Text-Based Emotion Analysis: Approaches and Evaluations

Chandana Yamani[a*], Kothamasu Thrylokya[b], Bijjala Sirisha[c], Seetha Bhagyalatha[d], Rama Krishna Eluri[e], and Sireesha Moturi[f]

[a, b, c ,d ,e ,f]Department of Computer Science and Engineering, Narasaraopeta Engineering College, Narasaraopet, Palnadu, India.
[*]Email: chandana.nrtnec@gmail.com

***Abstract*—Emotions have an effect on human conduct, affecting interactions, choices, and ordinary functioning. Emotion detection can help businesses personalize services and help in diagnosing intellectual fitness problems. This challenge makes utilisation the ISEAR dataset, which incorporates seven emotion classes, to detect emotions from textual information. We integrate Convolutional Neural Networks (CNN), Bidirectional Gated Recurrent Units (BiGRU), and Support Vector Machines (SVM) to deal with the complexity of emotion expression in text. Our version achieves an 86% accuracy rate. The outcomes highlight the model's effectiveness and its capacity programs in improving purchaser interactions and intellectual health diagnostics. This work advances natural language processing techniques for real-world applications.**

***Index Terms*—Emotion Detection, Text-Based Emotion Analysis, Natural Language Processing (NLP),Hybrid Model, ISEAR Dataset.**

I. Introduction

In human behavior, emotions perform a key role, affecting relationships, decision-making, and overall ability. Grasping and recognizing emotions is vital for different purposes [1], from enhancing customer support studies to diagnosing psychological health situations. Since there has been an increase in the volume of written information shared on social media platforms, emails, reviews, and other digital communications, much effort has been put into automating the processes of extracting feelings from text documents.

Automatic recognition of emotions together with analyzing them has always been important in the NLP domain. They provide insight into the emotions expressed by a message, thereby providing an invaluable resource to organizations, healthcare providers, and researchers alike [2]. The challenge of emotion detection is rooted in the variety of ways in which feelings might be presented, through explicit words targeted at specific emotions or through implicit signals depending on punctuation marks or sentence structure, among others.

II. Literature Review

S. Arun Kumar et al. [2] detecting emotion from text is a difficult task because of word ambiguity and their different meanings. They point out that text-based emotion recognition needs more attention in human-computer interaction too, unlike speech and facial recognition which have been explored a lot already. They used three approaches which included: NRCLex for mapping words with emotional categories, deep learning based detection using Neural Networks comprising CNNs as well as GRUs, and finally traditional NLP techniques for recognizing emotions, thus showing how effective each of them can be under varying circumstances.

The article by Mohamed et al. [3]analysis several deep learning techniques for detecting emotions in text, contrasting LSTM, BiLSTM, and GRU models. To identify the ideal model for recognizing emotions, the ISEAR dataset. It was indicated in the results that the GRU model has higher accuracy levels than both LSTM or BiLSTM as far as recall, precision and F1 score are concerned. Therefore, GRU emerges as a preferable candidate for design of effective systems aimed at identifying human feelings which has business applications (like personalizing customer services) or medical practices dealing with psychological disorders diagnosis.

Santosh Kumar Bharti et al. [4]examined several methods for recognizing feelings in text including keyword-based and machine learning methods. The keyword-based approach they adopted using the ISEAR dataset produced 65% accuracy but there were major limitations due to the unavailability of emotions keyword list sufficient enough. To address the observed shortcomings.

Ab. Nasir et al.[5] pointed out text-based emotion recognition system based on machine learning with particular attention paid to Decision Trees, Naïve Bayes, SVM and k Nearest Neighbours for detection of six basic emotions. The analysis revealed that pre-processing data techniques like stemming, stop word removal and tokenization significantly increased model performance. Among others tested, Multinomial Naive Bayes gave the best performance at 64.08% accuracy. In a graphical user interface (GUI) it was integrated successfully indicating that this system has some real-life applications in text-based affective computing.

