An Intelligent System to Identify the Emotions in the Text Using a Hybrid Deep Learning Model

Chevella Anil Kumar[1][a], Vinay Kumar Chikoti[2][b], Sree Ram Gandla[3][c], Akshitha Ganji[1], M. Dharma Teja[1] and B. Nikhilesh[1]

[1]Department of Electronics and Communication Engineering, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, Telangana, India

[2]Software Engineer, Softech International Resources Inc, North Carolina, U.S.A.

[3]E-Giants Technologies LLC, Texas, U.S.A.

Keywords: Natural Language Processing (NLP), Bidirectional Encoder Representations from Transformers (BERT), Hybrid Models, LSTM (Long Short-Term Memory), GRU (Gated Recurrent Unit), Contextual Embeddings, Attention Mechanism, Feature Extraction.

Abstract: This research introduces a novel deep learning approach to effectively identifying emotions in text. This model is a hybrid structure that integrates the advantages of BERT, LSTM, GRU, and Transformer encoder layers to detect nuanced emotional signals and intricate linguistic patterns. By incorporating attention mechanisms, we enhance the model's ability to focus on significant details and comprehend context. We trained our model using the dataset 'emotion_dataset.csv' to accurately classify emotions across different text formats. This approach has a wide range of applications, including sentiment analysis, complex storytelling, human-computer interaction, personalized content generation, and monitoring mental health.

1 INTRODUCTION

As digital communication grows rapidly, there is a wealth of unstructured textual data from sources such as social media, online chats, emails, and consumer reviews. This data contains significant emotional insights that are critical for applications such as human-computer interaction, mental health support, sentiment analysis (Salam, Gupta, et al. 2018), and tailored content production. By identifying emotions in text, AI systems across a range of fields can become far more perceptive and sympathetic. However, conventional emotion detection systems, which are usually based on rule-driven techniques or simple machine learning algorithms, struggle with the complex and context-sensitive nature of human emotions.

By capturing the complex linguistic patterns and relationships within a sentence, deep learning (Souza, Souza, et al. 2019)] models—such as, Bidirectional Encoder Representations from Transformers (BERT)—have demonstrated great promise in natural language processing (NLP) applications in recent years. Due to its bidirectional architecture, BERT can understand the underlying context from both forward and backward, which makes it useful for tasks requiring a thorough comprehension of context. Although BERT is widely used in many NLP tasks, its ability to decipher text's complex and multi-layered emotions is somewhat constrained by its inability to handle long-term dependencies and subtle emotional cues

In order to overcome these challenges, our study introduces a hybrid model that combines Transformer encoder layers with BERT and Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) architectures. The suggested approach seeks to increase the model's sensitivity to emotional nuances by utilizing the contextual benefits of BERT, the sequential memory capacities of LSTM and GRU,

[a] https://orcid.org/0000-0003-2277-8748
[b] https://orcid.org/0009-0006-3041-6519
[c] https://orcid.org/0009-0009-5495-2246

Anil Kumar, C., Chikoti, V. K., Gandla, S. R., Ganji, A., Dharma Teja, M. and Nikhilesh, B.
An Intelligent System to Identify the Emotions in the Text Using a Hybrid Deep Learning Model.
DOI: 10.5220/0013590700004664

In *Proceedings of the 3rd International Conference on Futuristic Technology (INCOFT 2025) - Volume 2*, pages 277-283
ISBN: 978-989-758-763-4

and the focused attention qualities of Transformer layers. By taking into account both the immediate context and the more complex, deeper linkages seen in emotionally charged language, our suggested approach seeks to improve emotion recognition. The suggested deep learning (Rashid, Iqbal, et al. 2020) approach has broad and important potential ramifications, with improvements anticipated in domains including automated customer support, personalized content distribution, and mental health support systems.

