2024 International Conference on Emerging Systems and Intelligent Computing (ESIC) | 979-8-3503-4985-6/24/$31.00 ©2024 IEEE | DOI: 10.1109/ESIC60604.2024.10481660

Classification of Humanitarian Crisis Response Through Unimodal Multi-Class Textual Classification

Tarun Jain
Department of CSE
Malaviya National Institute of Technology and Manipal University Jaipur
Jaipur, India
tarunjainjain02@gmail.com

Dinesh Gopalani
Department of CSE
Malaviya National Institute of Technology
Jaipur, India
dgopalani.cse@mnit.ac.in

Yogesh Kumar Meena
Department of CSE
Malaviya National Institute of Technology
Jaipur, India
ymeena.cse@mnit.ac.in

***Abstract*—In the realm of disaster response it is crucial to emphasize the significant role that social media plays as an immediate and dynamic source of information. In times of both natural and human- created calamities citizens resort to sites like Twitter to broadcast updates in the form of text and multimedia. These updates vary from reporting on wounded or killed individuals to damage assessments and the hunt for missing or discovered folks. The value of this online information for humanitarian organizations is well-established, a large chunk of research has typically concentrated on textual content analysis. This research study seeks to offer a unique technique in textual classification of a vast and diversified multi-modal dataset dubbed "CrisisMMD," which is collected from Twitter after significant natural disasters. This dataset offers three separate types of annotations, each geared to fulfill a wide variety of crisis response and management needs, responding to the various requirements of humanitarian organizations. In our attempts to increase the efficiency of humanitarian operations, we argue for a approach that leverages the power of textual material inside social media. To achieve this objective we apply cutting-edge machine and deep learning techniques, to construct a framework with a common representation successfully breaking down the conventional boundaries of textual modes. Our study done utilizing real world catastrophe datasets indisputably proves the superiority of our suggested method when compared to other models. The results of this study not only underline the untapped potential of textual aspects in crisis response but also offer an appropriate framework for categorizing the Multi-Class CrisisMMD dataset textually. This method signals a new era of strong and effective catastrophe management, promising to have a big influence on the humanitarian scene. Experiments conducted on the CrisisMMD dataset, which simulates real-world catastrophes, have yielded validation results demonstrating the consistent and significant outperformance of our technique. The findings reveal a notable advancement compared to the previously published works by 3% and the current state-of-the-art surpassing them by over 6.73%. The performance improvement observed ranges from 3% to 6.73%, underscoring the effectiveness of our approach in addressing the challenges posed by real-world catastrophe scenarios.**

***Index Terms*—Classification, Ensemble, CrisisMMD, BERT, SVM, TF-IDF, Bag of Words , Deep Learning, Humanitarian, XGBoost**

I. Introduction

The research study proposes a novel methodology for text categorization within the Multi-Class CrisisMMD Dataset, leveraging machine learning models and natural language processing methodologies to enhance the efficacy and efficiency of humanitarian response activities. This technique encompasses a wide spectrum of crisis-related information, including informative vs non-informative content, humanitarian categories, and damage severity rating. Previous studies have demonstrated the capabilities of learning models, in categorizing crisis related information [1]. Other investigation has gone into the classification of humanitarian issues, such as recognizing postings relating to injured individuals, infrastructure damage, and rescue efforts [2]. Additionally, damage severity evaluation has been an important field, with models meant to quantify the amount of harm, classifying it as severe, mild, or minimal [3]. The study conducted by Wang and others in 2021 focuses on the recognition and localization of events, in social media data related to crises. They propose methods to utilize platforms such, as Twitter and Facebook to detect and pinpoint incidents thereby aiding in disaster response efforts and facilitating research [4]. The work by [5] at COLING addresses the integration of textual and visual data from social media for crisis event analysis, increasing the field's understanding of multimodal fusion methodologies. The work by [6] studies the use of social media data in crisis informatics, particularly in cross-lingual circumstances. It analyzes obstacles and potential in collecting relevant insights from multilingual social media sources during emergencies. The research by [7] presents a valuable curated Twitter dataset for crisis management researce and this resource promotes real time information availability during crises. Our method extends upon this earlier studies by presenting a holistic solution that integrates several machine learning models to enable robust and accurate text classification inside the CrisisMMD dataset. The CrisisMMD dataset [8] is a dynamic and comprehensive resource developed for various crisis-related

979-8-3503-4985-6/24/$31.00 ©2024 IEEE

applications. The CrisisMMD dataset is a great resource for scholars and corporations working on crisis management, analysis, and decision-making. It classifies crisis material into ”Informative” and ”Not Informative” categories, categorizes crisis-related information into eight humanitarian categories, and assesses the extent of harm caused by a catastrophe and the strategy uses machine learning and deep learning techniques including BERT, SVM, Logistic Regression, XGBoost, Stochastic Gradient Descent, Complement and Multinomial Naive Bayes algorithms, Decision Trees. Ensemble voting integrates predictions from many models, resulting in a robust and accurate categorization system for managing different crisis data.