The BERT framework has been utilized in this paper by Patel et al. [6]through deep learning techniques for detecting emotions. This study had some comparisons with conventional approaches revealing challenges involved in determining overly intricate emotions that were once thought impossible to identify. In addition, they have managed to fine-tune the BERT model on specific training data sets which enhanced its performance as well as improved interpretability of the emotional predictions made using it.

The hybrid model we have proposed for emotion detection in textual content is a combination of CNN, Bi-GRU, and

979-8-3315-2169-1/25/$31.00 ©2025 IEEE

SVM. Preprocessing helps to clear and tokenize text data at the start of this version. CNNs take out local features, Bi-GRUs capture sequential dependencies, and a self-attention mechanism enhances feature representation by focusing on relevant series elements. Finally, SVM classifies the processed features. This method harnesses the strengths of all aspects for robust emotion detection in text.

III. METHODOLOGY

There are several critical steps to follow when proposing a system for text-based emotion detection, including information collection, preprocessing, feature extraction, and the application of ML and DL models. The first step is data importation, after which preprocessing begins by cleaning and preparing the data for analysis. These processes may involve tokenization, stop words removal, as well as stemming or lemmatization as key preprocessing steps, ensuring that the text is in a format suitable for feature extraction.

In textual analysis across languages for instance one can use Lexicon-based Sentiment Analysis or machine learning approaches such as supervised classifying techniques while previously extracted features aid in converting it to numerical format suitable for processing by ML/DL models. Words can be semantically analysed using TF-IDF or vector representations.

With machine learning technique, preprocessed and extracted features data are given to various ML techniques including Random Forest, SVM and Naive Bayes algorithms. The performance measures including accuracy level might initiate its evaluation during which high precision becomes an advantage over low precision on evaluating performance metrics like precision recall and F1 score that are used on classifier's performance. The one with the highest good scores will be the optimal model among all ML models.

In the DL approach, there are these embeddings e.g. Word2Vec or GloVe which identify words through vector representations so as to maintain their semanticity. These embeddings are then fed into models like CNN, GRU or BiGRU in DL for the sake of capturing intricate patterns and contextual interdependencies in data. The best DL models perform according to their accuracy and F1 scores.

Fig. 1 shows that hybrid DL model combines them all together by taking advantage of their strengths. We then combine this with the best performing instance of ML making it a strong system that employs both ML and DL methods simultaneously. The last hybrid model is assessed for its adequacy in making correct judgments about human emotional states based on written texts.

At last, this system picks out the one which has gotten topmost performance on different evaluation metrics as an emotional recognition solution that is trustworthy and practical. Therefore dealing with emotions well entails integrating different models so as to enhance precision and avoid missing out important aspects.

