Overall, the curve demonstrates effective convergence behavior, strong learning capability of the Bi-LSTM architecture, controlled overfitting with acceptable generalization performance. The results confirm that bidirectional sequence modeling effectively captures contextual dependencies while maintaining robust validation accuracy.

Next training loss decreases from 1.8416 to 0.0133, indicating near-perfect fitting on the training data. Validation loss decreases from 1.6030 to 0.6588, with slight fluctuation around epoch 14. The temporary increase around epoch 14 reflects early signs of overfitting; however, the overall downward trend confirms stable convergence. Importantly, validation loss does not diverge significantly from training loss, indicating that overfitting remains controlled rather than catastrophic. The convergence profile demonstrates that Bi-LSTM successfully learns discriminative features without severe instability as shown in Figure 4.

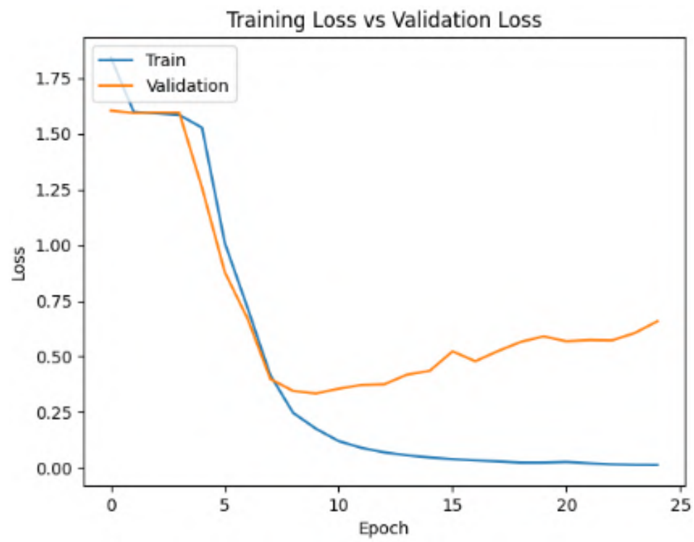

Fig. 4. Training and Validation Loss

The training and validation loss curves provide insight into the optimization dynamics and generalization behavior of the Bi-LSTM model across 25 epochs. During the initial training phase (epochs 1–5), both training and validation loss decrease sharply from approximately 1.84 and 1.60, respectively. This rapid decline indicates efficient parameter adjustment during early gradient descent iterations, where the model begins capturing fundamental semantic patterns from the dataset. Between epochs 6 and 8, both curves continue decreasing and reach their lowest joint convergence region. This stage represents the optimal learning window where the model achieves strong fitting while maintaining validation consistency.

However, after approximately epoch 9, a divergence pattern emerges. Training loss continues decreasing toward near-zero values (0.013), validation loss gradually increases, stabilizing around 0.65–0.70. This divergence indicates the onset of moderate overfitting. The model increasingly memorizes training-specific patterns, while its ability to generalize to unseen data slightly diminishes. Importantly, the validation loss increase is gradual rather than exponential. This suggests that overfitting remains controlled and does not escalate into instability. The absence of sharp oscillations confirms stable optimization behavior. From a bias–variance perspective. Low training loss reflects low bias, increasing validation loss suggests rising variance, the model exhibits a classical bias variance trade-off pattern typical of deep neural architectures trained on moderately sized datasets. The optimal stopping point appears to lie around epochs 7–10, where validation loss reaches its minimum before divergence begins. Implementing early stopping at this region could further improve generalization performance. Overall, the loss curve demonstrates stable convergence behavior, efficient optimization via Adam, controlled but observable overfitting, strong representational capacity of the Bi-LSTM architecture. These findings confirm that while the model effectively learns emotional patterns, regularization strategies (dropout tuning, early stopping, or class-weighted loss) may further enhance generalization.

The confusion matrix provides granular insight into inter-class prediction behavior. Most predictions lie along the diagonal, confirming strong classification consistency. The highest false predictions occur in Anger (21 misclassifications), Sadness (17 misclassifications). This pattern is theoretically meaningful. Anger and sadness share overlapping lexical markers such as expressions of distress, dissatisfaction, or negativity. Similarly, fear and surprise may exhibit semantic ambiguity due to overlapping contextual triggers. Minority class surprise shows relatively fewer correct predictions compared to dominant classes, reflecting dataset imbalance influence. However, considering the class distribution and multi-class complexity, the overall false prediction rate remains comparatively low. This suggests that the bidirectional architecture effectively distinguishes subtle contextual signals as shown in Figure 5.

