Multi-Label Emotion Prediction for Text using Glorot Entropy Kernel based Gated Recurrent Unit Approach

Rakesh Kancharla, Deepak Kumar and Ananda Kumar Kinjarapu

Abstract—**Multi-label emotion prediction plays a vital role in analyzing human emotions through textual data. Previous studies have primarily focused on polarity analysis for classifying emotions. However, interpreting human expressions through text is a complex task, as emotions often co-occur with significant correlations. To overcome the aforementioned challenges, this study proposes a Glorot Entropy Kernel-based Gated Recurrent Unit (GEK-GRU) model for multi-label text emotion prediction. The proposed GRU handles sequential data and effectively captures temporal relationships between words and sentences for text emotion prediction. In multi-label emotion prediction, a single text instance expresses multiple emotions simultaneously, thus requiring enhanced capability to learn temporal patterns to distinguish between closely related emotional states. This improves the model's overall classification accuracy. Textual data required for evaluation is acquired and preprocessed using tokenization and lemmatization techniques, empowering the model in eliminating meaningless words from sentences, followed by text vectorization carried out using Glove Vectorization. Furthermore, Pearson correlation coefficient technique is employed in the selection of significant vectorized features for precise emotion prediction. The experimental results of the GEK-GRU method's multi-label prediction performance demonstrates a macro-precision of 88.32% and 69.85% on the SemeEVal-2018 and RenCECps datasets. These outcomes prove the GEK-GRU's superiority over the existing prediction approaches namely, Transfer learning and Attention based Bidirectional Long Short-Term Memory with convolutional layer (AC-Bi-LSTM).**

Index Terms—**Multi-label emotion prediction, Gated Recurrent Unit, Glorot Entropy Kernel, Glove vectorization, Pearson Correlation Coefficient.**

I. Introduction

EMOTIONS play a significant role in a person's life, influencing decision-making, as well as their physical and mental well-being. Emotions can be expressed through actions or as pieces of text to convey what an author intends to express [1-2]. The key objective of text emotion classification is to analyze and interpret these emotions to better understand the underlying intent and mental state of the author [3-4]. Developing robust text emotion classification models is essential due to the extensive significance and impact of emotions. These models are widely employed in various fields, such as medicine, marketing, and education, where understanding emotional context is critical [5]. Text emotion classification is implemented to analyze emotions in a given statement and determine the underlying semantics, including emotions such as sadness, anger, joy, and worry [6-8]. Traditionally, text emotion classification relied on single-label models that limit predictions to a single dominant emotion per text. This gives rise to a major drawback in single-label text emotion prediction, making it difficult for the models to accurately define the exact emotional state of the authors [9]. Thus, multilabel text emotion prediction-based models are used to precisely analyze and predict various emotions in text [10].

As multi-label emotion prediction is a relatively challenging task due to the difficulty in identifying various emotions expressed by the same words with different meanings, artificial intelligence methods, such as machine learning and deep learning methods, are employed to learn emotions at a deeper level and effectively differentiate between them [11-12]. However, conventional machine learning methods often treat each label (emotion) independently and rely on heuristic optimization methods to handle multi-label predictions [13]. These approaches fail to capture context dependencies between emotions, such as "sadness" and "fear," thereby resulting in inaccurate prediction outcomes [14]. Thus, deep learning-based prediction methods are implemented to efficiently learn the contextual dependencies of different emotions for enhanced multilabel emotion prediction through accurate differentiation between overlapping emotions [15]. However, existing research continues to face challenges with accurate emotion prediction, owing to inadequate representation of multiple emotions, resulting in barriers in efficient differentiation. To overcome this limitation, the Glorot Entropy Kernel-Gated Recurrent Unit (GEK-GRU) is proposed for prediction of multi-label emotions in text. The primary contributions of this study are as follows:

- Preprocessing techniques namely, tokenization, stop word removal, and lemmatization are applied to further enhance processing by eliminating unnecessary text from sentences and converting raw text into a useful format.
- Glove vectorization, Term Frequency and Inverse Document Frequency (TF-IDF) are used to transform

Manuscript received February 20, 2025; revised August 14, 2025.

Rakesh Kancharla is an assistant professor of Computer Science and Engineering Department, NIT Meghalaya, Shillong, Meghalaya 793003 India. (corresponding author to provide phone: +91-8977014251; e-mail: rakeshkancharla@nitm.ac.in).

Deepak Kumar is an assistant professor of Computer Science and Engineering Department, NIT Meghalaya, Shillong, Meghalaya 793003 India. (e-mail: Deepak.kumar@nitm.ac.in).

Ananda Kumar Kinjarapu is professor of Computer Science and Engineering Department, Aditya University, Surampalem, Andhra Pradesh 533437 India. (e-mail: akkinjarapu@gmail.com).

the preprocessed text into vectorized feature representations for effective multi-label emotion prediction.

- The proposed GEK-GRU is employed for multi-labelled prediction of emotions in text by learning subtle differences between emotions based on the entropy kernel, further assisting the GRU model to accurately predict multiple emotions in text.

The remainder of the manuscript is structured as follows: Section 2 presents a literature review of existing models designed for emotion classification, and Section 3 details the proposed methodology introduced for multi-label emotion prediction. Section 4 presents the results and discusses the evaluation of the proposed model. Finally, Section 5 concludes the study.

II. Literature Survey

This section discusses the existing DL approaches currently used for multi-label emotion prediction tasks. These models are reviewed to identify their advantages and limitations over the proposed model.

Ameer et al. [16] designed three DL models for multi-label emotion classification using transfer learning ROBERTa, XLNet, and DistiBert approaches. The designed transfer-learning models were employed to classify multiple emotions of tweets written in English and Chinese. The main advantage of the designed transfer-learning methodology is the incorporation of multiple attention mechanisms, which augments the model to streamline its focus on accurate feature extraction with relevant meanings. However, the transfer learning models XLNet, DistiBert, and ROBERTa faced limitations with identifying relationships between classes and phrases, and in accurately differentiating between emotions, such as optimism and disgust.