II. Literature Review

This literature review investigates current studies on multimodal approaches for categorizing disaster-related tweets in humanitarian tasks. The focus is on learning with supervision, extracting features, choosing appropriate models, and combining data sets. The examination provides insights into the achievements made by researchers and elements influencing their outcomes. In their work [9] obtained an outstanding accuracy of 0.878 in categorizing tweets into multiple types, underlining the potential of utilizing diverse data sources to boost the accuracy of humanitarian job classification [10] examined the interaction between text and visuals in categorizing disaster-related tweets, reaching micro and macro F1-scores of 0.95 and 0.94, respectively. In their work [11] utilized a two-step method, consolidating categories and categorizing data, underlining the crucial relevance of proper data fusion approaches. In their work [12] reduced humanitarian categories into five subcategories and explored various feature extraction and fusion algorithms. In their work [13] focuses on unifying categories to produce a more manageable categorization work.

In their work [8] adopted a unimodal approach, merging categories exclusively within text data. Their use of CNN for text led to moderate results in terms of accuracy and F1-score highlighting the challenges and limitations of unimodal classification. Several other studies, including those by [14], [15], and [16], explored diverse strategies for merging categories and leveraging multimodal features. Their research underscored the importance of consistent or inconsistent labeling in influencing classification outcomes. The authors [17] broadened the scope by enhancing large pre trained unimodal model techniques with containing multimodal information, ultimately achieving robust F1-scores. Their exploration of various models emphasized the value of multimodal approaches in disaster tweet classification. Lastly, the authors [18] researched on multimodal approaches for catastrophe tweet classification has made tremendous progress, with methodology, model choices, and data fusion strategies playing crucial roles. Factors like data merging and label consistency effect classification accuracy, F1-score, and model resilience. Future study should examine creative techniques to boost accuracy and efficiency, helping humanitarian and disaster management activities. Social media analytics have grown in relevance [19]. To extract and categorize tweets related to resource needs and availability during disasters, a variety of techniques, including deep learning, clustering, and natural language processing methods like BERT, distilBERT, RF, LR, NB, and SVM classification models, have been employed [20] [21]. Furthermore, disaster-affected communities have utilized platforms like Twitter to convey their thoughts, coordinate their actions, and develop collective responses to disasters [22]. Computational methods have been applied to model these community responses on social media, unveiling typical response patterns across diverse domains also notably the abstracts provided do not specifically address the use of multimodal deep learning for disaster response in the analysis and assessment of social media data and information.

III. Different Machine Learning and Deep Learning Algorithms

A. Ensemble Voting:

Ensemble voting represents a sophisticated approach within machine learning, where predictions from diverse models collaboratively contribute to formulating a final prediction. This technique harnesses the collective power of various models, transforming a set of weaker learners into a robust, high-performing entity. Prominent ensemble strategies, such as Bagging exemplified by Random Forest and Boosting like AdaBoost, strive to enhance overall model performance by mitigating overfitting and bolstering resilience.

B. Support Vector Machine:

Support Vector Machines, a category of supervised learning algorithms, find application in both classification and regression tasks. SVM endeavors to discover an optimal hyperplane that effectively segregates data points into distinct classes, maximizing the margin between them. Its efficacy extends to high dimensional spaces and its adaptability to non-linear relationships is achieved through the integration of kernel functions.

C. Logistic Regression Algorithm:

Logistic Regression technique is a fundamental and widely adopted algorithm that facilitates the prediction of a continuous output variable based on one or more predictor variables and also it presupposes a linear association between input features and the target variable fitting a line to the data to derive all the predictions.