2 LITERATURE SURVEY

We strongly encourage authors to use this document Emotion recognition in text has emerged as a critical area of research, driven by its potential applications in fields such as human-computer interaction, sentiment analysis, and mental health assessment. Recent studies have introduced innovative models that leverage deep learning and hybrid approaches to improve emotion detection accuracy. In *"Hybrid Feature Extraction for Multi-Label Emotion Classification in English Text Messages" (Ahanin, Ismail, et al. 2023)* by Zahra Ahanin, Maizatul Akmar Ismail, Narinderjit Singh Sawaran Singh, and Ammar AL-Ashmori (2022), the authors propose a hybrid feature extraction model for multi-label emotion classification. This approach combines human-engineered features, such as sentiment polarity derived from lexical resources, with deep learning-based features generated by Bi-LSTM and BERT. The model addresses the challenges of small training datasets through data augmentation, enabling it to effectively capture linguistic and contextual information. The study achieved Jaccard accuracies of 68.40% on the SemEval-2018 dataset and 53.45% on GoEmotions, demonstrating that combining handcrafted and automated features can significantly enhance performance in emotion detection tasks. Another notable contribution is AHRNN: Attention-Based Hybrid Robust Neural Network for Emotion Recognition (Xu, Liu, et al. 2022) by Ke Xu, Bin Liu, Jianhua Tao, Zhao Lv, Cunhang Fan, and Leichao Song (2022). This study introduces the Attention-Based Hybrid Robust Neural Network (AHRNN), designed to improve semantic emotion recognition and cross-language sentiment analysis. The model integrates CNNs for extracting local semantic features, Bi-LSTM for capturing contextual dependencies, and attention mechanisms for emphasizing emotionally salient words. Pre-trained embeddings are used to infuse prior semantic knowledge, and the architecture exhibits robustness against noisy data. AHRNN achieved 86% accuracy in single-language tasks, improved fine-grained classification by 9.6%, and enhanced cross-language recognition by 1.5%. These results highlight the value of attention mechanisms and hybrid architectures in addressing complex emotion recognition challenges. Building on these advancements, we propose a novel hybrid architecture that integrates BERT, LSTM, GRU, and a Transformer Encoder layer to further improve emotion detection. Unlike previous studies that combine BERT with either LSTM or attention-based models, our model leverages all four components to simultaneously capture deep contextual, sequential, and global information. BERT provides robust contextual embeddings but lacks the sequential memory necessary for tasks involving gradual emotional shifts. To address this limitation, our model incorporates LSTM and GRU layers to capture long-term dependencies and track the progression of emotional cues. The Transformer Encoder layer introduces an attention mechanism, refining feature representations and enhancing the model's focus on critical emotional signals while balancing local and global context. The proposed architecture capitalizes on the unique strengths of each component: BERT for context-rich embeddings, LSTM and GRU for sequential memory, and Transformer Encoder for attention-based refinement. Initial experimental results demonstrated a promising accuracy of 73%, indicating the potential of this hybrid design to outperform simpler model combinations. By addressing key gaps in the literature, this architecture presents a comprehensive and well-rounded approach for emotion recognition, paving the way for future advancements in NLP applications requiring nuanced emotional understanding.

3 DATASET USED

In this research, we used a dataset emotion_dataset_2.csv of 34,785 sentences, each labeled with one of eight emotion categories: *anger*, *disgust*, *fear*, *joy*, *neutral*, *sadness*, *shame*, and *surprise*. This dataset was particularly selected to enable the model to recognize a wide range of human emotions, which is valuable in tasks such as sentiment analysis (Singh, Sharma, et al. 2023), social media monitoring, and mental health assessment.

The dataset has an imbalanced distribution, with certain emotions like *joy* (11,044 sentences) and *sadness* (6,721 sentences) having a higher frequency compared to underrepresented emotions such as *shame* (145 sentences) and *disgust* (855 sentences) as shown in table 1. Though the code does not explicitly apply techniques (Malagi, Y. R et al. 2023) such as data augmentation, oversampling, or class weighting during training, this limitation is mitigated in part by using a robust architecture that combines BERT embeddings with LSTM (Ren, and She, 2021) and GRU layers (Ren, and She, 2021) followed by a Transformer encoder. This model structure, combined with careful data splitting and validation procedures, helps capture both contextual and sequential dependencies in text data, enhancing the model's ability to generalize across diverse emotions.

To prepare this, we applied label encoding to convert categorical emotion labels into numerical values, ensuring compatibility with our model's loss function. For the text data, we used tokenization with the BertTokenizer from Hugging Face, converting each sentence into BERT-compatible tokens, with sequences either truncated or padded to a fixed length of 128. This standardization facilitated efficient processing within our BERT-LSTM-GRU-Transformer model, enabling it to effectively learn patterns within the emotion categories.