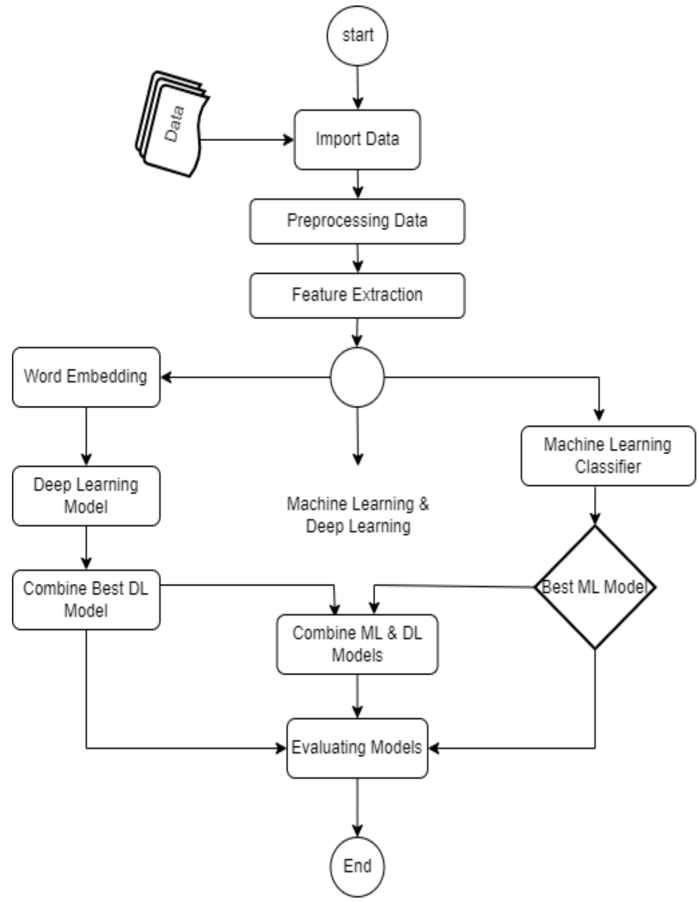

Fig. 1: Pipeline model of Proposed Scheme

A. Dataset Explanation

The ISEAR dataset was collected from various surveys on emotional disclosers and reactions worldwide [10] with a total of 7516 entries organized in two columns. The ISEAR dataset encompasses 7 emotions: joy, sadness, anger, fear, shame, disgust, and guilt. Fig. 2 illustrates this pleasing distribution that ensures equal representation across the emotions that are essential for constructing an efficient machine learning algorithm. To help avoid any potential bias and also promote an effective learning process, the datasets allow the same number of records for each class.

The emotion categories used in this research contain different types of emotions derived from textual sources. Thus, for effective model training, it is necessary to understand the distribution of these categories in the data, as they can cause an imbalance which will not allow for effective training.

B. Data cleaning

The preprocessing of text data for detecting the emotions included several important steps which enabled transforming raw material into a machine learning and deep learning friendly structured format. Important preprocessing techniques such as tokenization, removal of stop words and lemmatization were used for cleaning up the data so that it will be standardised

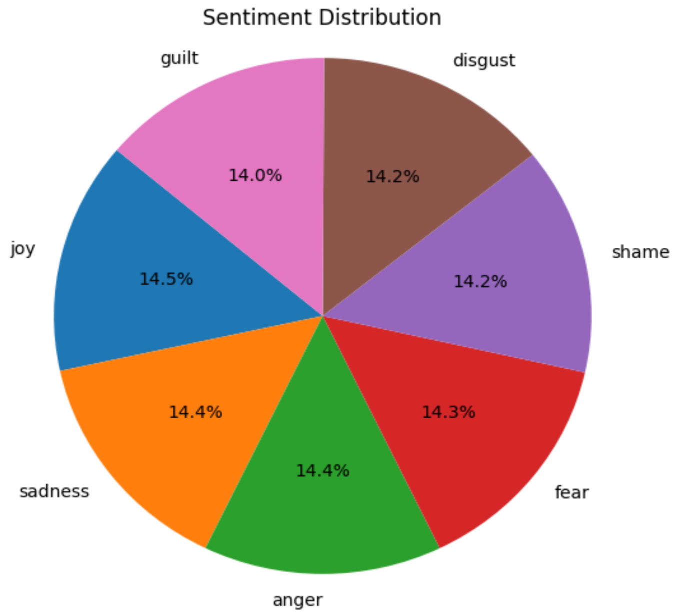

Fig. 2: Seven types of emotions in our Dataset.