Li et al. [17] explored a mixed emotion prediction model through the analysis of social media texts using feature extraction. The attention-based bidirectional Long Short-Term Memory with a convolutional layer (AC-Bi-LSTM) was employed to enhance feature extraction, which positively impacted the prediction and classification of mixed emotions. The key advantage of the AC-Bi-LSTM method is the integration of an emotion correlation technique, which is used to identify relationships among diverse emotions and differentiate them effectively. However, the AC-Bi-LSTM model failed to understand the contextual embedding of texts, which affected multi-label classification.

Deng and Ren [18] introduced a multi-label emotion detection model based on the BERT framework, which incorporates emotion-specific correlation learning and feature extraction. The BERT-based model included three sub-modules for correlation learning and feature extraction, containing data specific to each emotion. The BERT approach professed in its ability to extract emotional information from the extracted features, which were then fuses at the sentence, context, and correlation levels to improve multi-label emotion identification in text. However, the introduced BERT model struggled to effectively differentiate various emotions, impacting overall performance in multi-label emotion prediction.

Ameer et al. [19] presented a Graph Attention Network (GAN)-based emotion classification model for semantic representation of text. The proposed GAN model was utilized for the multi-label classification of various emotions expressed in text through semantic and syntactic representations. The GAN model placed its primary focus on analyzing semantic representations over syntactic representations to better understand the meanings and correlations between emotions in the text. However, the model failed to interpret mixed emotions within a single sentence, making it difficult for the detection model to accurately identify emotions.

Le et al. [20] introduced a deep-learning-based transformer architecture for multi-label and multimodal emotion recognition. The introduced transformer model was employed to enrich feature processing in video images for emotion recognition on social media. It effectively fused multimodal video data and achieved better accuracy than unimodal methods. However, because the modalities were processed independently during training, there was a lack of integration between data, which affected the feature-extraction process and led to suboptimal results.

III. Methodology

This study proposes the GEK-GRU model for multi-label emotion prediction in textual framework, consisting of four phases: dataset, preprocessing, feature extraction, and prediction. A block diagram of the functioning of the proposed multi-label emotion prediction is given in Fig. 1. The text samples are acquired from the publicly available datasets, SemEval-2018 and RenCECps, and preprocessed using tokenization, stop words, and removal of URL and emojis techniques. These preprocessed text samples are converted into word vectors and then fed as input vectors for feature extraction to obtain effective prediction results. This step ensures the extraction of significant features by collecting accurate information about emotions. Finally, after accessing these input feature vectors, emotions are predicted from the text using the proposed GEK-GRU emotion-prediction approach.

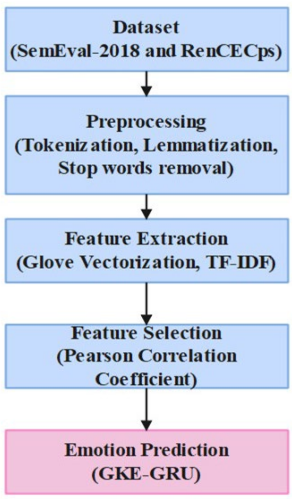

Fig. 1. Block diagram for proposed Multi-label emotion prediction

A. Dataset

SemEval-2018 and RenCECps are two publicly available datasets used to acquire textual samples for multi-label emotion prediction.

SemEval-2018

SemEVAL-2018 [21] is an open-source public dataset containing data on Twitter posts. The SemEval-2018 dataset consists of tweet posts data from 2016 and 2017 utilized to estimate the model's accuracy in the prediction of the emotions of Joy, Anger, Anticipation, Love, Surprise, Fear, Optimism, and Disgust. The SemEVAL-2018 dataset consists of tweets written in three languages: English, Arabic, and Spanish. This study utilizes an English corpus for multi-label emotion prediction.

Ren-CECps

The Ren-CECps dataset is sourced from the Chinese We blog, which contains 34,702 sentences that are further categorized into 27,008 training and 7,694 test sentences. It consists of sentences with eight types of emotions: joy, anxiety, expectation, surprise, hate, love, anger, and sorrow. These emotions fall within the range of [0, 1], where emotional states with values greater than 0 are assigned a label of 1 and all others are assigned a label of 0. These textual data are then fed for preprocessing to convert the raw data into an appropriate format to enhance the multi-label emotion prediction of the model.

B. Preprocessing

The text data are passed as input for preprocessing of text for further processing in emotion prediction and to enhance data quality. This technique is applied to remove unwanted and meaningless characters such as URL, hashtags, and emojis from text-based data. This shortens the length of sentence segments in the text data. The pre-processing techniques used in this study are described in the following sections.

C. Tokenization

The acquired text data consists of raw words that have similar meanings but express different emotions. In multi-label emotion prediction, the process of dividing long sentences into smaller text segments based on punctuation marks is referred to as tokenization. The goal of this method is to generate a list of words by breaking down longer sentences, which assists the model in understanding the meanings of the text more effectively during training, and facilitates accurate categorization of multi-label emotions.

D. Lemmatization

Lemmatization refers to the process of reducing words to their base or dictionary forms by understanding their intended meanings and context. This technique aligns words with their corresponding meanings and synonyms based on their morphological properties. The advantage of lemmatization is that it reduces the size and complexity of features used to train the proposed prediction model. The lemmatized text is then passed through a stop-word removal process to convert it into a more useful format.

E. Stop Words Removal

In a sentence, stop words refer to commonly occurring words such as "am," "is," and "the," which typically do not carry any emotional meaning. These words have limited significance in a sentence and are not useful for multi-label emotion prediction. Therefore, stop words in the tokenized text are removed to direct the model's focus towards words that carry meaningful emotional information. The preprocessed text is then fed as input for feature extraction.

F. Feature Extraction

The preprocessed data are fed into the feature extraction phase to capture the most significant features containing relevant information corresponding to various emotion classes. Feature extraction is the process of identifying and extracting relevant features from the preprocessed text for accurate multi-label prediction. This process transforms unstructured data into structured data, thereby enabling efficient emotion classification. In this study, GloVe-based word embedding and TF-IDF techniques are employed to extract relevant features from the preprocessed text.