Table 1: Dataset Distribution per emotion label.

Emotion	Number of Sentences
Joy	11,044
Sadness	6,721
Fear	5,409
Anger	4,297
Surprise	4,061
Neutral	2,253
Disgust	855
Shame	145
Total	34,785

The dataset was split into training and validation sets using an 80-20 ratio. This split allowed us to train the model on a majority of the data while reserving a substantial portion for performance evaluation. The pre-processed data was then converted into PyTorch datasets and subsequently loaded into Data Loaders for batch processing during training.

The model training involves tracking metrics such as accuracy, precision, recall, and F1 score. These metrics provide insight into the model's performance on each emotion category, allowing us to identify potential biases and limitations related to class imbalance.

4 METHODOLOGY

The suggested model uses a number of deep learning components to process and extract local and global textual data for the purpose of identifying emotions. The BERT model, a transformer-based model that has already been trained and offers contextual embeddings for every token in the input text, is the first component of the architecture. After that, these embeddings are sent to an LSTM (TGDK, Selvarai, et al. 2023), (Su, Wu, et al. 2018),
(Su, Wu, et al. 2018) layer, which records long-term dependencies between tokens and sequential information. A bidirectional GRU layer refines these features after the LSTM, with an emphasis on maintaining important sequential information while lowering computational cost. The Transformer Encoder layer, which uses self-attention methods to improve the model's capacity to capture global context, receives the output from the GRU.

To enable the model to concentrate on the most pertinent passages for emotion categorization, the Transformer encoder applies attention to different areas of the text based on the information that has been successively processed. This output conforms to the model's structure through dimension permutation, facilitating efficient feature extraction and preserving interoperability with subsequent layers.

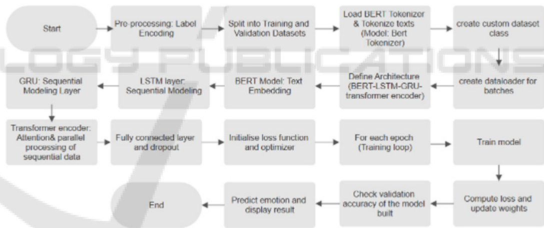

Figure 1: Workflow Diagram of the Proposed Model.

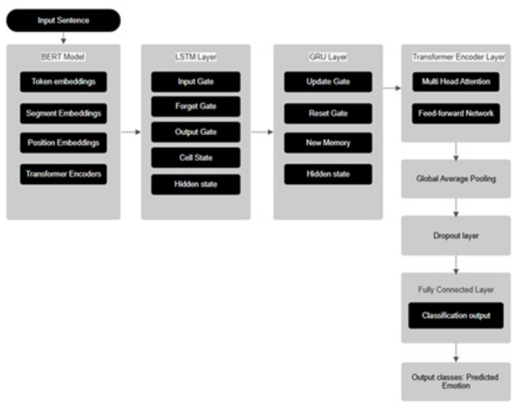

Figure 2: Hybrid Deep Learning Model for the Fine-Grained Emotion Recognition

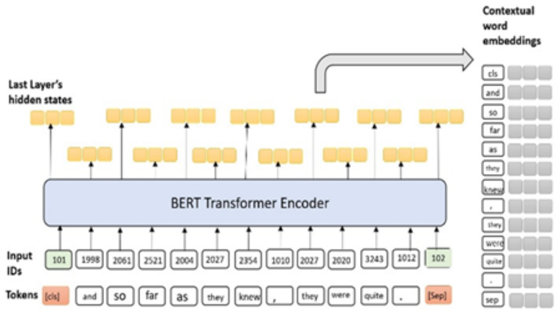

Figure 3: BERT Transformer Encoder Architecture

The architecture starts with the BERT model, which accepts tokenized text as its input. Bert model produces contextual embeddings that capture the meaning of each token in context. The resulting output shape from BERT is (number of tokens, 768), indicating the contextual representation of each token as shown in Figure 3.

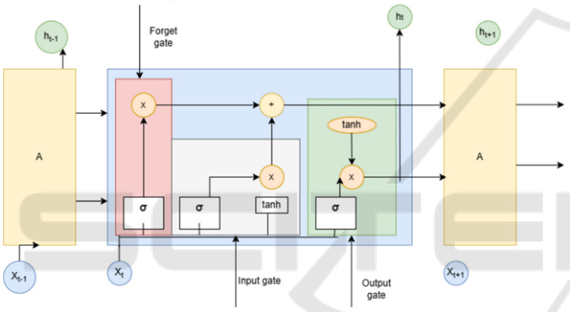

Figure 4: LSTM cell with Input, Forget, Output, and Cell Gates.