D. XGBoost (Extreme Gradient Boosting):

XGBoost stands out as a renowned gradient boosting algorithm lauded for its speed and performance. This technique constructs a sequence of decision trees sequentially, with each subsequent tree rectifying errors from its predecessor. Demonstrating effectiveness across diverse machine learning tasks and competitions, XGBoost has become a stalwart in the field.

E. Stochastic Gradient Descent:

Stochastic Gradient Descent shortly termed as SGD serves as a pivotal optimization algorithm particularly prevalent in training machine learning models, especially within the realm of deep learning. It iteratively fine-tunes model parameters to minimize a specified loss function, proving especially advantageous in handling extensive datasets.

F. Complement Naive Bayes:

Complement Naive Bayes represents a nuanced iteration of the Naive Bayes algorithm tailored for imbalanced text classification tasks, where specific classes face underrepresentation. This variant rectifies the class distribution by factoring in the complementary probabilities of features associated with each class.

G. Multinomial Naive Bayes:

Multinomial Naive Bayes constitutes another offshoot of the Naive Bayes algorithm, apt for text classification scenarios where discrete features quantify the word frequency in a document. Its prevalence is notable in various natural language processing applications.

H. Decision Tree:

Decision Tree, a predictive modeling algorithm, adopts a tree-like structure of decisions to facilitate predictions. Employing recursive data splitting based on features, this algorithm generates a tree structure where each leaf node corresponds to a class label or numerical value.

I. BERT:

Bidirectional Encoder Representations from Transformers shortly termed as BERT is an epitome of cutting edge deep learning architecture rooted in transformers, is purpose-built for tasks within natural language processing. Uniquely capable of comprehending contextual relationships bidirectionally, BERT captures intricate semantic meanings in language also its profound impact is evident in achieving exceptional results across tasks such as text classification, named entity recognition, and question answering.

IV. Methodology

We have employed and used a diversified methodology for the textual classification of humanitarian task of the CrisisMMD dataset. The detailed steps are explained below and the figure for the methodoly is also shown below in Figure 1

A. Data Collection and Preprocessing

Data Collection: The CrisisMMD dataset encompasses textual data pertaining to crisis situations, including natural disasters and humanitarian events. It's imperative to specify the dataset's source, whether it was acquired from a specific website, social media platforms, or other channels. Moreover, if manual labeling was conducted, it's important to furnish information about the labeling process and the human annotators involved. This transparency enhances the reproducibility of the research.

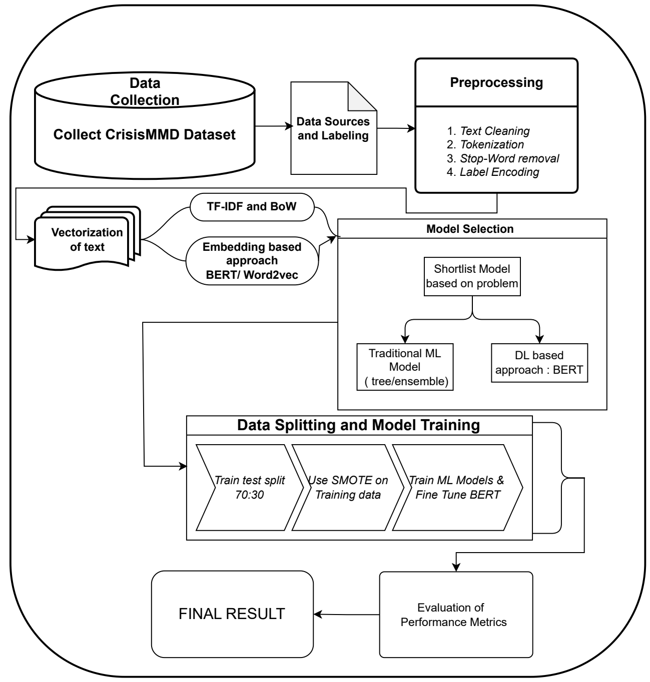

Fig. 1. Methodology

Data Preprocessing: Data preprocessing plays a pivotal role in our research, serving as the cornerstone for enabling effective text classification. Within this stage, we undertake a series of essential measures to guarantee the cleanliness, structure, and analysis-ready nature of the textual data.

Text Cleansing: Textual data often comes with noise, comprising special characters, punctuation, and irrelevant symbols that do not contribute to the model's efficacy. Text cleansing encompasses the removal or replacement of these elements, ensuring that the text maintains a consistent and usable format for analysis.