G. Word Vectorization

Word vectorization is a widely used method that generates continuous vector representations of high quality for the acquired text, and captures the semantic similarity between words. In this process, each word in a text is represented by a vector based on its occurrence and context in a predefined dictionary. In this study, the pre-trained GloVe vectorization method [22] is used to generate a vector matrix with numerical values for each preprocessed text. The mathematical representation of the m-dimensional vectors generated for each word is given by (1).

$$D = \{w_1, w_2, \dots, w_p\} \in V^{p*m} \tag{1}$$

Here, D represents the input data and V denotes the vector matrix for dimension m. Because the text length often varies for each input, the data varies in lengths which are sized into uniform ranges with a size S for vectorization. When the size of the input text data is smaller than S, the size is increased using the zero-padding technique, whereas if the input size exceeds S, it is reduced accordingly. In this study, the length of the word vector is fixed at 35 for each text or preprocessed token. The size of the input text data in the vector matrix is expressed by (2).

$$D = \{w_1, w_2, \dots, w_p\} \in V^{S*m} \tag{2}$$

Where S denotes the size of the input text data. These vectorized data are then passed through the TF-IDF process to obtain a numerical text representation.

H. TF-IDF

After vectorization, the TF-IDF method is utilized to convert text into meaningful numerical representations for feature extraction, which helps enhance multi-label emotion prediction using the proposed DL-based prediction model. This method is used as a metric to determine the significance of text-based terms on its term frequency, and is mathematically expressed as in (3).

$$TF = \frac{No\ of\ times\ term\ T\ appears\ in\ doc\ d}{total\ no\ of\ terms\ in\ doc\ d} \tag{3}$$

Where, TF represents the frequency term which estimates the frequency of text appearance in a sentence or phrase. The TF-IDF model also incorporates the inverse document frequency method, which is utilized to estimate how a phrase or text is relevant to an emotion. This IDF technique also assigns an additional weight d, as expressed in (4).

$$IDF = log\frac{total\ no.of\ documents}{total\ no.of\ document\ with\ term\ t} \tag{4}$$

Therefore, the final weight of the text in dataset d is calculated using (5):

$$TF - IDF = TF \times IDF \tag{5}$$

In multi-label emotion prediction for textual data, the advantage of utilizing GloVe word vectorization and TF-IDF as feature extraction methods lies in their ability to efficiently identify relevant words in a sentence, thereby aiding the proposed prediction model in better understanding the underlying meanings of the text effectively. These vectorized and numerically represented features are then forwarded to the proposed feature selection phase to enhance the multi-label emotion prediction performance with textual data.

I. Feature Selection

After feature extraction, word embedding techniques, GloVe and TF-IDF, are employed in the feature selection phase to identify relevant features containing significant information for accurate multi-label emotion prediction in text. The primary function of the feature selection process is to select the most informative features from the extracted set, thereby minimizing the time required for emotion prediction. Furthermore, the Pearson Correlation Coefficient (PCC) is used for feature selection, which reduces feature dimensionality and selects the most important features for accurately predicting different emotions, thereby improving the performance of the proposed multi-label emotion prediction model.

Pearson Correlation Coefficient

The PCC method is employed for accurate feature extraction for improved multi-label emotion prediction through determining the relationship between the extracted features. The correlation between features is calculated using Equation (6).

$$R_{pq} = \frac{\sum_{i=1}^{n}(p_i - \bar{p})(q_i - \bar{q})}{\sqrt{\sum_{i=1}^{n}(p_i - \bar{p})^2(q_i - \bar{q})}} \tag{6}$$

Where, R denotes the degree of linear correlation between features ranging from -1 to +1. If the value is greater than zero, two features x and y are positively correlated, meaning that as the value of the variable increases, the other variable also increases. If the value is less than zero, the features are negatively correlated, indicating that as one variable increases, the other decreases. The strength of the correlation increases with an absolute value of R. An absolute value close to -1 or +1 represents a strong correlation, whereas a value near zero denotes a weak correlation between the features. This correlation helps select significant features effectively. The features chosen by the PCC method are then fed as inputs to the proposed multi-label emotion prediction model.

J. Proposed Multi-Label Emotion Prediction Model

The extracted features are passed as input to the proposed GEK-GRU multi-label prediction of emotions in textual data. A GRU is a type of recurrent neural network and an improvised version of the LSTM. The proposed GEK-GRU model allows for accurate emotion prediction by utilizing past contextual information between words. Unlike models that process each feature individually, GRU processes inputs sequentially, making it more stable than models that process each feature individually. The architecture of the proposed GRU model involves two gates: reset gate and the update gate [23] [24]. These gates operate based on the previous hidden state $h < sub > t-1 </sub >$ and the current input $x < sub > t </sub >$, which are mathematically expressed in (7) and (8), respectively.

The reset gate controls how much of the past information is forgotten while the update gate selects information from the memory that is to be retained up to the current moment [23] [24]. These two gates determine the last transmitted hidden state h_{t-1} and the input x_t of the current node, as mathematically formulated in (7) and (8):

$$z_t = \sigma(W_{xz}x^t + W_{hz}h^{t-1} + b_z) \tag{7}$$
$$r_t = \sigma(W_{xr}x^t + W_{hr}h^{t-1} + b_r) \tag{8}$$

where σ denotes the sigmoid function, x^t indicates the input text, z_t and r_t represent the update and reset gates, h^{t-1} represents the hidden layer output, W_{xz} and W_{xr} denote the weights of the update and reset gates, and b_r and b_z represent the reset and update gates.

The information flow extracted from the selected features in the network is regulated by the two gates in the GRU. The previously mentioned update and reset gates are responsible for determining which information should be retained and which should be discarded, while also alleviating gradient descent-related issues. However, GRU models have certain limitations, such as overfitting, difficulty in capturing long-term dependencies, and limited expressiveness in handling multi-label emotion predictions simultaneously. To address these issues, the proposed model incorporates Glorot initialization for weight distribution, and employs an Entropy Kernel (EK) as the activation function within the GRU. The EK improves the learning process of neurons, thereby enhancing model reliability.

K. GEK Technique

The limitations of the GRU are addressed by the proposed GEK technique in four major steps, as explained below.