The output generated by BERT is then fed into an LSTM (Long Short-Term Memory) layer as shown in Figure 4. This layer captures the temporal dependencies among the tokens, yielding an output shape of (number of tokens, 256). It aids in identifying patterns and relationships within the sequence.

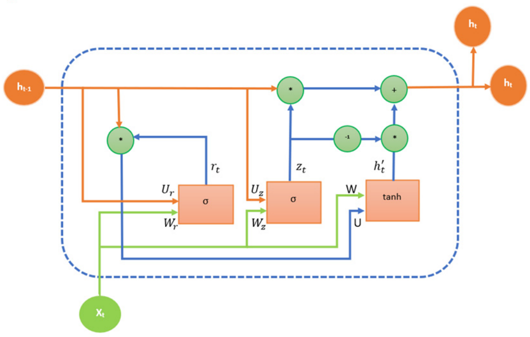

Figure 5: GRU: Capturing Long Term dependencies in Sequential Data with Reduced Complexity.

A GRU (Gated Recurrent Unit) layer processes the output after the LSTM layer. By generating an output shape of (number of tokens, 128), the GRU improves comprehension of the links between tokens. This layer enhances the model's ability to represent complex dependencies.

Subsequent to the GRU layer, the output is directed into a Transformer Encoder layer, which employs multi-head attention to enrich the contextual embeddings. The output shape remains (number of tokens, 128). The following layers include Global

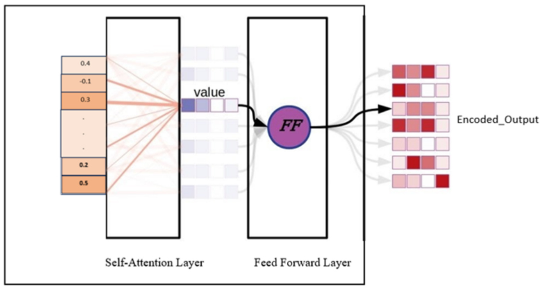

Figure 6: Transformer Encoder Layer with Self-Attention and Feed-Forward Neural Network

Average Pooling, Dropout, and a Fully Connected layer, which ultimately yields class scores with an output shape of (1, number of classes), representing the predicted sentiment classification.

Deep learning (Baruah Chutia, et al. 2024) concepts utilized in our model include Cross-Entropy Loss, which measures the difference between predicted probabilities and true labels, Adam W Optimizer, an adaptive learning rate optimization algorithm, Accuracy, Precision, Recall, and F1-Score for evaluation, Matrix Multiplication for neural network operations, ReLU Activation Function for introducing non-linearity, Tensor Operations (element-wise addition, subtraction, multiplication, and division) for data manipulation, Permutation for rearranging tensor dimensions, and Mean Pooling for reducing dimensionality. The Adam W Optimizer uses moving averages, hyperparameters, and gradients for efficient training. Precision, Recall, and F1-Score evaluate classification performance. Matrix Multiplication enables neural network layer interactions. ReLU Activation Function introduces non-linearity, allowing complex pattern learning. Tensor Operations facilitate data processing, while Permutation and Mean Pooling enable dimension rearrangement and reduction.

Let us consider an example sentence “I am so happy today!” using a BERT tokenizer. The tokenizer converts words into token IDs.

The tokenized output is:

[101, 1045, 2572, 2061, 3407, 2651, 102]

Where [101] and [102] are [cls] and [sep] tokens.

These tokens are embedded into vectors. Assume each token has an embedding vector of dimension d=768, so the sentence embedding matrix X has the shape (7,768), where each row represents a word.

The self-attention layer computes attention weights using:

$$Attention(Q,K,V) = softmax(\frac{QK^T}{\sqrt{d_k}})\,V$$

For example, if our embedding dimension is 768, we may set dk=64 per attention head. Suppose after calculating softmax on the attention weights, we get a weighted sum that represents the context for each token based on other tokens.

