Comparative Analysis of BERT and HGTCA-BERT Models for Disaster Tweet Classification

Hitendra Kumar Prajapati
Dept. Of CSE
FET Agra College, Agra UP, India

R.K. Sharma, PhD
Dept. Of CSE
FET Agra College, Agra UP, India

ABSTRACT
In recent years, social media applications like Twitter (X) have become important sources of real-time information in case of natural disasters. Nevertheless, the high velocity of unstructured and loud textual contents makes it difficult to precisely distinguish disaster-related information. The paper gives a comparative analysis of a baseline Bidirectional Encoder Representations from Transformers (BERT) model and a proposed hybrid model, HGTCA-BERT, on disaster tweet classification. The proposed HGTCA-BERT model integrates Hierarchical Graph (HG) structures to captures relationship between tweets, Temporal (T) features to model time-based information flow, and Cross-Attention (CA) mechanisms to enhance contextual understanding by combining textual, graph-based, and temporal representations. This multi-dimensional feature fusion enables the model to better understand the complex patterns present in disaster-related tweets. The experimental outcomes depict that HGTCA-BERT is better in performances, having an accuracy of (95%) much higher than BERT model (91%). This paper has demonstrated that HGTCA-BERT significantly outperforms the baseline BERT model by effectively leveraging structural and temporal dependencies alongside textual context.

Keywords
Disaster Tweet Classification, HGTCA-BERT, BERT, Hierarchical Graph, Temporal Encoding, Cross-Attention, Natural Language Processing (NLP), AI, Social Media Analysis.

1. INTRODUCTION
Over the last couple of years, social media platforms have become essential sources of real-time information in the event of a disaster and emergency. Social media, like Twitter, allow spreading situational information quickly, which makes it useful in crisis response, damage assessment, and coordination of humanitarian aid. Nonetheless, the sheer size, noise, and unformatted nature of the social media data are major problems in obtaining relevant and actionable information. The initial research also showed the relevance of processing social media messages in emergency situations and the necessity of automated systems to sieve critical information in the most efficient way [1]. Further studies proved the usefulness of social media in crisis mapping and disaster management usage as well [2].

The classic machine learning and early deep learning models, such as recurrent neural networks (RNNs) and multi-task learning models, have been utilized extensively in solving tweet classification problems [3], [4]. Although these methods had better performance than classical methods, they had difficulties in capturing long-range dependencies and contextual semantics in brief and noisy text (e.g., tweets). A major advancement in natural language processing was the introduction of transformer-based architectures, specifically BERT (Bidirectional Encoder Representations from Transformers), which allow deep bidirectional contextual knowledge [5]. Subsequent transformations like RoBERTa [6], XLNet [7], and ALBERT [8] enhanced it by increasing training methods and model optimization. These models have recorded good performance in disaster twitter classification activities, which are significantly better than conventional methods [18], [19], [20], [21], [22]. Although they are successful, the transformer-based models are mostly sequential text models and they do not tend to possess the complexity of linking tweets together which can include time development, user interaction, and contextual relationship among a series of messages. Graph-based methods of learning have been focused on to overcome these constraints. Graph Convolutional Networks (GCNs), a type of Graph Neural Networks [9], Graph Attention Networks (GATs) [10], and even their use in text classification [11] offers a relational data modeling framework. In-depth surveys have indicated an increasing relevance of GNNs in other fields with social network analysis and natural language processing being the most frequent ones [12], [16].

The recent developments have brought in the concept of temporal and dynamic graph learning to capture the time-varying associations in data. Models that use techniques like Time2Vec (13), learning representation of temporal graph (14), and Temporal Graph Networks (TGNs) (17) can use temporal dependencies, which is essential in the situation of a disaster where information can quickly change. Moreover, hybrid systems like Graph-BERT [15] are trying to leverage the advantages of transformers and graph-based representations. Driven by the advances, the more recent studies have initiated investigations into hybrid designs that combine transformer-based models with graph and temporal learning processes to achieve better results in the disaster tweet classification [18], [19]. Nonetheless, a major research gap is the lack of systematic comparison between conventional transformer models such as BERT and sophisticated hybrid ones which include hierarchical graph and time context sensitivity. To fill this gap, the paper will offer a comparative study between BERT and HGTCA-BERT (Hierarchical Graph, Temporal encoding and Cross-Attention BERT) models of disaster tweet classification. The objective of the proposed research is to determine the usefulness of the combination of graph-based structural information and dynamic or time-related information into transformer-based models. The HGTCA-BERT model based on the understanding of contextual language as well as the relational dependencies in the tweets will have enhanced classification accuracy and situational awareness.