Step 1: Initially, the selected features are fed as inputs to the GRU-based prediction model, as represented in (9).

$$\Psi_t = \{\Psi_{i1}, \Psi_{i2}, \Psi_{i3}, \dots, \Psi_{in}\} \quad for\ t = t1, t2, t3, \dots, tn \tag{9}$$

where Ψ represents the extracted features, and tn denotes the number of time steps. Furthermore, the proposed Glorat weight estimation technique's functioning is mathematically expressed in (10).

$$\omega = \sqrt{\frac{2}{\phi_{in}-\phi_{op}}} \qquad (10)$$

Here, ϕ_{in} and ϕ_{op} denote the input and output of the learnable features of the gates, respectively. These weight-initialized input features are fed in the next step.

Step 2: In the hidden layers of the GRU, the reset and update gates are determined based on the input features. The mathematical representation of these two gates is given in (11) and (12), respectively.

$$z_t = \sigma(W_{xz} + h^{t-1} + \Psi_z) \qquad (11)$$
$$r_t = \sigma(W_{xr} + h^{t-1} + \Psi_r) \qquad (12)$$

W_{xz} and W_{xr} represent the weights of the update and reset gates, respectively, and σ indicates the sigmoid function. The weight-based update and reset gates are passed for further processing.

Step 3: The optimal hidden state is determined by multiplying the previous hidden state by the update and reset gates, along with the proposed activation function, which is used to enhance multi-label emotion prediction simultaneously. The hidden state is mathematically formulated in (13):

$$hs_{\tilde{t}} = \Phi\left(W_{hs} * (r_t * hs_{t-1}, \Psi_t)\right) \qquad (13)$$

where Φ denotes the entropy-centric kernel function and W_{hs} indicates the weight of the hidden state of the GRU. The proposed entropy kernel-based activation is determined using (14).

$$\Phi = \sum_{k=1}^{e} E_k \ log_2(E_k) \qquad (14)$$

Where, E_k refers to the kernel element and e represents the number of kernels. These kernel-based elements are fed into the final step, as described below.

Step 4: The output from the hidden state layers in the GRU network is represented by (15).

$$hS_t = (1-U_t) * \ hs_{t-1} + \ U_t * \ hs_{\tilde{t}} \qquad (15)$$

Here, U_t represents the output from hidden states. The proposed GEK improves the GRU by integrating the kernel entropy activation function and Glorot initialization technique. The integration of the entropy technique enhances the performance of the GRU model employed for multi-label prediction of emotions in text by analyzing the entropy of the prediction performance of the GRU model to effectively differentiate between multiple emotions. This technique refines the boundaries between closely related emotions, allowing the model to distinguish emotions, such as fear, sadness, joy, and surprise, with greater accuracy. Algorithm 1 represents the overall process of the proposed GEK-GRU prediction model employed for multi-label emotion prediction in the text.

Algorithm 1:

```
Input: Extracted attribute features
Output: Multi-label emotion prediction
Begin
  Initialize weight parameters ω_ℜ, ω_U, ω_hs, maximum
iteration Itr_max
  Set Iteration Itr = 1
  while Itr ≤ Itr_max do
    for each time step (t1 to tn) do
      Compute Reset gate (ℜ_t)
        Compute Update gate (U_t)
      Evaluate candidate hidden state function (hs_t~)
      hs_t ≅ Φ(w_hs * [ℜ_t * hs_{t-1}, Ψ_t])
      Calculate hidden state (hs_t)
       hs_t = (1 − U_t) + hs_{t-1} + U_t * hs_t~
       If (hs_t = 1) then
      {
      Predict
      }
       else
      {
       Predict
      }
       end if
  end for
end while
end=0
```

IV. Results and Discussion

The experimental results of the GEK-GRU model for emotion prediction are presented in this section. The proposed prediction model is implemented using Python 3.9, with the simulation run on a system configured with Windows 10 OS, Intel i5 processor, and 16 GB RAM. The performance metrics used to evaluate the GEK-GRU method includes accuracy and three macro-level metrics: macro-recall, macro-precision, and macro F1-score. Mathematical representations of these performance metrics are provided in Eqs.(16)–(19).

$$Accuracy = \frac{TP+FN}{TP+TN+FP+FN} \qquad (16)$$

$$macro-precision = \frac{1}{8}\sum_{i=1}^{8} Pre \qquad (17)$$

$$macro-recall = \frac{1}{8}\sum_{i=1}^{8} Rec \qquad (18)$$

$$macro-f1 = \frac{1}{8}\sum_{i=1}^{8} f1 \qquad (19)$$

Where, TP denotes True Positive, TN denotes True Negative, FP denotes False Positive, and FN denotes False Negative.

A. Quantitative and Qualitative Analysis

The performance of the GEK-GRU prediction model is evaluated and compared with existing models used for text-based multi-label emotion prediction. Evaluation measures, namely accuracy, macro-recall, macro-precision, and macro F1-score, are used to evaluate the effectiveness of existing methods, including the Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), LSTM, and GRU. Table I presents the performance of the GEK-GRU method utilizing the SemEval-2018 dataset. Table II presents an analysis on the RenCECps dataset.

TABLE I
PERFORMANCE ANALYSIS OF THE PROPOSED METHOD IN THE SEMEVAL-2018 DATASET

Methods	Macro-Precision (%)	Macro-Recall (%)	Macro-F1-score (%)
CNN	75.39	74.67	75.02
RNN	76.21	75.89	76.04
LSTM	78.87	77.27	78.06
GRU	81.40	80.33	80.86
Proposed GEK-GRU Method	88.32	87.23	86.86

The performance of the GEK-based initialization method with the GRU model is evaluated and compared with that of various other initialization techniques. Normal weight initialization, random weight initialization, and weight initialization methods are considered for evaluation alongside the proposed GEK method. Fig 2 illustrates the performance of the GEK-GRU method in comparison with these initialization techniques on the SemEval-2018 dataset. Fig 3 presents a similar case on the RenCECps dataset.