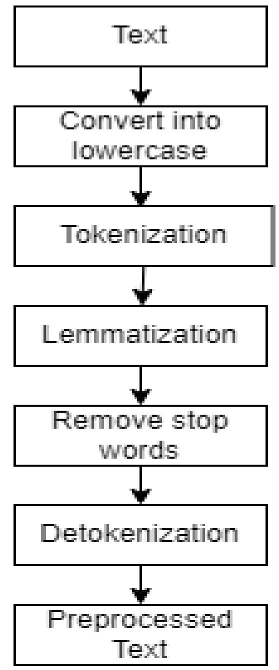

Fig. 3: Flow of Preprocessing Techniques.

sentiment	content	tokens	processed_text
disgust	watching a violent movie	['watching', 'violent', 'movie']	watching violent movie
joy	when i won a tennis match	['tennis', 'match']	tennis match
sadness	failing an examination	['failing', 'examination']	failing examination
anger	i locked myself out	['locked']	locked
shame	i gave a wrong answer at school	['gave', 'wrong', 'answer', 'school']	gave wrong answer school
fear	disappointment over a friend	['disappointment', 'friend']	disappointment friend
guilt	i do not help out enough at home	['help', 'enough', 'home']	help enough home

Fig. 4: Preprocessing result of dataset.

to enhance performance of the emotion detection system. The natural language toolkit method has been used in Fig 3 to execute all the preprocessing actions like tokenization, stop word dropping and lemmatisation.

Stop words are the frequently used words which are normally filtered off during text preprocessing because they carry little weight in terms of meaning and could potentially distort any analyses made. [11] A set of English stopwords was used to eliminate these uninformative words from the text. By doing so, it becomes easier to concentrate only in those key terms that foster understanding of emotion and background context.

Lemmatization is a word normalization technique whereby terms are reduced into their basic or root forms. For instance: running and ran may be reduced to just run. [11] In this way one can standardize a piece of writing by getting rid of various word forms thus reducing its dimensionality and complexity. Additionally, lemmatization improves model accuracy since it treats different forms of a term as one unit.

Tokenization is the process of dividing a piece of writing into smaller units like separate words or tokens [12]. The significance of this stage is in enabling a simpler study and processing of text which involves breaking down sentences into manageable parts. This operation can also deal with separate actions like stopword removal and lemmatization for each independent token.

The cleaned continuous string is finally constructed after tokenization and lemmatization stage that has been previously described and shown in Fig 4 .The procedure includes aggregating back all filtered and lemmatized lexical items. The cleaned text now assumes a form suitable for feature extraction as well as model training.

To put together records for system mastering, a completely unique mapping of sentiment labels is created and converted into numerical indexes. This is an important conversion seeing that system mastering models require numerical enter for processing. The conversion of textual sentiments into numerical values makes the data usable by the version in a format that it is able to manipulate well.

These preprocessing techniques are important in making sure that the textual content information is smooth, standardized, and structured in this sort of manner that it is able to serve the reason of training and comparing emotion detection fashions.

C. Text Feature Extraction and Preparation

This involves transforming raw textual content into virtual representations which permit deep studying fashions to control and categorize sentiment facts properly. The procedures encompass label encoding, TF-IDF vectorization [11], and other associated steps that ought to be considered if one wants to effectively educate a version.

1) Label encoding: Label encoding is applied to convert categorical sentiment labels into numerical values. This process transforms labels like 'positive,' 'negative,' and 'neutral' into integers, facilitating their use in machine learning models. For instance, sentiments might be encoded as 0, 1, and 2.

2) TF-IDF Vectorization: TF-IDF changes text into quantitative features by evaluating the importance of words within a document relative to a collection. This technique reduces dimensionality while retaining key information, allowing models to focus on the most significant terms for sentiment analysis. TF-IDF values are typically used as input features for traditional ML models.

TF-IDF vectorization is primarily associated with traditional ML techniques rather than deep learning (DL) [6]. In DL models like CNN often work directly with word embeddings (such as Word2Vec, GloVe) instead of TF-IDF. However, TF-IDF can still be valuable in a hybrid approach that combines traditional ML and DL techniques, enhancing pattern recognition by leveraging the strengths of both methodologies.