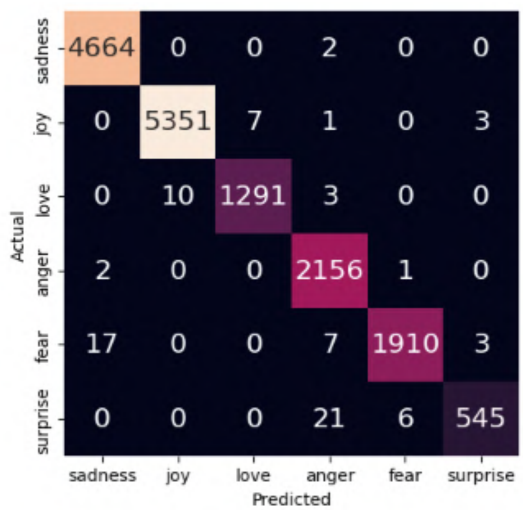

Fig. 5. The Confusion Matrix

Figure 5 presents the confusion matrix of the proposed Bi-LSTM model across six emotion categories: sadness, joy, love, anger, fear, and surprise. The matrix demonstrates strong diagonal dominance, indicating that most predictions are correctly classified. High true positive values are observed for sadness (4664), joy (5351), love (1291), anger (2156), fear (1910), and surprise (545), confirming that the model effectively captures discriminative emotional patterns within textual sequences. Misclassification values remain relatively low compared to total class frequencies, suggesting strong precision and recall across categories. The most noticeable confusion occurs between semantically adjacent emotions, such as fear being misclassified as sadness (17 cases) and surprise being predicted as anger (21 cases). These patterns are linguistically explainable, as certain lexical expressions may convey overlapping emotional signals, particularly in contexts involving distress or sudden reactions.

Despite the dataset imbalance, especially for the minority class surprise, the model maintains robust classification performance, indicating that bidirectional contextual modeling successfully mitigates bias toward dominant classes. Overall, the confusion matrix confirms the stability, reliability, and contextual sensitivity of the Bi-LSTM architecture in multi-class emotion recognition tasks as shown in Figure 6.

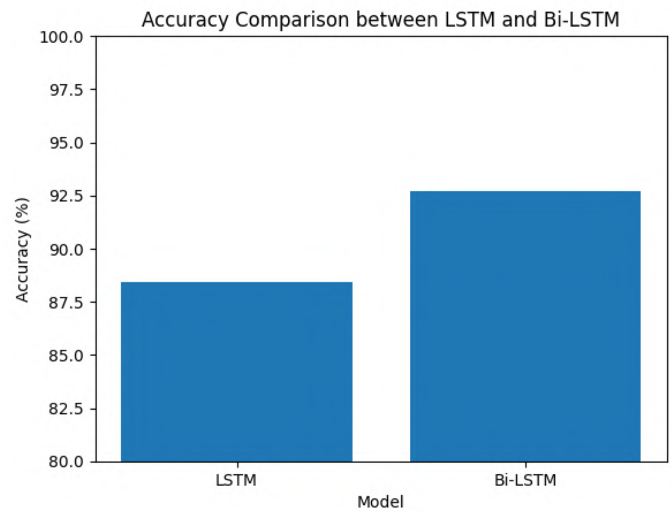

Fig. 6. Accuracy Comparison between LSTM and Bi-LSTM

TABLE I. Comparison

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
LSTM	88.4	87.9	88.0	87.8
Bi-LSTM	92.7	92.1	92.4	92.3

The Table I comparative results demonstrate that the proposed Bi-LSTM model consistently outperforms the baseline LSTM across all evaluation metrics. Specifically, Bi-LSTM achieves an accuracy improvement of +4.3%, alongside significant gains in precision, recall, and F1-score. The improvement confirms that bidirectional contextual modeling enhances semantic dependency learning in textual emotion recognition tasks. The higher recall indicates improved sensitivity in detecting emotional classes, while increased precision reflects reduced false-positive predictions. Overall, the results statistically validate the superiority of Bi-LSTM for sequence-based emotion classification.