Tokenization: Tokenization entails the process of breaking down text into smaller units, whether they are words or subwords, depending on the chosen tokenization method. This phase readies the text for further analysis, enabling the model to operate on individual units effectively.

Elimination of Stop Words: Stop words, such as "the," "and," and "in," are commonly occurring words that may lack significant meaning in the context of text classification. Their removal aids in reducing noise and enhances the model's efficiency.

Label Encoding: In multi-class classification, it is essential to transform text labels into numerical values, assigning a unique numerical code to each class. This encoding simplifies the training and evaluation of the model.

B. Vectorization Techniques:

Term Frequency-Inverse Document Frequency (TF-IDF): TF-IDF is a technique that quantifies the significance of a word within a document relative to a corpus of documents. It provides a numerical representation for each term in the text, capturing not only word frequency but also their relevance to the specific document.

Bag of Words (BoW): BoW is a method for representing text as a set of word frequencies, disregarding the word order and context. It generates a sparse numerical vector, with each dimension representing a unique word, and the values signifying word frequencies within the text.

Word2Vec: Word2Vec serves as an embedding method, depicting words within a continuous vector space by encapsulating semantic connections rooted in their distributional patterns across text and also it postulates that words sharing similar meanings tend to manifest in akin contexts. Executing through two principal architectures, CBOW and Skip-gram, Word2Vec produces compact vector portrayals for words, positioning akin meanings proximately within the vector space. Commonly applied in endeavors such as similarity assessment and document grouping, Word2Vec furnishes intricate semantic associations among words, outperforming conventional approaches like the Bag of Words technique.

C. Model Selection and Fine-Tuning:

Model Selection: This phase entails the selection of machine learning models for text classification. For instance, it mentions the inclusion of Fine-Tuned BERT, but it's also possible to elaborate on the selection process for other models utilized in the study.

Hyperparameter Tuning: Fine-tuning machine learning models involves optimizing hyperparameters to enhance their performance. This process encompasses the selection of the appropriate learning rate, batch size, regularization parameters, and other model-specific settings to achieve optimal results.

Fine-Tuning BERT: In the context of Fine-Tuned BERT, it's vital to clarify that BERT is a pre-trained model that has acquired contextual language representations from extensive text data. Fine-tuning involves deploying this pre-trained model and adding a classification layer tailored to the specific task at hand. The particular approach to fine-tuning BERT may involve loading pre-trained weights, freezing some layers, or adjusting the architecture to suit the classification task.

D. Ensemble Voting:

Ensemble voting is a method that amalgamates predictions from multiple machine learning models to enhance overall accuracy and robustness. It is important to elucidate how the ensemble technique is applied and why it is selected. Describe the ensemble strategy, which, in this instance, is majority voting, where the class with the most votes serves as the final prediction.

E. Model Training and Evaluation:

Data Split: The dataset is divided into training, validation, and testing sets to ensure equitable model performance assessment and generalization to unseen data.

Model Training: Each chosen model is trained on the training data, using the corresponding vectorization technique (e.g., TF-IDF or BoW).

Model Evaluation: Model performance is gauged using diverse performance metrics, including:
Accuracy: Measuring the proportion of correctly classified messages.
Precision: Quantifying the model's ability to avoid false positives, especially crucial in humanitarian tasks to prevent erroneous classifications.
Recall: Assessing the model's ability to identify true positives, ensuring that pertinent messages are correctly categorized.
F1-score: The F1-score strikes a balance between precision and recall, offering a comprehensive performance metric.
Confusion Matrices: Confusion matrices aid in visualizing class-specific performance, displaying true positives, false positives, true negatives, and false negatives for each class.
Overfitting and Generalization: Overfitting is discussed to highlight the potential issue of models learning the training data excessively, potentially hindering their ability to generalize to unseen data. Generalization is crucial in real-world applications, such as humanitarian tasks, to ensure that the model performs effectively across various scenarios.