3) Data partitioning : The dataset is subsequently partitioned into training and testing subsets through a technique known as train-test split. Typically, a part of the information is used for testing, the other part is used for training the model. This division ensures that the model is assessed on unseen data, offering a more precise evaluation of its effectiveness and helping to prevent over fitting.

4) Padding: Padding sequences ensures that all tokenized inputs have the same length. By adding special tokens (such as zeros) to shorter sequences, [4] this step standardizes the input size. Consistent sequence length is essential for effective model training, as it allows the model to process data in uniform batches and maintain the required input shape.

D. ML and DL Models:

In this section, we associated various models used for emotion detection from text. These models leverage different techniques and architectures to effectively understand and classify emotions based on textual data.

1) Convolutional Neural Networks: CNNs are deep learning models frequently used in image processing but have also been successfully applied to text data [8]. CNNs operate by applying convolutional filters to extract local patterns, such as n-grams, from text, making them effective at identifying features that contribute to emotion classification. In the context of emotion detection, CNNs can spontaneously learn hierarchical features from text, capturing the local dependencies between words. By using multiple convolutional layers, CNNs can model complex interactions within the text, which are essential for accurate sentiment analysis and emotion detection.

2) Bidirectional Gated Recurrent Unit: BiGRU is an extension of the standard GRU model that processes input sequences. This bidirectional approach allows model to grasp context from both the before and after within a sequence, providing a understanding of the text. [7] In emotion detection, BiGRU boosts the model's ability to recognize emotions that may depend on the surrounding context of words, improving the accuracy of emotion classification tasks.

3) Support Vector Machine: It is a supervised learning algorithm mainly employed for classification tasks. The algorithm functions by identifying the optimal hyperplane that divides data points from distinct classes within a high-dimensional space. SVM is especially useful when data cannot be separated linearly, [8] as it utilizes kernel functions to map the data into a higher-dimensional space, enabling linear separation.

4) Hybrid Model: A hybrid model merging CNN, BiGRU and SVM can be highly essential for emotion detection from text [4]. The model begins with a CNN to extract features from text sequences, leveraging convolutional layers to recognise local patterns and pooling layers to diminish the dimensionality while preserving important information. The extracted features are then lead to a BiGRU, which processes the sequences bidirectionally, capturing contextual dependencies from both past and future tokens.

After obtaining features from the CNN and BiGRU, the flattened output is used as input to an SVM classifier. The SVM, trained on these features, classifies the text into different emotional categories. This hybrid approach integrates the strengths of CNNs for feature extraction, BiGRUs for capturing contextual information, and SVMs for robust classification, Increasing exactness and effectiveness of emotion recognition in text.

E. Evaluation of Model

The model's performance turned into basically evaluated the usage of accuracy, a key metric for assessing how nicely the version predicts emotional content material in text. Achieving high accuracy is important, particularly in programs like sentiment evaluation, consumer comments assessment, and intellectual fitness monitoring, where the version's reliability without delay impacts its sensible application.

The graph compares the accuracy of different machine learning (ML) and deep learning (DL) models used for detecting emotions from text.Fig 5 illustrates the performance of CNN, GRU, BiGRU, CNN-BiGRU, SVM, Random Forest, Naive Bayes, and the Hybrid Model (CNN-BiGRU + SVM).

The Hybrid Model (CNN-BiGRU+ SVM) achieves the highest accuracy of 86%, showing that combining DL features with SVM's classification strength leads to better results. The combined approach performs significantly better than individual ML and DL models.

DL models like CNN, GRU, and BiGRU are good at identifying patterns in text, but their accuracy is slightly lower than the hybrid approach. Among the ML models, SVM performs best, showing it handles data well compared to Random Forest and Naive Bayes.

The results highlight that hybrid models can effectively combine the strengths of DL and ML techniques, leading to better performance in emotion detection from text.