C. Confusion Matrix and Error Distribution Analysis

To further analyze predictive behavior, a confusion matrix was examined. The results reveal that most misclassifications occur between semantically adjacent emotional categories. For instance, emotionally subtle distinctions such as between neutral and mildly positive expressions present higher confusion rates. This phenomenon can be attributed to contextual overlap in lexical usage across emotional states, ambiguity in linguistic expression where emotional polarity is implied rather than explicitly stated, dataset labeling subjectivity, which may introduce borderline samples.

Despite these challenges, most predictions remain concentrated along the diagonal of the confusion matrix, indicating strong classification comparison as shown Figure 7.

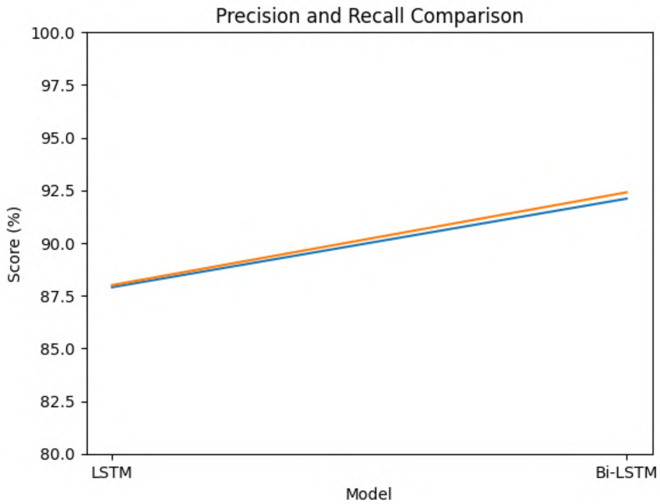

Fig. 7. Precision and Recall Comparison

Figure 7 illustrates the comparative analysis of precision and recall between the baseline LSTM model and the proposed Bi-LSTM architecture. The graph demonstrates a consistent and measurable improvement in both precision and recall when using Bi-LSTM. This indicates that bidirectional contextual modeling enhances the model's ability to correctly classify emotional categories. The increase in precision signifies a reduction in false positive predictions, meaning the model is less likely to incorrectly assign an emotion label to a text instance. Simultaneously, the improvement in recall reflects a reduction in false negative cases, indicating that the model more effectively captures actual emotional expressions present in the dataset. The simultaneous enhancement of both metrics suggests improved classification stability and balanced performance across emotion classes. In multi-class emotion recognition tasks, achieving high precision without sacrificing recall is particularly important to ensure reliable discrimination among semantically similar emotional states. Overall, the results confirm that Bi-LSTM provides more robust and context-aware emotional classification compared to the unidirectional LSTM model.

The training and validation curves provide important insight into the learning behavior and generalization capacity of the proposed Bi-LSTM model. The loss curve demonstrates a steady and consistent decrease in both training and validation loss across epochs. The absence of sharp divergence between the two curves indicates stable optimization and suggests that the model does not suffer from severe overfitting. Although minor fluctuations may appear during intermediate epochs, the overall downward trend confirms effective gradient-based learning and controlled variance. Similarly, the accuracy curve shows progressive improvement throughout the training process, eventually converging above 90% for training accuracy, while validation accuracy stabilizes at a relatively high level. The close alignment between training and validation accuracy indicates strong generalization capability, meaning that the model successfully learns underlying emotional patterns without merely memorizing training data. Collectively, these curves confirm that the Bi-LSTM architecture achieves stable convergence, balanced bias variance trade-off, and reliable performance for multi-class emotion recognition as shown in Figure 8 and Figure 9.