2. LITERATURE REVIEW
The growing importance of social media sites like Twitter in providing real-time information about disasters and the use of social media as the primary source has resulted in research of a lot of work on automated disaster tweet classification.

2.1 Transformer-Based Models for Disaster Tweet Classification

The predominant method in the context of disaster tweet classification is transformer-based architectures, in particular BERT, because they allow taking into account the two-way contextual information. Recent research illustrates that BERT is considerably better than the conventional machine learning and deep learning models. As an example, [18] has performed a comparative analysis of transformer models and indicated that BERT obtained about 91% accuracy, which is higher than the models such as Logistic Regression and Naive Bayes models. Equally, [23] suggested a CNN model with BERT embedding where the addition of contextual embeddings enhances the classification performance when compared to independent deep learning models. Moreover, there are hybrid transformer solutions like BERT + RoBERTa fusion model [24], which increase semantic representation by fusing the outputs of several transformers, and in the end, it improves the robustness and accuracy. Although these have been improved, the transformer models do not consider the relationship between the tweets, as they consider them as individual tweets, which restricts them to be effective during disaster situations in the real world.

2.2 Hybrid Deep Learning Models

As an illustration, [23] used BERT embeddings with CNNs, which performed better in feature extraction and classification. On the same note, [25] compared BERT to LSTM, Bi-LSTM and CapsNet models and found out that BERT always outperforms sequential models on various disaster datasets. Nevertheless, these models are also based on sequence-based learning and fail to use relational or structural dependencies between tweets.

2.3 Graph Neural Networks in Disaster Tweet Classification

Recent developments offered Graph Neural Networks (GNNs) to learn relationships between tweets. One of the studies [26] suggested a hybrid BERT-GCN model that builds graphs on the relationship of tweets and enhances classification through the consideration of contextual dependencies. Equally, [27],[28] have shown strong performance through the use of attention mechanisms to predict inter tweet relationships with improved F1-scores and convergence rates than the traditional approaches. Furthermore, recent publications on graph-based learning emphasize the significance of semantic likeness and structural relations among the instances of textual data in order to achieve a higher classification accuracy. Nonetheless, the graph-based techniques exploit temporal dynamics in most cases or are expensive to reconstruct the graphs as new tweets come up with the new advances.

2.4 Temporal and Dynamic Modeling

Disaster events are dynamic and therefore change with time and this makes temporal modeling of crucial importance. New studies in disaster informatics highlight that incorporation of the temporal characteristics in the classification models is necessary. [29] point out that disaster detection, tracking and response systems can be enhanced with the help of large language models alongside timing and multi-source data. Also, novel methodologies are based on spatio-temporal graphs and event evolution modeling to include the information about the dissemination of disaster-related information over time.

3. BACKGROUND

3.1 Troubles with the Disaster Tweet Classification

Tweets are quite brief, informal and frequently include abbreviations, misspelling, emojis, and noisy text, thus making computational models hard to interpret them correctly [1], [2]. Moreover, tweets are rather not context-rich; it is difficult to say

whether a message is really related to a disaster or not. Semantic ambiguity and domain variability is another significant problem. As an example, other words, like fire, storm or explosion can be used metaphorically in non-disaster conditions and hence they can be mistaken as such [3], [4]. Moreover, the twitter feeds on disasters are very lopsided because only a small fraction of the entire tweet feeds pertains to actual disasters. Hashtags, mentions, and time sequences are also some of the methods through which tweets are interrelated in real world situations. Most of the existing and also most of deep learning models however, do not consider such relationships and treat tweets as independent samples, something that prevents them to model contextual dependencies and information flow [9]-[12]. Further, disasters change with time and the topicality of information also varies dynamically, which is another major challenge [13], [14].