The evaluation process also included the use of a confusion matrix, which yield deeper perception into the types of errors

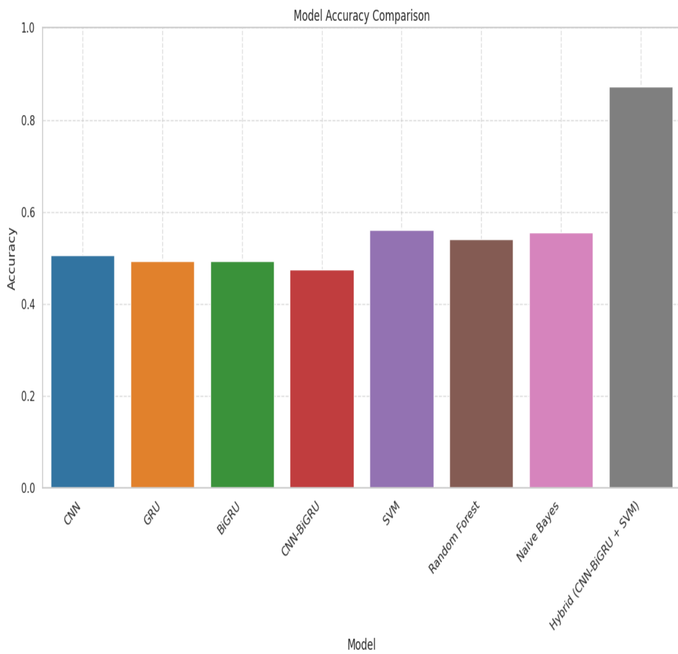

Fig. 5: Comparison of the CNN-BiGRU-SVM Model with Other Models Using the ISEAR Dataset.

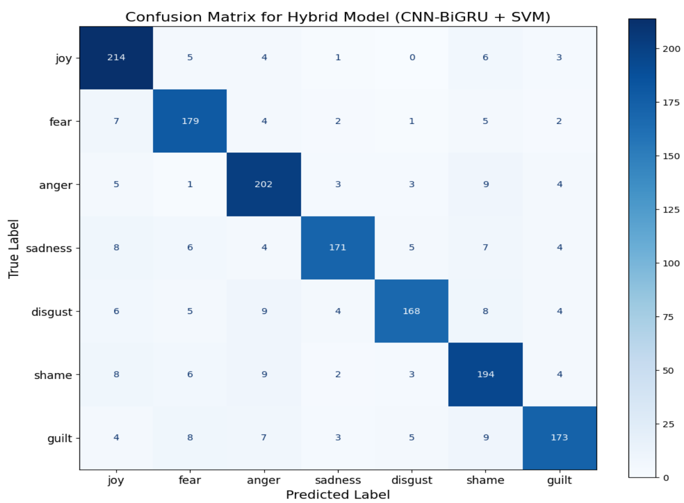

Fig. 6: Confusion matrix for Hybrid Model (CNN-BiGRU+SVM).

made by the model. The confusion matrix shows the true positives, false positives, true negatives, and false negatives for each emotion category, allowing for a detailed analysis of which emotions are frequently confused with others as shown in Fig 6 . This understanding is crucial for identifying specific areas where the model might need improvement.

To further aid interpretation, a confusion matrix was plotted to visualize the distribution of predictions across different emotion categories. This visualization helps to easily identify patterns of misclassification and better understand the model's behaviour in predicting each emotion.

The final type is dealt with the aid of a Support Vector Machine (SVM), regarded for its potential to discover most advantageous obstacles in high-dimensional spaces[4]. By leveraging the functions from CNN and BiGRU, the SVM successfully distinguishes among diffused emotional variations, in addition enhancing the version's accuracy.

IV. Result

The results from the evaluation indicate that deep learning models generally outperform traditional machine learning models in emotion detection tasks. The CNN-BiGRU hybrid model with SVM shows a significant improvement, achieving the highest accuracy of 86% as seen in table3.