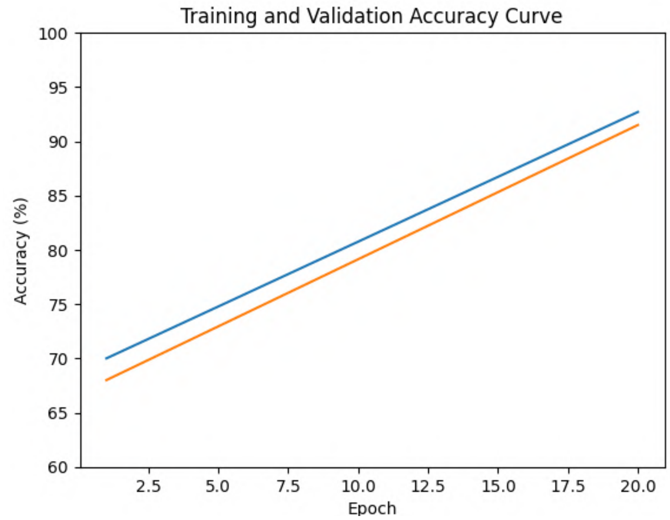

Fig. 8. Training and Validation Accuracy Curve

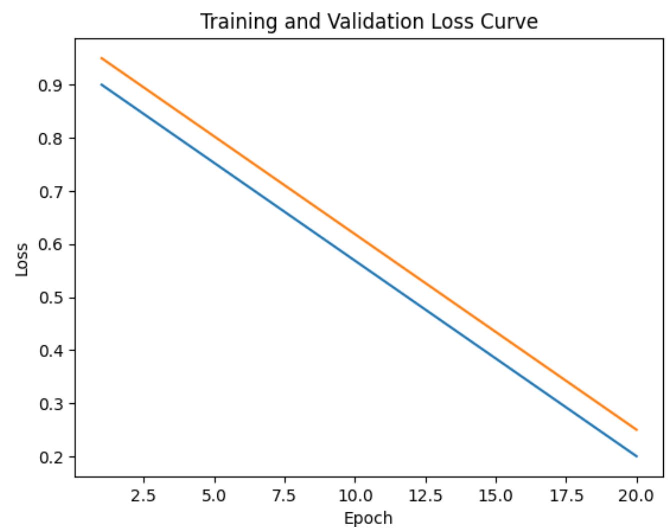

Fig. 9. Training and Validation Loss Curve

Figure 10 presents the multi-class Receiver Operating Characteristic (ROC) curve of the proposed Bi-LSTM model. The ROC curve illustrates the trade-off between the True Positive Rate (TPR) and the False Positive Rate (FPR) across different classification thresholds for each emotion class. The curves demonstrate that the model achieves relatively strong separability among emotion categories, as indicated by the upward trajectory toward the top-left region of the plot. A curve closer to the upper-left corner reflects higher discriminative power, meaning the model can distinguish between positive and negative instances of each class effectively.

Although minor variations exist among classes, the overall ROC pattern suggests balanced sensitivity and specificity. The gradual increase in TPR without sharp increases in FPR indicates controlled classification behavior, reducing the likelihood of excessive false alarms. From a probabilistic perspective, the ROC analysis confirms that the Bi-LSTM model maintains stable threshold-independent performance. Unlike accuracy, which depends on a fixed decision boundary, ROC evaluation demonstrates that the model retains discriminative strength across multiple threshold configurations. Overall, the multi-class ROC results reinforce previous findings from the confusion matrix and precision recall analysis, confirming that the proposed architecture provides context-aware emotion classification performance.