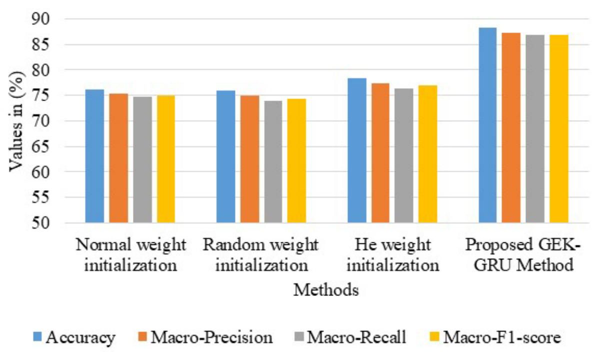

Fig. 2. Performance analysis of the proposed initialization method in SemEval-2018 dataset

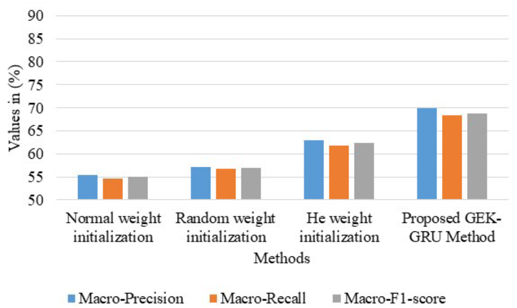

Fig. 3. Performance analysis of the proposed initialization method in RenCECps dataset

TABLE II
PERFORMANCE ANALYSIS OF THE PROPOSED METHOD IN RENCECPS DATASET

Methods	Macro-Precision (%)	Macro-Recall (%)	Macro-F1-score (%)
CNN	56.41	55.42	55.91
RNN	57.53	56.53	57.02
LSTM	59.82	58.82	59.31
Bi-LSTM	62.93	61.93	62.42
Proposed GEK-GRU Method	69.85	68.34	68.79

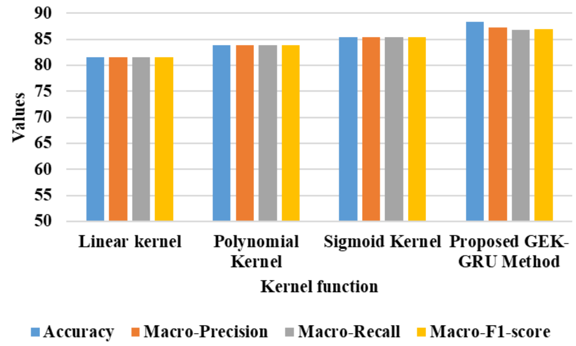

Fig. 4. Performance analysis of the proposed kernel function in SemEval-2018 dataset

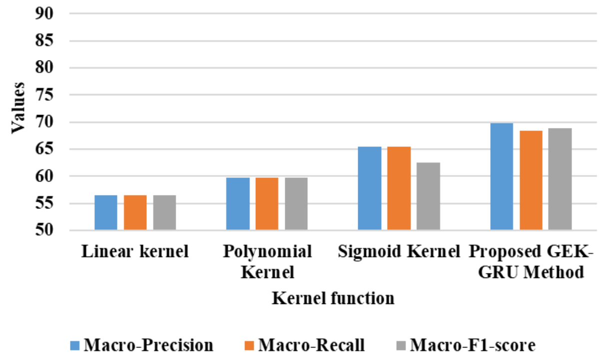

Fig. 5. Performance analysis of the proposed kernel function in RenCECps dataset

The GEK-GRU prediction model is evaluated against existing kernel function techniques on the SemEval 2018 dataset, as shown in Fig 4. Initialization techniques such as the linear kernel, polynomial kernel, and sigmoid kernel are compared with the GEK-GRU model for multi-label emotion prediction in text. Similarly, in Fig. 5, on the RenCECps dataset, the performance of the proposed GEK-GRU prediction model is evaluated and compared against the linear, polynomial, and sigmoid kernels.

The performance of the proposed GEK-GRU-based prediction model is evaluated against existing feature-extraction techniques on the RenCECps and SemEval-2018 datasets, as presented in Table III. For feature extraction from text, word vectorization and TF-IDF are utilized in this research and are compared with state-of-the-art methods, namely, Continuous Bag of Words (CBoW) and skip-gram.

TABLE III
PERFORMANCE ANALYSIS OF FEATURE EXTRACTION METHOD USED FOR MULTI-LABEL EMOTION PREDICTION FOR TEXTS

Dataset	Methods	Accuracy (%)	Macro Precision (%)	Macro Recall (%)	F1-score
RenCECps	CBoW	83.49	83.08	82.90	82.98
	Skip gram	84.34	82.18	81.94	82.05
	GloVe and TFIDF	88.32	87.23	86.86	86.92
SemEval-2018	CBoW	65.33	65.12	64.95	65.03
	Skip gram	66.51	65.98	65.61	65.79
	GloVe and TFIDF	69.85	68.34	68.79	69.85

TABLE IV
PERFORMANCE EVALUATION OF GEK-GRU IN TERMS OF COMPUTATIONAL TIME AND MEMORY CONSUMPTION FOR MULTI-LABEL EMOTION PREDICTION

Methods	Datasets	Computational time (ms)	Memory consumption (KB)
CNN	SemEval-2018	102	136
RNN		95	123
LSTM		92	114
GRU		85	105
Proposed GEK-GRU Method		83	102
CNN	RenCECps	153	145
RNN		123	134
LSTM		98	108
GRU		89	103
Proposed GEK-GRU Method		85	99

TABLE V. EVALUATION OF PROPOSED GEK-GRU METHOD BASED ON CLASS WISE RESULT USING SEMEVAL-2018 DATASET

Classes	Accuracy (%)	Macro-Precision (%)	Macro-Recall (%)	F1-score (%)
Anger	87.8	87.0	86.9	86.9
Anticipation	88.4	87.0	87.2	87.2
Disgust	89.0	87.4	87.4	87.3
Fear	88.2	86.9	86.9	86.8
joy	88.7	87.3	87.3	87.2
Love	88.3	87.0	87.0	87.0
Optimism	88.3	86.6	86.7	86.7
Surprise	88.4	87.5	87.5	87.2
Overall	88.32	87.23	86.86	86.92

TABLE VI. EVALUATION OF PROPOSED GEK-GRU METHOD BASED ON CLASS WISE RESULT USING RENCECPS DATASET

Classes	Macro-Precision (%)	Macro-Recall (%)	F1-score (%)
Love	69.26	68.13	68.61
Surprise	71.63	68.55	69.55
Joy	69.77	68.37	68.77
Anxiety	69.69	68.45	68.92
Expect	69.48	68.22	68.78
Sorrow	69.34	68.16	68.57
Anger	69.34	68.13	68.78
Hate	68.35	68.34	68.61
Overall	69.85	68.34	68.79