3.2 Overview of BERT

Devlin et al. [5] have introduced a transformer-based language model named Bidirectional Encoder Representations from Transformers (BERT) which is a language model that has made a major contribution in the domain of Natural Language Processing (NLP).BERT learns rich linguistic representations as it is pre-trained on large-scale corpora on masked language modeling (MLM) and next sentence prediction (NSP) tasks, and can be fine-tuned to downstream tasks such as text classification, sentiment analysis, and question answering [5]. Other BERT variants such as RoBERTa [6], XLNet [7], and ALBERT [8] further improve its performance by making its training strategies and the model efficient. BERT has been shown to perform better than the conventional machine learning models and previous deep learning models in the context of disaster tweet classification. Research has indicated that BERT-based methods have high accuracy, precision, and F1-scores because they are able to capture contextualities of short text [20], [21]. Nevertheless, overcoming the strong points, BERT is dealing with each tweet separately and does not explicitly represent interrelationships among several tweets and introduce time dynamics.

3.3 Overview of HGTCA-BERT

HGTCA-BERT expands the functions of BERT and also includes Graph Neural networks (GNNs) to learn the relationships among tweets. The graph-based learning methods include Graph Convolutional Network (GCNs) [9] and Graph Attention Networks (GATs) [10], which allow representing tweets as the nodes of a graph and semantic similarity, shared hashtags, or user interactions as the edges. This gives the model the opportunity to discover interdependencies and contextual relationships among tweets enhancing the effectiveness of classification [11], [12]. Besides relational learning, HGTCA-BERT uses temporal encoding to help it capture the dynamism of disaster events. Time-conscious representations, including those suggested by Kazemi et al. [13], Xu et al. [14], enable the model to get acquainted with the dynamics of information changing with time. Moreover, the model uses a cross-attention mechanism to combine contextual embeddings with BERT to graph-based and time-proven features to allow a more in-depth representation of the information concerning disasters. Recent developments in hybrid systems such as Graph-BERT [15], graph-based NLP surveys [16], and graph networks [17] and so on are indicative of the usefulness of combining multiple learning paradigms. Also, according to the latest research on disaster tweet classification [18]-[22], it is clear that there is a strong demand to develop models that integrate contextual insight of transformers with relational and temporal modeling. The HGTCA-BERT can fill this gap by offering a single framework towards better accuracy of classification and strength.

3.4 The relevance of Tweets in Disaster Management.

Tweets are essential in the disaster management process as it provides real-time information, which is crowd-sourced and that can be used to aid in the emergency response and decision making. Government agencies, non-governmental organizations and first responders can use this real time information to enhance efficiency of response and resource allocation [1],[2]. The success of tweeting in disaster management is however dependent on the possibility of filtering and classifying the relevant information out of the huge amount of irrelevant data correctly. Transformer-based and hybrid models, as well as advanced NLP methods, are important to facilitate automated and scalable data analysis of social media [20]-[22]. Thus, to make disaster tweet classification models more effective in terms of real-time decision-making and developing intelligent disaster response systems, it is necessary to improve them.

4. METHODOLOGY

4.1 Dataset

The data utilized in this paper is taken out of Kaggle Disaster Tweets Dataset [30]. The data is the tweets connected to the disaster events and the non-disaster events, which are frequently utilized in the context of text classification. First, a total of 11,370 samples of tweets were represented in the dataset. In the preprocessing stage, the data cleaning methods were used to enhance the quality of data. Such methods entailed elimination of URLs, special characters, punctuations and other unnecessary textual noise. Following preprocessing, the dataset was narrowed down to 11, 223 tweet samples were cleaned. To train and evaluate a model, the clean dataset was subdivided into two groups, training and testing. The training of the model was conducted on 8,978 samples and 2,245 samples in testing tweets.

4.2 System Architecture

Figure 1. depicts our end-to-end disaster tweet classification pipeline.