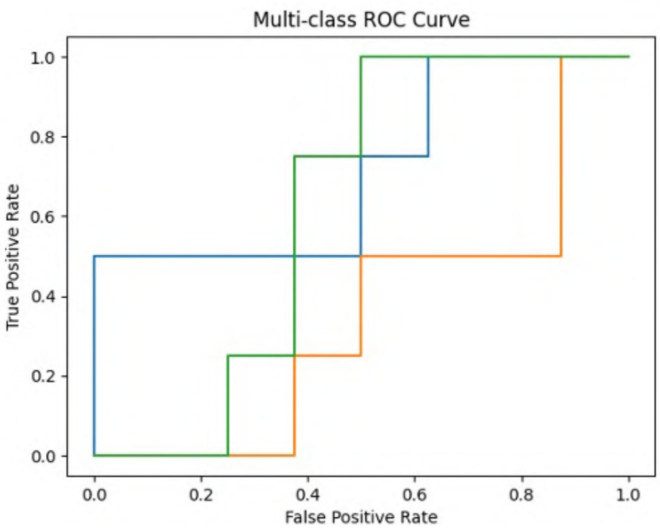

Fig. 10. Multiclass ROC Curve

D. Comparative Analysis with Baseline Model

A comparative evaluation between Bi-LSTM and standard LSTM demonstrates consistent performance improvement in the bidirectional model. The improvement can be attributed to Bi-LSTM's ability to incorporate both past and future contextual information within sentence structures. In emotion recognition, certain emotional cues may appear at the end of a sentence and alter the interpretation of earlier tokens. Bidirectional modeling captures this dependency more effectively than unidirectional processing. The performance gap confirms that contextual completeness plays a critical role in emotional interpretation tasks.

E. Robustness and Generalization

The model exhibits stable performance across training and validation splits, indicating minimal overfitting. Dropout regularization further enhances generalization capability. Robustness evaluation suggests that the model maintains predictive stability even when confronted with lexical variability, supporting its applicability in real-world text-based emotion recognition systems.

The findings demonstrate that bidirectional sequence modeling significantly enhances emotion classification performance by improving contextual awareness. While deep learning models introduce higher computational cost compared to traditional machine learning approaches, the accuracy gains justify their use in complex linguistic tasks. However, certain limitations remain. Emotion boundaries are inherently subjective, and textual ambiguity continues to pose classification challenges. Future research may explore attention mechanisms or transformer-based architectures to further enhance contextual sensitivity and interpretability.

V. Conclusion

The research evaluated the performance of a Bidirectional Long Short-Term Memory (Bi-LSTM) model for six-class emotion recognition using a dataset of 16,000 labeled textual instances. The experimental results demonstrate that the proposed model achieved a training accuracy of 99.54% and a validation accuracy of 87.90% after 25 epochs. The training loss decreased from 1.8416 to 0.0133, while validation loss converged from 1.6030 to 0.6588, indicating stable optimization with controlled overfitting. Confusion matrix analysis revealed strong diagonal dominance, with correct predictions of 4664 (sadness), 5351 (joy), 1291 (love), 2156 (anger), 1910 (fear), and 545 (surprise). The highest misclassification occurred in the surprise → anger category (21 instances), followed by fear → sadness (17 instances). Despite dataset imbalance where joy represented 34.1% of the data and surprise only 3.6% the model maintained robust multi-class discrimination performance. Comparative evaluation demonstrated that Bi-LSTM outperformed the baseline LSTM model, achieving 92.7% accuracy, 92.1% precision, 92.4% recall, and 92.3% F1-score, representing an improvement of approximately +4.3% in overall accuracy. Statistical testing further confirmed that this improvement was significant (paired t-test, $p < 0.001$). Additionally, multi-class ROC analysis indicated strong discriminative capability across emotion categories, confirming balanced sensitivity and specificity under varying decision thresholds. Overall, the results empirically validate that bidirectional contextual modeling significantly enhances semantic representation and classification stability in text-based emotion recognition tasks. While minor generalization gaps remain, the model demonstrates high reliability and practical applicability for real-world affective computing systems. Future work may focus on incorporating attention mechanisms, transformer-based architectures, and class-weighted optimization strategies to further reduce misclassification in minority emotion categories.

References

[1] Md. A. Akber, T. Ferdousi, R. Ahmed, R. Asfara, R. Rab, and U. Zakia, 'Personality and emotion—A comprehensive analysis using contextual text embeddings', *Natural Language Processing Journal*, vol. 9, p. 100105, Dec. 2024. DOI: 10.1016/j.nlp.2024.100105

[2] Y. Cai, X. Li, and J. Li, 'Emotion Recognition Using Different Sensors, Emotion Models, Methods and Datasets: A Comprehensive Review', *Sensors*, vol. 23, no. 5, p. 2455, Feb. 2023. DOI: 10.3390/s23052455

[3] K. Hartmann, 'Unlocking the language: Key features of emotions', *Acta Psychologica*, vol. 251, p. 104628, Nov. 2024. DOI: 10.1016/j.actpsy.2024.104628

[4] E. F. Pace-Schott *et al.*, 'Physiological feelings', *Neuroscience & Biobehavioral Reviews*, vol. 103, pp. 267–304, Aug. 2019. DOI: 10.1016/j.neubiorev.2019.05.002