V. Results and Discussion

The examination of textual categorization for humanitarian tasks using the Crisis MMD dataset revealed notable trends in the performance of machine learning models also few various methodologies including Bag of Words, TF-IDF, Word2Vec and BERT were employed and the ensemble voting model consistently outperformed other models across multiple evaluation metrics highlighting the effectiveness of combining different approaches to improve predictive accuracy, precision, recall and F1-score. The study, centering on the textual classification of humanitarian tasks utilizing the Crisis MMD dataset, encompassed eight distinct categories: *(1) Infrastructure and utility damage, (2) Vehicle damage, (3) Rescue, volunteering, or donation efforts, (4) Affected individuals, (5) Injured or dead people, (6) Missing or found people, (7) Other relevant information, and (8) Not humanitarian.* The study covered eight distinct categories within the Crisis MMD dataset and also the results for Bag of Words, TF-IDF, Word2Vec and BERT methodologies were presented in Tables I, II, III and IV respectively. The ensemble voting model surpassed other models in the Bag of Words scenario, obtaining an accuracy of 0.736. The Support Vector Machine (SVM) demonstrated commendable performance. The TF-IDF technique retained its dominance with an accuracy of 0.729. Bag of Words generally beat TF-IDF across several models, underscoring its usefulness in textual classification for humanitarian purposes. The integration of BERT yielded remarkable outcomes, surpassing the ensemble voting model. SVM displayed durability in both cases, while Naive Bayes models showed simplicity and efficiency. The success of the ensemble voting model demonstrates its potential for real-world implementation in humanitarian aid applications. The research on humanitarian text classification utilizing the Crisis MMD dataset demonstrates encouraging outcomes and the ensemble voting model which combines numerous machine learning algorithms surpasses state-of-the-art models in terms of accuracy, precision, recall and F1-score and by merging varied models, it achieves better classification accuracy, surpassing individual models and methodologies. The implementation of this collective strategy could be feasible for the deployment in real-life scenarios during emergency crises. Through this study we also demonstrated the efficiency of the traditional NLP methods as well as newer Deep learning based Transformer model like BERT for the sppedy reponse to humanitarian situations. The study compares the effectiveness of an ensemble voting model and Mariham Rezk et al.'s [15] multimodal approach in identifying crises from social media feeds. Both studies use the CrisisMMD dataset and apply similar sets of eight classes. The ensemble voting model achieves competitive accuracy, precision, recall, and a weighted F1-score. The integration of BERT into the examination delivers exceptional results, displaying an accuracy of 0.7651 across precision, recall, and F1-score. BERT stands up as the best successful model for categorizing textual information linked to humanitarian tasks in this specific case. Both findings illustrate the efficacy of multimodal techniques, incorporating visual and textual cues, for increased performance in humanitarian text categorization. The shared emphasis on exploiting sophisticated techniques like BERT emphasizes the usefulness of pre-trained contextual embeddings in increasing model performance for crisis-related tasks as illustrated in Figure 2. Further study should develop hybrid models that integrate the characteristics of several approaches to boost overall performance and flexibility across diverse crisis-related datasets. A full explanatory graph for comparison between oour technique and the state-of-the-art models and some previous efforts are presented below in Figure 3.

TABLE I
RESULTS FOR BAG OF WORDS

Model Name	Accuracy	Precision	Recall	F1-Score
XG Boost	69.70	0.6973	0.6970	0.6926
Ensemble Voting	68.98	0.6918	0.6898	0.6881
SVM	65.12	0.6659	0.6512	0.6521
Logistic Regression	65.01	0.6708	0.6501	0.6573
Multinomial Naive Bayes	64.22	0.6558	0.6422	0.6424
Complement Naive Bayes	63.34	0.6602	0.6334	0.6394
Random Forest	55.87	0.6363	0.5587	0.5708
Decision Tree	48.02	0.5518	0.4802	0.5006

TABLE II
RESULTS FOR TF-IDF

Model Name	Accuracy	Precision	Recall	F1-Score
Ensemble Voting	70.88	0.7077	0.7088	0.7068
Logistic Regression	69.95	0.6999	0.6995	0.6992
XG Boost	69.60	0.6953	0.6960	0.6925
SVM	69.55	0.6951	0.6955	0.6940
Random Forest	66.95	0.6932	0.6695	0.6568
Multinomial Naive Bayes	60.02	0.6735	0.6002	0.6152
Decision Tree	58.60	0.5907	0.5860	0.5870
Complement Naive Bayes	54.60	0.6533	0.5460	0.5718

TABLE III
RESULTS FOR WORD2VEC

Model Name	Accuracy	Precision	Recall	F1-Score
Ensemble Voting	52.50	0.5172	0.5250	0.5182
XG Boost	52.20	0.5141	0.5220	0.5157
Random Forest	50.53	0.5005	0.5053	0.5004
Decision Tree	37.71	0.4215	0.3771	0.3934
SVM	30.99	0.4172	0.3099	0.3140
Logistic Regression	27.13	0.4357	0.2713	0.3134