TABLE VII. CROSS VALIDATION RESULTS OF PROPOSED GEK-GRU METHOD IN MULTI-LABEL EMOTION PREDICTION

Methods	Datasets	Accuracy (%)	Macro-Precision (%)	Macro-Recall (%)	F1-score (%)
k=3.00	SemEval-2018	86.81	85.79	84.66	83.57
k=5.00		**88.32**	**87.23**	**86.86**	**86.92**
k=7.00		87.35	86.89	85.61	85.45
k=9.00		84.74	83.67	82.79	81.54
k=3.00	RenCECps	N/A	66.41	65.25	64.17
k=5.00		N/A	**69.85**	**68.34**	**68.79**
k=7.00		N/A	67.36	66.41	65.72
k=9.00		N/A	65.31	64.94	63.90

The performance of the GEK-GRU model in terms of the computational time and memory usage is presented in Table IV. From the results, it is clear that the GEK-GRU model achieves superior computational efficiency when compared to traditional models such as CNN, RNN, LSTM, and GRU on both the SemEval-2018 and RenCECps datasets. This highlights the fact that the model is most appropriate for resource-constrained environments.

The performance evaluation of the proposed method based on class wise emotions using the SemEval-2018 dataset is presented in Table V. The performance of the emotion classification model is evaluated based on the metrics of accuracy, macro-precision, macro-recall, and F1-score across eight different emotion classes. These results indicate the robustness and effectiveness of the model in handling multiclass emotion recognition tasks.

The evaluation of the GEK-GRU performance based on class-wise emotions using the RenCecps dataset is illustrated in Table VI. The model's performance on eight emotion classes are assessed using macro-, macro-, and F1-score, with overall values of 69.85%, 68.34%, and 68.79%, respectively. The class-wise emotion performance results of the proposed model indicate a balanced performance across distinct emotions. Although moderate, the consistent metrics reflect the model’s capability in multiclass emotion detection.

Table VII presents the cross-validation results by varying the hyperparameter and evaluating the performance of the GEK-GRU model on the SemEval-2018 and RenCECps datasets. From the results, it is clear that k = 5 folds, yielding best overall performance with 88.32% accuracy and the highest macro-precision, recall, and F1-score across both datasets. The value k = 5 is considered balanced and optimal as it achieves a superior trade-off between model generalization and discriminative ability across both the datasets.

Statistical Analysis

Generally, statistical analysis involves the use of statistical methods to analyze data, identify patterns, and obtain meaningful conclusions from collected information. These statistical analyses are crucial to ensure the reliability and validity of the research findings, enabling the testing of hypotheses with predictions about a larger population. Tables VIII, IX, and X represent the statistical analysis of the proposed GEK-GRU method and feature extraction method on the SemEval-2018 and RenCECps datasets. The proposed GEK-GRU method, which integrates the Glorat Entropy Kernel (GEK) with the GRU architecture, showcases superior performance compared to traditional deep learning models such as CNN, RNN, LSTM, and GRU on both the SemEval-2018 and RenCECps datasets. This is evident from the highest mean performance scores of 87.47% and 68.99 % and lowest p-values 0.04 and 0.03 for the two datasets respectively, indicating statistical significance. GEK enhances GRU by embedding entropy-based mechanisms, ensuring the model to prioritize emotionally rich features and eliminate irrelevant information, which is especially effective in multi-label emotion classification.

Ablation study

An ablation study is performed to determine the contributions of each component of the proposed model system by removing irrelevant components, and then analyzing the resultant performance of the proposed multi-label emotion prediction model. Table XI presents the ablation study of the proposed GEK-GRU method based multi-label emotion prediction on the SemEval and RenCECps datasets. The results clearly demonstrate that the integration of GloVe and the TF-IDF-based feature extraction method with the GRU model significantly improves the performance by providing semantically rich and weighted features.

TABLE VIII. STATISTICAL ANALYSIS OF PROPOSED GEK-GRU METHOD IN MULTI-LABEL EMOTION PREDICTION USING SEMEVAL-2018 DATASET

Methods	Mean (%)	SD	p-value
CNN	75.03	0.36	0.08
RNN	76.05	0.13	0.09
LSTM	78.07	0.80	0.07
GRU	80.86	0.43	0.06
Proposed GEK-GRU Method	87.47	0.75	0.04

TABLE IX. STATISTICAL ANALYSIS OF PROPOSED GEK-GRU METHOD IN MULTI-LABEL EMOTION PREDICTION USING RENCECPS DATASET

Methods	Mean (%)	SD	p-value
CNN	55.91	0.40	0.09
RNN	57.09	0.41	0.08
LSTM	59.32	0.41	0.07
GRU	62.43	0.41	0.06
Proposed GEK-GRU Method	68.99	0.62	0.03

TABLE X. STATISTICAL ANALYSIS OF THE FEATURE EXTRACTION METHOD USED IN MULTI-LABEL EMOTION PREDICTION USING THE RENCECPS DATASET

Methods	Datasets	Mean (%)	SD	p-value
CBoW	SemEval	83.11	0.22	0.06
Skip gram		82.63	1.04	0.05
GloVe and TFIDF		87.33	0.61	0.03
CBoW	RenCECps	65.11	0.14	0.07
Skip gram		65.97	0.37	0.04
GloVe and TFIDF		69.21	0.66	0.02

Additionally, the proposed GRU model's GEK mechanism enhances the its ability to focus on emotionally relevant inputs by emphasizing the high-entropy features. The model performs superiorly by combining GloVe + TF-IDF with GEK-GRU, leveraging both contextual embeddings and entropy-based attention, resulting in higher accuracy, precision, recall, and F1-score across both the SemEval and RenCECps datasets. This confirms the complementary strength of feature richness and entropy-guided learning in multi-label emotion prediction.