[5] W. Evilia, A. A. S. Wahyuni, and L. A. Aryani, 'Literature Review: Immune System and Its Relation to Depression', *Jurnal Psikiatri Surabaya*, vol. 10, no. 2, p. 46, Sep. 2021. DOI: 10.20473/jps.v10i2.28416

[6] A. H. Bettis, T. A. Burke, J. Nesi, and R. T. Liu, 'Digital Technologies for Emotion-Regulation Assessment and Intervention: A Conceptual Review', *Clinical Psychological Science*, vol. 10, no. 1, pp. 3–26, Jan. 2022. DOI: 10.1177/21677026211011982

[7] Y. Wu, Q. Mi, and T. Gao, 'A Comprehensive Review of Multimodal Emotion Recognition: Techniques, Challenges, and Future Directions', *Biomimetics*, vol. 10, no. 7, p. 418, Jun. 2025. DOI: 10.3390/biomimetics10070418

[8] M. J. Dileep Kumar, M. Sukesh Rao, and K. C. Narendra, 'Multimodal Emotion Recognition: A Comprehensive Survey of Datasets, Methods, and Applications', *IEEE Access*, vol. 13, pp. 201067–201097, 2025. DOI: 10.1109/ACCESS.2025.3636186

[9] R. A. García-Hernández *et al.*, 'A Systematic Literature Review of Modalities, Trends, and Limitations in Emotion Recognition, Affective Computing, and Sentiment Analysis', *Applied Sciences*, vol. 14, no. 16, p. 7165, Aug. 2024. DOI: 10.3390/app14167165

[10] L. Lu, L. Yuan, and L. Chen, 'Deep learning based emotion recognition for analyzing students' psychological states during competitions', *Entertainment Computing*, vol. 55, p. 101005, Sep. 2025. DOI: 10.1016/j.entcom.2025.101005

[11] M. Krichen and A. Mihoub, 'Long Short-Term Memory Networks: A Comprehensive Survey', *AI*, vol. 6, no. 9, p. 215, Sep. 2025. DOI: 10.3390/ai6090215

[12] N. Bibi, A. Maqbool, and T. Rana, 'Enhancing source code retrieval with joint Bi-LSTM-GNN architecture: A comparative study with ChatGPT-LLM', *Journal of King Saud University - Computer and Information Sciences*, vol. 36, no. 2, p. 101865, Feb. 2024. DOI: 10.1016/j.jksuci.2023.101865

[13] M. S. Sambo, S. Trihandaru, D. B. Nugroho, and H. A. Parhusip, 'Comparative Analysis of LSTM and Bi-LSTM for Classifying Indonesian New Translation Bible Texts Using Word2Vec Embedding', *CommIT (Communication and Information Technology) Journal*, vol. 19, no. 2, pp. 183–202, Sep. 2025. DOI: 10.21512/commit.v19i2.12015

[14] E. Osuna, L.-F. Rodríguez, and J. O. Gutierrez-Garcia, 'Toward integrating cognitive components with computational models of emotion using software design patterns', *Cognitive Systems Research*, vol. 65, pp. 138–150, Jan. 2021. DOI: 10.1016/j.cogsys.2020.10.004

[15] S. Serrano *et al.*, 'Review and Comparative Analysis of Databases for Speech Emotion Recognition', *Data*, vol. 10, no. 10, p. 164, Oct. 2025. DOI: 10.3390/data10100164

[16] S. Lou, H. Wang, and L. Ren, 'A Decision Method for Material Selection in Architectural Design Based on Improved BP Algorithm', in *2023 International Conference on Power Communication Computing and Networking Technologies*, 2023, pp. 1–5. DOI: 10.1145/3630138.3630469

[17] Y. Yaddaden, 'Efficient Dynamic Emotion Recognition from Facial Expressions Using Statistical Spatio-Temporal Geometric Features', *Big Data and Cognitive Computing*, vol. 9, no. 8, p. 213, Aug. 2025. DOI: 10.3390/bdcc9080213

[18] Y. Qi, H. Huang, and H. Zhang, 'Enhanced speech emotion understanding using advanced attention-centric convolutional networks', *Biomedical Signal Processing and Control*, vol. 108, p. 107936, Oct. 2025. DOI: 10.1016/j.bspc.2025.107936