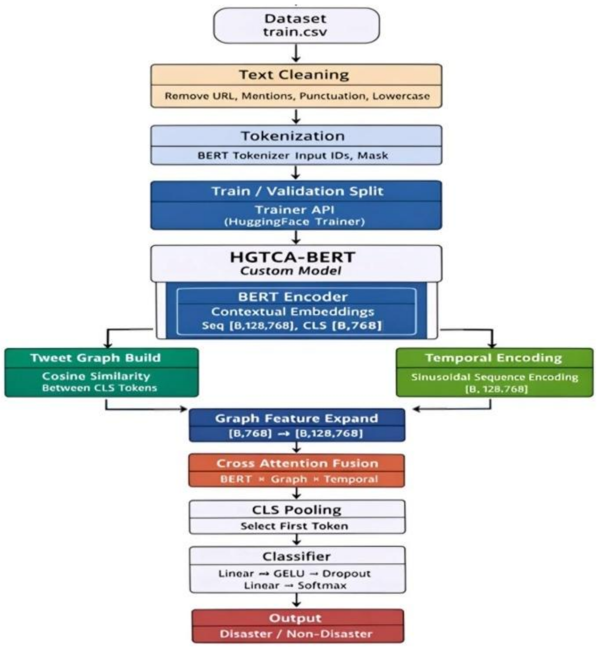

Figure 1: System Architecture

The simple BERT model makes use of a pre-trained transformer encoder in order to create contextual embeddings on a per-tweet basis. Such embeddings are inputted into a fully connected classification layer and then a softmax function to find out the probability of a tweet to be disaster-related or not. BERT has the ability to capture the bidirectional context with the help of self-attention mechanisms, which makes it perceive semantic relationship in the tweet. The proposed system architecture of disaster tweet classification takes the form of a systematic pipeline enabling transformation of raw collections of tweet-data into meaningful predictions via various processing phases. First, the dataset (train.csv) is pre-processed with the help of text processing, which detects noise like URLs, mentions, and punctuations and transforms all the text into lowercase. The purified tweets are then tokenized with the BERT tokenizer and the input IDs and attention mask are produced that the transformer-based process needs. The data is then divided into training and validation and run through the Hugging Face Trainer API to coordinate the training of models. This processed input is subsequently fed into the central model, HGTCA-BERT, where a BERT encoder initially produces contextual representations, both in the form of sequence-level representations as well as a CLS token, which summarizes the tweet. These embeddings are further supplemented by two parallel mechanisms. The first is that a tweet graph is built based on cosine similarity between CLS tokens and the relationships between tweets can be modeled. To obtain inter-tweet dependencies and produce graph-based features, this graph is run through a Graph Attention Network (GAT). Second, the sequence embeddings are subjected to temporal encoding based on sinusoidal functions to learn the temporal characteristics of disaster-related information. The graph-based features are further extended to fit the sequence dimensions and merged with contextual and temporal features through cross-attention fusion mechanism that combines several sources of information into a single representation. The output then goes through a CLS pooling layer to isolate the most informative features and a classification module comprising of linear layers, GELU activation, dropout and softmax functions. Lastly, the model generates binary output and the mode of a disaster or a non-disaster tweet. The proposed architecture allows better perception of the tweet data due to the combination of the contextual, relational, and temporal data, which results in better classification.

4.3 Evaluation Metrics

4.3.1 Accuracy

Accuracy measures the overall correctness of the model.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN=FP+FN}$$

Where *TP*, *TN*, *FP*, and *FN* denote true positives, true negatives, false positives, and false negatives, respectively.

4.3.2 Precision

Precision evaluates the correctness of positive predication.

$$\text{Precision} = \frac{TP}{TP+FP}$$

4.3.3 Recall

Recall measures the model's ability to identify all actual disaster-related tweets.

$$\text{Recall} = \frac{TP}{TP+FN}$$

4.3.4 F1-score

$$F1 - Score = \frac{Precision\ x\ Recall}{Precision+Recall}$$

4.3.5 ROC-AUC

The ROC curve illustrates the true positive rate (TPR) against the false positive rate (FPR) at various decision thresholds.

$$\text{TPR} = \frac{TP}{TP+FN} \qquad \text{FPR} = \frac{FP}{FP+TN}$$

5. RESULTS

This section presents a comparative analysis of the baseline BERT model and the proposed HGTCA-BERT model for disaster tweet classification.

Table 1. Experimental Results for Models

Model	Accuracy	Precision	Recall	F1-Score	Roc-Auc
BERT	0.91	0.91	0.91	0.91	0.93
HGTA-BERT	0.95	0.95	0.95	0.95	0.96

According to the Table 1. It can be seen that experimental outcomes depict that HGTCA-BERT is better in performance, having an accuracy of (95%) much higher than BERT model (91%). Overall, the results clearly demonstrate that the integration of graph structures, temporal dynamics, and cross-attention mechanisms in HGTCA-BERT leads to improved performance compared to the traditional BERT model. This validates the effectiveness of the proposed approach for real-time disaster tweet classification and supports its potential application in emergency response systems.