TABLE XI. ABLATION STUDY OF PROPOSED GEK-GRU METHOD UTILIZED IN MULTI-LABEL EMOTION PREDICTION WITH VARIOUS CONFIGURATIONS

Methods	Datasets	Accuracy (%)	Macro-Precision (%)	Macro-Recall (%)	F1-score (%)
GRU	SemEval	78.64	78.49	74.33	74.19
Glove and TF-IDF +GRU		81.54	81.33	81.14	81.26
GEK-GRU		84.51	84.48	84.21	84.32
Glove and TF-IDF + GEK-GRU		88.32	87.23	86.86	86.92
GRU	RenCECps	N/A	60.83	60.73	6.62
Glove and TF-IDF +GRU		N/A	63.29	63.16	62.98
GEK-GRU		N/A	66.78	66.61	66.47
Glove and TF-IDF + GEK-GRU		N/A	69.85	68.34	68.79

B. Comparative Analysis

The proposed GEK-GRU approach is evaluated alongside various emotion prediction models, including Transfer Learning [17], BERT [18], and UCCA-GAT [19], on both SemEval-2018 and RenCECps datasets. Tables XII and XIII present the comparative analysis of the GEK-GRU approach for multi-label emotion prediction in text on both the SemEval-2018 and RenCECps datasets.

TABLE XII. COMPARATIVE ANALYSIS OF THE PROPOSED METHOD

Methods	Dataset	Accuracy (%)	Macro-Precision (%)	Macro-Recall (%)	F1-score (%)
Transfer Learning methods [16]	SemEval-2018 dataset	62.4	N/A	N/A	60.3
UCCA-GAT [19]		61.2	N/A	N/A	57.8
Proposed GEK - GRU		88.32	87.23	86.86	86.92

TABLE XIII. COMPARATIVE ANALYSIS OF PROPOSED METHOD

Methods	Dataset	Macro-Precision (%)	Macro-Recall (%)	F1-score (%)
BERT [19]	RenCECps dataset	46.10	52.21	48.31
Proposed GEK -GRU		69.85	68.34	68.79

Research Implication

The proposed GEK-GRU model offers significant advancements in multi-label emotion prediction by learning subtle emotional distinctions using an entropy kernel integrated with a GRU. The combination of GloVe vectorization and TF-IDF enhanced feature richness boost generalizability across different datasets. The proposed GEK kernel's ability supports the GRU model's ability to distinguish overlapping emotions, reducing misclassification, thereby leading to improved model robustness. Additionally, the model demonstrates superior performance in terms of accuracy, macro-precision, and F1-score when compared to traditional models. Efficient feature extraction further contributes to lower computational costs and memory consumption, rendering the model scalable.

C. Discussion

GEK-GRU based Emotion Prediction

The proposed GEK-GRU-based multi-label emotion prediction model achieves superior text classification results after being evaluated using both prediction-related and macro-based metrics. Transfer Learning [16] has limitations in identifying relationships between classes and phrases and faces challenges in differentiating between emotions such as optimism and disgust. The AC-Bi-LSTM [17] model fails to understand the contextual embedding of texts, which negatively impacts m59.32ulti-label classification. BERT [18] struggles to effectively differentiate various emotions, limiting its ability to predict multi-label emotions in texts. The GAN [19] method fails to understand mixed emotions within a single sentence, making it difficult for the detection model to identify emotions accurately. To address these limitations, the GEK-GRU model is proposed for precise multi-label emotion prediction to efficiently differentiate between emotions. The GEK method analyzes the entropy of the prediction performance of the GRU model to distinguish multiple emotions. Additionally, the proposed GEK technique, an entropy-based scaling method, enhances performance in sequence tasks such as multi-label emotion prediction, as a result of its stable and improved weight distribution. This technique helps refine boundaries between closely related emotions to clearly distinguish between the emotions of fear, sadness, joy, and surprise. By capturing these subtle differences, the GRU model predicts multi-label emotions in text with greater precision.

Feature extraction and Feature selection

To extract emotion-related features from the preprocessed textual data, the GloVe and TF-IDF techniques are used for multi-label emotion prediction. The GloVe method has the advantage of identifying emotion-related words even when the words do not appear frequently in the dataset. In RenCECps, which contains mixed Chinese and English data, GloVe efficiently extracts features that contain more emotion-relevant information. Additionally, the TF-IDF model ensures the extraction of the most important words that contain rich emotional information, thus enhancing multi-label emotion prediction. By utilizing these two methods for feature extraction, the model extracts more significant features and contextual similarities from text, thereby improving the performance of the emotion prediction model. However, feature extraction models such as CBoW and Skip-gram still face drawbacks as it ignores rare words with limited context, which affects the extraction of relevant information and impacts the prediction results.

V. CONCLUSION

To overcome this problem, a Glorot Entropy Kernel-based Gated Recurrent Unit (GEK-GRU) model is proposed for multi-label emotion prediction in text. The proposed GRU model handles sequential data and captures temporal relationships between words and sentences to implement text-based emotion prediction effectively. In multi-label emotion prediction, a single text can represent multiple emotions simultaneously, making it challenging for the prediction model to accurately predict various emotions. Thus, the GRU model is used in multi-label emotion prediction, as it efficiently learns the temporal patterns of diverse emotions, helping differentiate between emotions that overlap in time or context. The sample text is acquired and preprocessed using tokenization and lemmatization techniques to eliminate unnecessary text from the sentences. Using the Pearson correlation coefficient technique, significant vectorized features are selected for precise emotion prediction using the GEK-GRU method. The experimental results of the proposed method for multi-label prediction prove that it achieves macro-precisions of 88.32% and 69.85% for the SemEval-2018 and RenCECps datasets, demonstrating visibly higher results than those of the existing prediction approaches such as RoBERTa and AC-Bi-LSTM. In the future, a transformer-based approach with an optimization algorithm may be implemented to further enhance multi-label emotion prediction in the text.

REFERENCES

[1] Y. Zhou, X. Kang, and F. Ren, "Prompt consistency for multi-label textual emotion detection," IEEE Trans. Affective Comput., vol. 15, no. 1, pp. 121–129, Mar. 2023.

[2] D. Park, S. Lim, Y. Choi, and H. Oh, "Depression Emotion Multi-Label Classification Using Everytime Platform with DSM-5 Diagnostic Criteria," IEEE Access, vol. 11, pp. 89093–89106, Aug. 2023.

[3] F. Qi, H. Zhang, X. Yang, and C. Xu, "A Versatile Multimodal Learning Framework For Zero-shot Emotion Recognition," IEEE Transactions on Circuits and Systems for Video Technology, vol. 34, no. 7, pp. 5728–5741, 2024.

[4] 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, Aug. 2023.