The LSTM takes in the last hidden state from BERT (size 7×768) and processes it through its gates at each timestep t.

For the token "happy" at the input "I am so happy today!", the hidden state is processed as:

Forget gate: $f_t = \sigma(W_f \cdot [h_{t-1,} x_t] + b_f)$, controlling how much of the previous state is retained.

Input gate: $i_t = \sigma(W_i \cdot [h_{t-1,} x_t] + b_i)$, which decides the amount of new information to add.

Cell update:

$$C_t = f_t \odot C_{t-1} + i_t \odot \tanh(W_c \cdot [h_{t-1,} x_t] + b_c)$$

These gates collectively update the hidden state for each word, capturing sequential dependencies in the text. After processing, LSTM produces output with hidden size 128 (bidirectional output size 256). Later, the LSTM output serves as input to the GRU layer, which further refines sequential representations.

For the "happy" token,

Update gate: $z_t = \sigma(Wz \cdot [h_{t-1,} x_t] + b_z)$ controls how much of the previous state is retained.

Reset gate: $r_t = \sigma(W_r \cdot [h_{t-1,} x_t] + b_r)$ controls what previous information to ignore.

The output for each token is calculated using:

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot h'_t$$

Where h'_t is the candidate hidden state at t.

For "happy," let's assume this layer outputs a vector of shape (7, 256), where 256 is the bidirectional hidden size.

The GRU [14][15] output shape of (7,256) is transposed to (7, batch size,256) for the transformer. With 4 attention heads, the Transformer applies multi-head self-attention:

$$\mathrm{MultiHead}(Q, K, V) = \mathrm{Concat}(\mathrm{head}_{1,}\ldots, \mathrm{head}_4)W^o$$

where each head computes an attention score.

For "happy," the final output vector (after all heads and FFN layers) is a contextualized representation that considers other tokens in the sentence. This helps capture complex relationships in the text.

After passing through the Transformer, the output is pooled by averaging across the sequence:

$$\text{pooled\_output} = \frac{1}{7}\sum_{t=1}^{7} transformer\_output\ t$$

Suppose this pooled vector has a dimension of 256.

Fully Connected Layer: The pooled output is then passed through a fully connected layer to produce class logits. For a 4-class problem (Happy, Angry, Joy, Sad):

$$\text{output} = \text{pooled\_output} \cdot W + b$$

where W has dimensions 256×4, resulting in logits for each emotion.

With the logits, we apply cross-entropy loss:

for the example "I am so happy today!", the predicted logits are [0.2, -0.5, 1.0, -1.5]

$$\textbf{s}\text{oftmax}([0.2, -0.5, 1.0, -1.5]) = [0.24, 0.12, 0.54, 0.10]$$

Suppose the correct label is "Happy" (label 0). The loss is:

$$\text{loss} = -\sum_{c=1}^{C} y \log(\text{softmax}(\text{output})_c)$$

$$\text{loss} = -\log(0.24) \approx 1.43$$

This setup, along with these metrics, provides a comprehensive analysis of the model's performance across the training and validation datasets.

5 RESULTS AND DISCUSSION

In this study, we evaluated the performance of four BERT-based model architectures for emotion classification: BERT+CNN, BERT+ biLSTM, BERT+GRU, and a proposed model that combines BERT with LSTM, GRU, and Transformer Encoder layers. The goal was to identify the most effective model for accurately classifying emotions based on text input, measured by validation accuracy, F1 score, precision, and recall. The BERT+CNN model showed moderate performance, achieving a validation accuracy of 0.71, with an F1 score, precision, and recall all at 0.71. This suggests that the CNN layer effectively captures local text features but may lack the ability to process deeper sequential context, which is crucial in understanding the nuances of emotional expressions. The BERT+biLSTM model, by comparison, performed slightly lower, with a validation accuracy of 0.69, an F1 score of 0.68, precision of 0.67, and recall of 0.69. The bi-directional LSTM layer within this configuration captures dependencies in both forward and backward directions, enhancing contextual understanding, but it may increase computational complexity and susceptibility to overfitting, which could account for its slightly lower performance.