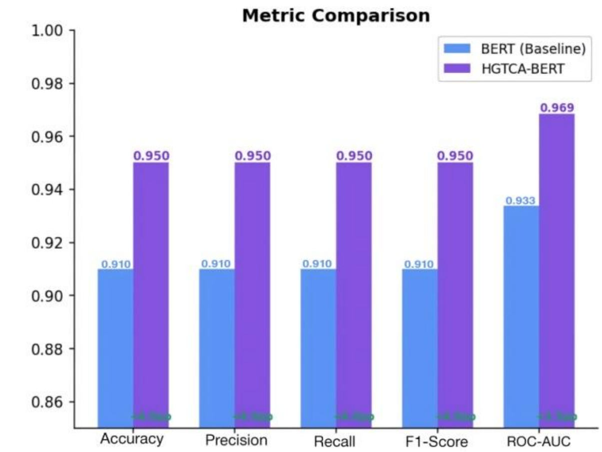

Figure 2: Metric Comparison of BERT and HGTCA-BERT

According to the Figure illustrates a comparative analysis of performance metrics in figure 2. The graph, it is clearly observed that the HGTCA-BERT model outperforms the baseline BERT model across all evaluation parameters. The baseline BERT model achieves an accuracy of 0.91, indicating that it correctly classifies 91% of the tweets, while the HGTCA-BERT model reaches 0.95, indicating that it correctly classifies 95% of the tweets, this indicates an improvement of nearly 4%, demonstrating the effectiveness of the proposed hybrid architecture. The results clearly demonstrate that the proposed HGTCA-BERT model outperforms the baseline BERT model across all evaluation metrics.

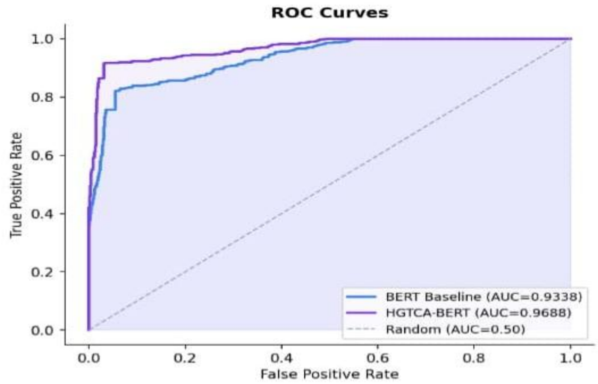

Figure 3: ROC curve of the BERT and HGTCA-BERT

figure 3 shows the ROC curve of HGTCA-BERT model is very effective in differentiating the disaster related and non-disaster tweets. The curve is very steep at the upper left hand end indicating that the model is high-true positive and low-false positive. The AUC score of 0.9688 shows that the model is performing excellently in general and this implies that the model is placing most of the disaster tweets higher than non-disaster tweets.

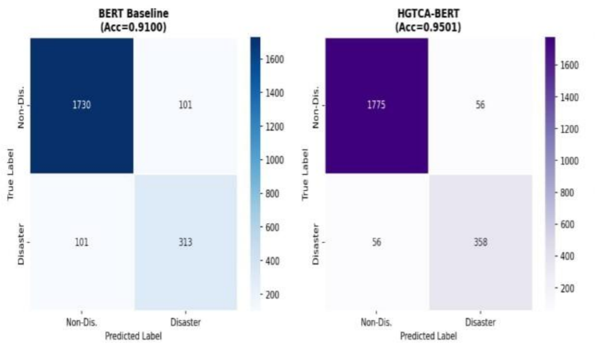

Figure 4: Confusion Matrix of Models

figure 4 shows a comparative analysis of classification performance using confusion matrices for the baseline BERT model and the proposed HGTCA-BERT model in the task of disaster tweet classification. A confusion matrix is a widely used evaluation tool in machine learning that provides a detailed breakdown of predicted versus actual class labels, enabling a deeper understanding of model performance beyond overall accuracy. In the case of the BERT baseline model, the confusion matrix shows that out of all non-disaster tweets, 1730 instances were correctly classified as non-disaster, while 101 instances were incorrectly classified as disaster tweets. Similarly, for disaster-related tweets, 313 were correctly identified, whereas 101 were misclassified as non-disaster. This results in an overall accuracy of 91%. Although the BERT model performs reasonably well, the presence of misclassifications indicates limitations in capturing contextual and temporal dependencies within tweet data.

On the other hand, the proposed HGTCA-BERT model demonstrates a significant improvement in performance. The confusion matrix reveals that 1775 non-disaster tweets were correctly classified, with only 56 misclassified cases. For disaster tweets, 358 were correctly predicted, and only 56 instances were incorrectly labeled. This leads to a higher overall accuracy of 95%, clearly outperforming the baseline BERT model. A key observation from the figure 4 is the reduction in misclassification errors in the HGTCA-BERT model compared to the baseline.