[5] M. Lim, M. Yi, P. Kim, and J. Shin, "Multilabel emotion recognition technique considering the characteristics of unstructured conversation data," Mobile Inf. Syst., vol. 2022, No. 1, p. 2057198, May 2022.

[6] U. Malik, S. Bernard, A. Pauchet, C. Chatelain, R. Picot-Clemente, and J. Cortinovis, "Pseudo-Labeling with Large Language Models for Multi-Label Emotion Classification of French Tweets," IEEE Access, vol. 12, pp. 15902–15916, Jan. 2024.

[7] Y. Zhu and O. Wu, "Elementary discourse units with sparse attention for multi-label emotion classification," Knowledge-Based Syst., vol. 240, p. 108114, Mar. 2022.

[8] R. C. Cabral, S. C. Han, J. Poon, and G. Nenadic, "MM-EMOG: Multi-Label Emotion Graph Representation for Mental Health Classification on Social Media," Robotics, vol. 13, no. 3, p. 53, Mar. 2024.

[9] Z. Li, X. Li, H. Xie, F. L. Wang, M. Leng, Q. Li, and X. Tao, "A novel dropout mechanism with label extension schema toward text emotion classification," Inf. Process. Manage., vol. 60, no. 2, p. 103173, 2023.

[10] R. Štrimaitis, P. Stefanovič, S. Ramanauskaitė, and A. Slotkienė, "A Combined Approach for Multi-Label Text Data Classification," Comput. Intell. Neurosci., vol. 2022, no. 1, p. 3369703, Jun. 2022.

[11] Y. Ma, X. Liu, L. Zhao, Y. Liang, P. Zhang, and B. Jin, "Hybrid embedding-based text representation for hierarchical multi-label text classification," Expert Syst. Appl., vol. 187, p. 115905, Jan. 2022.

[12] Y. Yan, F. A. Liu, X. Zhuang, and J. Ju, "An R-transformer_BiLSTM model based on attention for multi-label text classification," Neural Process. Lett., vol. 55, no. 2, pp. 1293–1316, Jun. 2023.

[13] M. Y. Saidabad, H. Hassanzadeh, S. H. S. Ebrahimi, E. Khezri, M. R. Rahimi, and M. Trik, "An efficient approach for multi-label classification based on Advanced Kernel-Based Learning System," Intell. Syst. Appl., vol. 21, p. 200332, Mar. 2024.

[14] J. L. Huan, A. A. Sekh, C. Quek, and D. K. Prasad, "Emotionally charged text classification with deep learning and sentiment semantic. Neural Comput. Appl., vol. 34, no. 3, pp. 2341–2351, Sep. 2022.

[15] N. Shelke, S. Chaudhury, S. Chakrabarti, S. L. Bangare, G. Yogapriya, and P. Pandey, "An efficient way of text-based emotion analysis from social media using LRA-DNN," Neurosci. Inf., vol. 2, no. 3, p. 100048, Sep. 2022.

[16] I. Ameer, N. Bölücü, M. H. F. Siddiqui, B. Can, G. Sidorov, and A. Gelbukh, "Multi-label emotion classification in texts using transfer learning," Expert Syst. Appl., vol. 213, p. 118534, Mar. 2023.

[17] Y. Li, J. Chan, G. Peko, and D. Sundaram, "Mixed emotion extraction analysis and visualisation of social media text," Data Knowl. Eng., vol. 148, p. 102220, Nov. 2023.

[18] J. Deng and F. Ren, "Multi-label emotion detection via emotion-specified feature extraction and emotion correlation learning," IEEE Trans. Affective Comput., vol. 14, no. 1, pp. 475–486, Oct. 2020.

[19] I. Ameer, N. Bölücü, G. Sidorov, and B. Can, "Emotion classification in texts over graph neural networks: Semantic representation is better than syntactic," IEEE Access, vol. 11, pp. 56921–56934, May 2023.

[20] H. D. Le, G. S. Lee, S. H. Kim, S. Kim, and H. J. Yang, "Multi-label multimodal emotion recognition with transformer-based fusion and emotion-level representation learning," IEEE Access, vol. 11, pp. 14742-14751, Feb. 2023.

[21] SemEval 2018 Dataset: https://www.kaggle.com/datasets/context/semeval-2018-task-ec (Accessed on October 2024)

[22] D. Mahto and S. C. Yadav, "Emotion prediction for textual data using GloVe based HeBi-CuDNNLSTM model," Multimedia Tools Appl., Vol. 83, no. 7, pp. 18943–18968, Jul. 2024.

[23] H. Shi and D. Zhao, "License plate recognition system based on improved YOLOv5 and GRU," IEEE Access, vol. 11, pp. 10429–10439, Jan. 2023.

[24] X. Gao, C. Xue, W. Jiang, and B. Liu, "Attention-GRU Based Intelligent Prediction of NOx Emissions for the Thermal Power Plants," IAENG Int. J. Comput. Sci., vol. 51, no. 8, pp. 1171–1181, 2024.

Rakesh Kancharla is currently working as an Assistant Professor at KL University in Vijayawada and is pursuing his Ph.D. at NIT Meghalaya in Shillong, India.

He earned his B.Tech. degree in computer science and Engineering and his M.Tech. degree in computer science and technology from CRR College of Engineering in Eluru, India. He successfully qualified for the APSET-2019 examination conducted by Andhra University in Visakhapatnam, India. His research interests include data mining, text mining, pattern recognition, and related areas.

Deepak Kumar has been working as an Assistant Professor at the National Institute of Technology Meghalaya, India, since 2013.

He earned his B.E. degree in Electronics and Communication Engineering from VTU, Belgaum, and his M.E. degree in Computer Science and Technology from Bengal Engineering and Science University, Shibpur, India. He completed his Ph.D. at NIT Meghalaya. His research interests include computational mathematics, digital signal processing and machine learning.

Ananda Kumar Kinjarapu has been serving as a Professor at Aditya University in Kakinada, India.

He completed his M.Tech. in computer science and technology from JNTUH in Hyderabad and his B.Tech. in computer science and engineering from CRR College of Engineering in Eluru, India. He also holds a Ph.D. from JNTUH in Hyderabad. His research interests include image processing, signal processing, and pattern recognition.