Similarly, the BERT+GRU model achieved a validation accuracy of 0.69, with an F1 score of 0.69, precision of 0.71, and recall of 0.69. GRU, with a simpler architecture compared to LSTM, demonstrated efficiency but did not significantly improve the model's generalization capacity. In contrast, the proposed hybrid model combining BERT with LSTM, GRU, and Transformer Encoder layers achieved the highest performance metrics, with validation accuracy, F1 score, precision, and recall each reaching 0.73. This hybrid approach leverages the strengths of each component, with BERT providing robust contextual embeddings, LSTM and GRU capturing sequential information, and the Transformer Encoder enhancing global context understanding. This combined architecture appears to effectively balance model complexity and generalization, resulting in superior classification accuracy and robustness, as evidenced by the higher scores across all metrics.

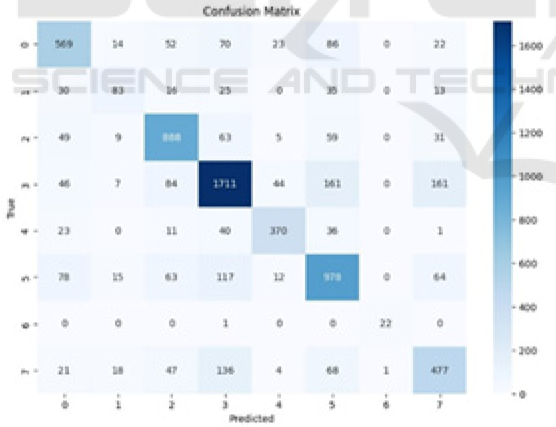

Figure 7: Confusion matrix for proposed BERT- based hybrid model for emotion classification.

The confusion matrix analysis provides further insight into the model's classification performance. The diagonal entries in the matrix indicate correctly classified instances, while off-diagonal entries represent misclassifications. Notably, the model shows strong classification accuracy for the most frequently occurring emotions in the dataset, especially for classes with substantial data representation. However, there were misclassifications between emotions with similar linguistic patterns, such as "joy" and "neutral," suggesting that overlapping language expressions in certain emotions contribute to these classification challenges. An illustrative example input, "I am feeling very excited today!" was accurately classified by the model as "joy," demonstrating its capability in real-world usage scenarios such as sentiment analysis, where accurate emotion recognition is essential.

Overall, The proposed model which is a combination of BERT+LSTM+GRU+Transformer encoder demonstrated not only high performance metrics but also balanced precision and recall, underscoring its suitability for real-world applications that require robust generalization.

Table 2: Comparative Performance Metrics of BERT-based Models for Emotion Classification as per our dataset emotion_dataset_2.csv.

Model	Validation Accuracy	F1 score	Precision	Recall
BERT+CNN	0.71	0.70	0.71	0.71
BERT+bilstm	0.69	0.68	0.67	0.69
BERT+GRU	0.69	0.69	0.71	0.69
Proposed model-BERT+LSTM+GRU +Transformer encoder	0.73	0.73	0.73	0.73

Compared to traditional machine learning classifiers, which often struggle with overfitting and generalization to unseen data, this deep learning approach provides enhanced contextual and sequential processing, making it well-suited for complex emotion recognition tasks in diverse applications such as mental health assessments, customer feedback analysis, and conversational sentiment analysis. The superior performance of this model highlights the potential of advanced hybrid architectures in accurately detecting nuanced emotions in text data. Future work could focus on expanding the dataset with a broader variety of emotional expressions and increasing the number of training epochs to further improve accuracy. Additionally, integrating other advanced components, such as attention mechanisms within the recurrent layers, may further enhance classification performance. These findings highlight the promise of BERT-based hybrid architectures in providing reliable and accurate emotion detection, crucial for applications that demand high precision in sentiment recognition from textual data

6 CONCLUSION

In terms of potential developments, this initiative could be broadened by investigating different enhancements and fine-tuning methods designed for the emotional classification model's unique needs. Integrating alternative pre-trained language models, such as RoBERTa or DeBERTa, may enhance the model's capacity for contextual comprehension, while applying ensemble techniques could bolster its robustness by merging the outputs from various architectures. Furthermore, implementing advanced data augmentation approaches could offer a wider range of linguistic contexts for training, thereby increasing the model's resilience to varied inputs. Looking into hyperparameter optimization techniques, including grid search or Bayesian optimization, might refine the training parameters, resulting in improved accuracy and efficiency.