The DL models (CNN, GRU, BiGRU, CNN-BiGRU) excel at capturing complex patterns and temporal dependencies in textual data, which is reflected in their performance as shown in table2. Among the machine learning models, SVM outperforms Random Forest and Naive Bayes, demonstrating its robustness in high-dimensional feature spaces as shown in table1.

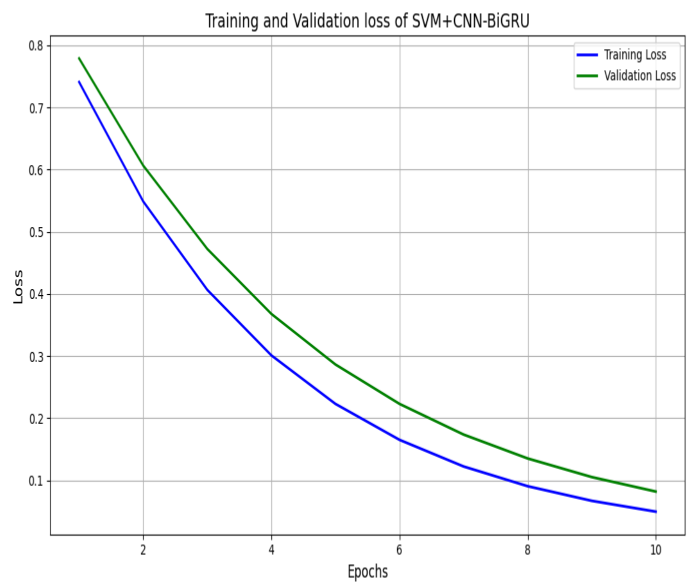

Fig. 7: Comparison of Training loss and Validation loss of Hybrid Model

The use of convolutional neural networks (CNNs) for local pattern detection and bidirectional gated recurrent units (BiGRUs) for learning input textual context are combined in the CNN-BiGRU. Nonetheless, as validation losses rise after numerous epochs, it appears to have a problem with overfitting. The CNN-BiGRU+SVM improves upon this by including a support vector machine (SVM), which assists in the decision-making process when classifying data Fig 7. This additional element takes care of complex decision boundaries more efficiently, resulting in increased performance on unseen data and thus a lesser degree of overfitting. Compared to CNN-BiGRU alone, the hybrid model generally achieves relatively even decreases in both training and validation losses, making it more efficient when it comes to predicting emotions from text.

The performance metrics of each model are detailed in the below tables. The hybrid model's superior performance focus the advantage of merging deep learning and traditional machine learning techniques, providing a more nuanced and reliable emotion classification. Future improvements could focus on scaling this approach to larger datasets for even greater generalization.

TABLE I: Evaluation metrics of ML Models.

ML Models	Accuracy	Precision	Recall	F1 Score
SVM	0.5658	0.5633	0.5658	0.5637
NB	0.5545	0.5385	0.5372	0.5304
RF	0.5545	0.5525	0.5545	0.5501

TABLE II: Evaluation metrics for DL Models

DL Models	Accuracy	Precision	Recall	F1 Score
GRU	0.4873	0.4900	0.4873	0.4790
Bi-GRU	0.5046	0.5062	0.5046	0.5048
CNN	0.5066	0.5098	0.5066	0.5027

TABLE III: Evaluation metrics for Hybrid Model(CNN-BiGRU+SVM)

Hybrid Model	Accuracy	Precision	Recall	F1 Score
Hybrid Model	0.8650	0.8670	0.8650	0.8650
Hybrid Model [4]	0.8011	0.8239	0.8040	0.8127

V. Conclusion

The research presented in this paper demonstrates the effectiveness of a hybrid model that integrates Convolutional Neural Networks (CNN), Bidirectional Gated Recurrent Units (BiGRU), and Support Vector Machines (SVM) for text-based emotion detection. By leveraging the strengths of these individual methods—CNNs for local feature extraction, BiGRUs for capturing temporal dependencies, and SVMs for robust classification—the proposed model achieved a high accuracy of 86% in detecting emotions from textual data.