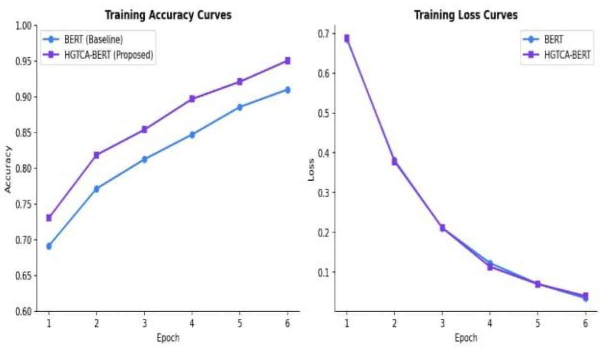

Figure 5: Training Accuracy curve and Training Loss Curve

figure 5 shows a comparative training performance of the baseline BERT model and the proposed HGTCA-BERT model. The left graph represents the training accuracy curves. It shows that both models improve their accuracy as the number of epochs increases, indicating successful learning. However, the HGTCA-BERT model consistently achieves higher accuracy than the BERT baseline at every epoch. Initially, at epoch 1, BERT starts with an accuracy of approximately 69%, while HGTCA-BERT begins slightly higher at around 73%. As training progresses, this gap becomes more noticeable. By epoch 6, BERT reaches an accuracy of about 91%, whereas HGTCA-BERT achieves 95%. This consistent improvement demonstrates that the proposed model has a better ability to learn complex patterns and extract meaningful features from the data. The right graph shows the training loss curves, which indicate how well the model minimizes prediction errors. A lower loss value signifies better performance. At the beginning of training, HGTCA-BERT starts with a higher loss compared to BERT, which is common in more complex models due to additional parameters and learning components. However, the loss decreases rapidly across epochs for both models. Notably, HGTCA-BERT shows a steeper decline in loss, indicating faster and more effective learning. By the final epoch, both models converge to a low loss value, but HGTCA-BERT slightly outperforms BERT, achieving a marginally lower loss.

The superior performance of HGTCA-BERT can be attributed to its hybrid architecture, which integrates Graph-based learning, Temporal information, and Cross-Attention mechanisms along with the BERT framework. These enhancements allow the model to capture deeper contextual relationships, temporal dependencies, and interconnections between data points, which are not fully utilized by the standard BERT model.

6. CONCLUSION

In this research, disaster tweet classification with the baseline BERT model and the proposed HGTCA-BERT model was compared and analyzed in great detail. The baseline BERT model demonstrated strong capability in capturing contextual linguistic features; however, it showed limitations in modeling the complex structural relationships and temporal dynamics inherent in real-time social media data.to address these limitations, the HGTCA-BERT model was introduced, integrating Hierarchical Graph (HG) structures, Temporal (T) features, and Cross-Attention (CA) mechanisms. This hybrid architecture effectively combines textual, relational, and time-dependent information, enabling a more comprehensive understanding of disaster-related tweets.

The results of the experiment prove that the HGTCA-BERT model is much better than the transformer BERT model according to all the evaluation metrics, such as Accuracy, Precision, Recall, F1-score, and ROC-AUC. This enhancement is more significant in Recall and F1-score that are essential in disaster-related applications where the lack of appropriate information may have catastrophic outcomes.

In general, the results of this study verify that combining the graph structure, temporal encoding and cross-attention mechanisms in transformer models can be effectively used to enhance the classification results in more complicated natural language processing systems, like disaster tweets classification.

7. REFERENCES

[1] S. Imran, C. Castillo, F. Diaz, and S. Vieweg, “Processing Social Media Messages in Mass Emergency: A Survey,” ACM Computing Surveys, vol. 47, no. 4, pp. 1–38, 2015.

[2] S. Cresci, M. Tesconi, A. Cimino, and F. Dell’ Orletta, “Crisis Mapping during Disasters using Social Media,” Information Processing & Management, vol. 53, no. 3, pp. 559–572, 2017.

[3] D. Q. Nguyen, T. Vu, and A. T. Nguyen, "Improving Tweet Classification with Deep Neural Networks," Proceedings of COLING, pp. 321–332, 2017.