To summarize, The Proposed model(BERT-based model, augmented with LSTM, GRU, and Transformer layers)shows remarkable potential for emotion classification within text. Its multi-faceted approach merges BERT's contextual capabilities with the sequential analysis of LSTM and GRU, alongside the potent sequence-to-sequence functions of Transformers. This integration effectively utilizes the strengths of each layer, thereby enhancing the precision of emotion detection. Future endeavours focused on adapting the model to different datasets and contexts will further boost its flexibility, making it a significant asset in sentiment analysis, mental health assessments, and various NLP applications.

REFERENCES

S. A. Salam and R. Gupta, ‘‘Emotion detection and recognition from text using machine learning,’’ Int. J. Comput. Sci. Eng., vol. 6, no. 6, pp. 341–345, Jun. 2018.

A. D. Souza and R. D. Souza, ‘‘Affective interaction based hybrid approach for emotion detection using machine learning,’’ in Proc. Int. Conf. Smart Syst. Inventive Technol. (ICSSIT), Nov. 2019, pp. 337–342.

Rashid, U., Iqbal, M.W., Skiandar, M.A., Raiz, M.Q., Naqvi, M.R. and Shahzad, S.K. (2020) Emotion Detection of Contextual Text Using Deep Learning. 2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Istanbul, Turkey, 22-24 October 2020, 1-5. https://doi.org/10.1109/ISMSIT50672.2020.9255279

Z. Ahanin, M. A. Ismail, N. S. S. Singh, and A. AL-Ashmori, "Hybrid Feature Extraction for Multi-Label Emotion Classification in English Text Messages," Sustainability, vol. 15, no. 16, p. 12539, 2023, doi: 10.3390/su151612539.

K. Xu, B. Liu, J. Tao, Z. Lv, C. Fan, and L. Song, "AHRNN: Attention-based hybrid robust neural network for emotion recognition," Cogn. Comput. Syst., vol. 4, no. 1, pp. 85–95, Mar. 2022, doi: 10.1049/ccs2.12038.

V. Singh, M. Sharma, A. Shirode, and S. Mirchandani, "Text emotion detection using machine learning algorithms," in Proc. 8th Int. Conf. Commun. Electron. Syst. (ICCES), Coimbatore, India, 2023, pp.1264-1268.

R. R. Malagi, Y. R., S. P. T. K., A. Kodipalli, T. Rao, and R. B. R., "Emotion detection from textual data using supervised machine learning models," in Proc. 5th Int. Conf. Emerg. Technol. (INCET), Belgaum, India, 2023, pp. 1–5.

F. Ren and T. She, "Utilizing external knowledge to enhance semantics in emotion detection in conversation," IEEE Access, vol. 9, 2021.

S. TGDK, V. Selvarai, U. Raj, V. P. Raiu, and J. Prakash, "Emotion detection using bi-directional LSTM with an effective text pre-processing method," in Proc. 12th Int. Conf. Comput. Commun. Netw. Technol. (ICCCNT), Kharagpur, India, 2021.

M.-H. Su, C. H. Wu, K.-Y. Huang, and Q.-B. Hong, "LSTM-based text emotion recognition using semantic and emotional word vectors," in Proc. Conf. Affect. Comput. Intell. Interact. (ACII Asia), Beijing, China, 2018.

M.-H. Su, C. H. Wu, K.-Y. Huang, and Q.-B. Hong, "LSTM-based text emotion recognition using semantic and emotional word vectors," in Proc. Conf. Affect. Comput. Intell. Interact. (ACII Asia), Beijing, China, 2018.

"Contextual emotion detection in text using deep learning and big data," in Proc. 2nd Int. Conf. Comput. Sci., Eng. Appl. (ICCSEA), Gunupur, India, 2022, pp. 1–5.

N. Baruah and T. Chutia, "An overview of deep learning methods for detecting emotions," Artif. Intell., 2024.

K. Sakthiyavathi and K. Saruladha, "Design of hybrid deep learning model for text-based emotion recognition," in Proc. 2nd Int. Conf. Adv. Comput. Intell. Commun. (ICACIC), Puducherry, India, 2023.

N. Priyadarshini and J. Aravinth, "Emotion recognition based on fusion of multimodal physiological signals using LSTM and GRU," in Proc. 3rd Int. Conf. Secure Cyber Comput. Commun. (ICSCCC), Jalandhar, India, 2023, pp. 1–6.