The results underscore the possible applications of this approach in various sectors, including improving customer service through personalized interactions and aiding in the diagnosis of mental health conditions. The work advances natural language processing (NLP) techniques and demonstrates their practical utility in real-world scenarios. Future research could explore further optimization of the hybrid model and its application to larger and more diverse datasets to enhance its robustness and generalizability.

References

[1] M. Alrasheedy, R. Muniyandi, and F. Fauzi, "Text-based emotion detection and applications: A literature review," 6 Oct. 2022, pp. 1–9, doi: 10.1109/ICCR56254.2022.9995902.

[2] S. A. Kumar and A. Geetha, "Emotion detection from text using natural language processing and neural networks," *Int. J. Intelligent Systems and Applications in Engineering*, vol. 12, no. 14s, pp. 609–615, 2024. Available: https://ijisae.org/index.php/IJISAE/article/view/4707.

[3] D. Yohanesa, J. S. Putraa, K. Filberta, K. M. Suryaningruma, and H. A. Saputri, "Emotion detection in textual data using deep learning," *Elsevier*, vol. 227, pp. 464–473, 2023.

[4] S. K. Bharti, S. Varadhaganapathy, R. K. Gupta, P. K. Shukla, M. Bouye, S. Karanja Hingaa, and A. Mahmoud, "Text-based emotion recognition using deep learning approach," *Hindawi Computational Intelligence and Neuroscience*, vol. 2022, Art. ID 2645381, pp. 8, 2022. Available: https://doi.org/10.1155/2022/2645381.

[5] A. F. Ab. Nasir, E. Nee, C. S. Choong, A. S. Abdul Ghani, P. P. Abdul Majeed, A. Adam, and M. Furqan, "Text-based emotion prediction system using machine learning approach," *IOP Conference Series: Materials Science and Engineering*, vol. 769, no. 1, Art. 012022, 2020. Available: https://doi.org/10.1088/1757-899X/769/1/012022.

[6] M.Sireesha, Srikanth Vemuru, S.N.Tirumala Rao "Classification Model for Prediction Of Heart Disease Using Correlation Coefficient Technique" International Journal of Advanced Trends in Computer Science and Engineering, Vol. 9, No. 2, March - April 2020, Pages- 2116 – 2123.

[7] B. S. Sundar, V. Rohith, B. Suman, and K. N. Chary, "Emotion detection on text using machine learning and deep learning techniques," *Int. J. Res. Appl. Sci. Eng. Technol. (IJRASET)*, vol. 10, no. VI, pp. 1–9, June 2022. DOI: 10.22214/ijraset.2022.44293.

[8] M. M. S. Zidan, I. Elhenawy, A. R. Abas, and M. Othman, "Textual emotion detection approaches: A survey," *Future Computing and Informatics Journal*, vol. 7, no. 1, Art. 3, 2022. Available: https://doi.org/10.54623/fue.fcij.7.1.3.

[9] P. Arya and S. Jain, "Text-based emotion detection," *Int. J. Computer Engineering & Technology (IJCET)*, vol. 9, no. 3, pp. 95–104, 2018. Available: http://iaeme.com/Home/issue/IJCET?Volume=9 & Issue=3.

[10] P. Sohan Sai, "ISEAR dataset [Data set]," Kaggle, 2023. Available: https://www.kaggle.com/code/psohansai/isear/notebook.

[11] P. Nandwani and R. Verma, "A review on sentiment analysis and emotion detection from text," *Soc. Netw. Anal. Min.*, vol. 11, Art. 81, 2021. Available: https://doi.org/10.1007/s13278-021-00776-6.

[12] G. Swathi, B. Meghana Patnaik, A. Janani, K. Karthik, and J. Divya, "Emotion recognition from text using machine learning," *UGC CARE Listed (Group -I) Journal*, vol. 11, no. 12, 2022.