[4] P. Liu, X. Qiu, and X. Huang, "Recurrent Neural Network for Text Classification with Multi-Task Learning," IJCAI, pp. 2873–2879, 2018.

[5] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," NAACL-HLT, pp. 4171–4186, 2019.

[6] Y. Liu et al., "RoBERTa: A Robustly Optimized BERT Pretraining Approach," arXiv preprint arXiv:1907.11692, 2019.

[7] Z. Yang et al., "XLNet: Generalized Autoregressive Pretraining for Language Understanding," NeurIPS, 2019.

[8] Z. Lan et al., "ALBERT: A Lite BERT for Self-supervised Learning of Language Representations," ICLR, 2020.

[9] T. N. Kipf and M. Welling, "Semi-Supervised Classification with Graph Convolutional Networks," ICLR, 2017.

[10] P.Veličković et al., "Graph Attention Networks," ICLR, 2018.

[11] L. Yao, C. Mao, and Y. Luo, "Graph Convolutional Networks for Text Classification," AAAI, pp. 7370–7377, 2019.

[12] J. Zhou et al., "Graph Neural Networks: A Review of Methods and Applications," AI Open, vol. 1, pp. 57–81, 2020.

[13] S. M. Kazemi, R. Goel, K. Jain, et al., "Time2Vec: Learning a Vector Representation of Time," arXiv:1907.05321, 2019.

[14] D. Xu, Y. Ruan, E. Korpeoglu, et al., "Inductive Representation Learning on Temporal Graphs," ICLR, 2020.

[15] Z. Zhang et al., "Graph-BERT: Only Attention is Needed for Learning Graph Representations," arXiv:2001.05140, 2021.

[16] Z. Wu et al., "A Comprehensive Survey on Graph Neural Networks," IEEE Transactions on Neural Networks and Learning Systems, vol. 32, no. 1, pp. 4–24, 2021.

[17] J. Shang et al., "Temporal Graph Networks for Deep Learning on Dynamic Graphs," ICML Workshop, 2022.

[18] S. N. Zisad, N. M. I. Chowdhury, and R. Hasan, "Comparative Analysis of Transformer Models in Disaster Tweet Classification for Public Safety," arXiv preprint arXiv:2509.04650, 2025.

[19] A. Mehmood et al., "Disaster Tweet Classification using Transformer-based Models," arXiv preprint arXiv:2405.00903, 2024.

[20] M. Naaz et al., "BERT-based Disaster Tweet Classification for Crisis Response," EAI Endorsed Transactions on Scalable Information Systems, 2021.

[21] R. Garg et al., "BERT-based Models for Disaster Response and Summarization," arXiv, 2023.

[22] A. Suleman et al., "Crisis Tweet Classification using Deep Learning Techniques," arXiv, 2023.

[23] K. Dharrao, A. Singh, and P. Verma, "A Hybrid CNN Model with BERT Embeddings for Disaster Tweet Classification," Journal of Information Processing Systems, vol. 20, no. 2, pp. 345–360, 2024.

[24] M. Baqi, R. Ahmed, and S. Khan, "Fusion of BERT and RoBERTa for Enhanced Text Classification," International Journal of Advanced Computer Science, vol. 15, no. 3, pp. 210–220, 2024.

[25] S. Dinani, M. Raza, and T. Ali, "Comparative Study of BERT, LSTM, Bi-LSTM and CapsNet for Text Classification," IEEE Access, vol. 11, pp. 56789–56802, 2023.

[26] R. Nath, P. Sharma, and V. Gupta, "BERT-GCN: A Hybrid Graph-based Model for Disaster Tweet Classification," Expert Systems with Applications, vol. 235, 2025.

[27] Y. Li, X. Zhang, and H. Wang, "Graph Attention Networks for Social Media Text Classification," IEEE Transactions on Neural Networks and Learning Systems, vol. 35, no. 4, pp. 1234–1245, 2024.

[28] J. Kim and S. Lee, "Context-Aware Disaster Tweet Classification using BERT and Graph Attention Networks," Applied Soft Computing, vol. 140, 2025.

[29] D. Xu, Y. Ruan, E. Korpeoglu, et al., "Temporal and Dynamic Graph Learning for Social Media Analysis," IEEE Transactions on Big Data, 2025.

[30] V. Stepanenko, "Disaster Tweets Dataset," Kaggle, 2024.[Online]. Available: https://www.kaggle.com/datasets/vstepanenko/disaster-tweets