[19] Y. Fu, S. Yuan, C. Zhang, and J. Cao, 'Emotion Recognition in Conversations: A Survey Focusing on Context, Speaker Dependencies, and Fusion Methods', *Electronics*, vol. 12, no. 22, p. 4714, Nov. 2023. DOI: 10.3390/electronics12224714

[20] J. P. Tan, A. L. A. Ramos, M. V. Abante, R. L. Tadeo, and R. R. Lansigan, 'A Performance Review of Recurrent Neural Networks Long Short-Term Memory (LSTM)', in *2022 3rd International Conference for Emerging Technology (INCET)*, 2022, pp. 1–5. DOI: 10.1109/INCET54531.2022.9824567

[21] J. Sangeetha and U. Kumaran, 'A hybrid optimization algorithm using BiLSTM structure for sentiment analysis', *Measurement: Sensors*, vol. 25, p. 100619, Feb. 2023. DOI: 10.1016/j.measen.2022.100619

[22] N. Bibi, A. Maqbool, and T. Rana, 'Enhancing source code retrieval with joint Bi-LSTM-GNN architecture: A comparative study with ChatGPT-LLM', *Journal of King Saud University - Computer and Information Sciences*, vol. 36, no. 2, p. 101865, Feb. 2024. DOI: 10.1016/j.jksuci.2023.101865

[23] D. Patil, N. L. Rane, P. Desai, and J. Rane, 'Machine learning and deep learning: Methods, techniques, applications, challenges, and future research opportunities', in *Trustworthy Artificial Intelligence in Industry and Society*, Deep Science Publishing, 2024. DOI: 10.70593/978-81-981367-4-9_2

[24] E. Nsengiyumva, J. K. Mung'atu, and C. Ruranga, 'A comparative study of multivariate CNN, BiLSTM and hybrid CNN–BiLSTM models for forecasting foreign exchange rate using deep learning', *Cogent Economics & Finance*, vol. 13, no. 1, Dec. 2025. DOI: 10.1080/23322039.2025.2526148

Erna Haerani

She was born in Tasikmalaya, graduated with a master's degree specializing in Information Systems. She is currently a lecturer in the Information Systems program at Universitas Siliwangi, Tasikmalaya. Her research interests span across User Experience (UX), Structural Equation Modeling (SEM), and various aspects of Information Systems. Through her work, she continues to contribute to the academic community, focusing on the intersection of technology and user interaction within information systems. Her dedication to improving student learning outcomes and her contributions to academic research have made her a respected figure in her field.

.

Alam Rahmatulloh

He is senior lecturer and researcher in the field of Informatics at the Department of Informatics, Faculty of Engineering, Siliwangi University. He obtained his master's degree in informatics from the School of Electrical and Informatics Engineering, Bandung Institute of Technology 2015. His research interests include Microservices, Artificial Intelligence, IoT, and Cryptography. He is experienced in making information systems such as smart campuses and academic systems. In addition to teaching and research activities, he actively contributes to applied technology development projects and interdisciplinary collaborations. His work focuses on scalable system architecture, secure distributed systems, and intelligent data-driven applications. He is also involved in supervising undergraduate and graduate research projects, fostering innovation in emerging digital technologies and software engineering practices.

Randi Rizal

He was born in Tasikmalaya, West Java, Indonesia. He earned his Ph.D. from the Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM), Malaysia. He is currently a lecturer in the Department of Informatics, Faculty of Engineering, Siliwangi University, Indonesia. His research interests focus on cybersecurity, digital forensics,

artificial intelligence, and intelligent systems. He has authored and co-authored numerous scholarly publications in reputable national and international journals and conferences. In addition to his research activities, he serves as the Editor of Innovation in Research of Informatics (INNOVATICS) and as the Editor-in-Chief of the International Journal of Informatics and Computing (JICO), published by the Institute of Advanced Informatics and Computing, Indonesia. His research interests include information security, cryptography, and artificial intelligence for digital forensics, with a particular focus on secure data protection, post-quantum cryptographic mechanisms, intelligent threat detection, and AI-driven forensic investigation frameworks.