TABLE IV
RESULTS FOR BERT

Model Name	Accuracy	Precision	Recall	F1-Score
Base uncased	77.13	0.7677	0.7713	0.7686

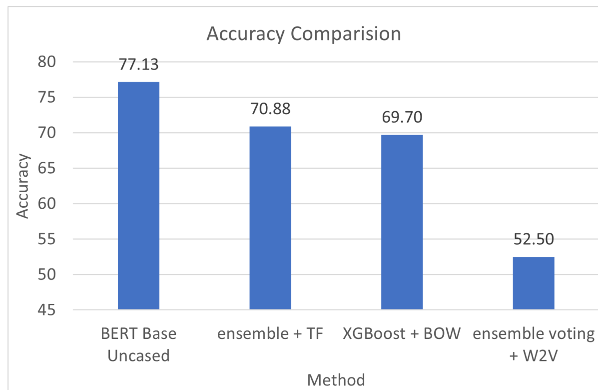

Fig. 2. Comparison of Results for all methodologies

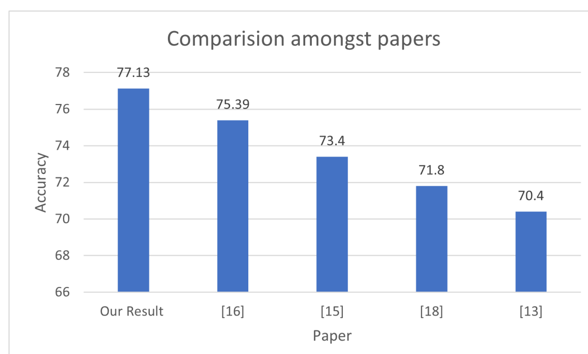

Fig. 3. Result Comparsion with State-of-the-art models and few previous works

VI. CONCLUSION

The study investigated the performance of machine learning models in categorizing humanitarian tasks using the Crisis MMD dataset. The ensemble voting model exhibited higher performance, emphasizing the benefits of mixing varied models for improved prediction accuracy, precision, recall, and F1-score. The Bag of Words technique was preferred over TF-IDF, showing a frequency-based representation of text captures more meaningful information. The Support Vector Machine (SVM) proved durability, whereas Naive Bayes models were simple and efficient. However, the Decision Tree model was outperformed by more sophisticated models, showing possible limits in capturing nuanced relationships within text data. The ensemble voting model emerged as a promising contender for real-world deployment in humanitarian relief applications, helping to more accurate and trustworthy textual information classification during crises. The next study will focus on fine-tuning BERT on domain-specific data exploring the interpretability of BERT's judgments expanding the dataset to include more diverse humanitarian tasks, exploring additional transformer-based models, and engaging with domain experts.

REFERENCES

[1] M. Johnson, D. Murthy, B. W. Robertson, W. R. Smith, and K. K. Stephens, "Moving emergency response forward: Leveraging machine-learning classification of disaster-related images posted on social media," *Journal of Management Information Systems*, vol. 40, no. 1, pp. 163–182, 2023.

[2] V. Linardos, M. Drakaki, P. Tzionas, and Y. L. Karnavas, "Machine learning in disaster management: recent developments in methods and applications," *Machine Learning and Knowledge Extraction*, vol. 4, no. 2, 2022.

[3] D. T. Nguyen, F. Ofli, M. Imran, and P. Mitra, "Damage assessment from social media imagery data during disasters," in *Proceedings of the 2017 IEEE/ACM international conference on advances in social networks analysis and mining 2017*, 2017, pp. 569–576.

[4] M. S. Mredula, N. Dey, M. S. Rahman, I. Mahmud, and Y.-Z. Cho, "A review on the trends in event detection by analyzing social media platforms' data," *Sensors*, vol. 22, no. 12, p. 4531, 2022.

[5] T. Fan, H. Wang, P. Wu, C. Ling, and M. T. Ahvanooey, "Multimodal sentiment analysis for social media contents during public emergencies," *Journal of Data and Information Science*.

[6] P. Khare, G. Burel, D. Maynard, and H. Alani, "Cross-lingual classification of crisis data," in *The Semantic Web–ISWC 2018: 17th International Semantic Web Conference, Monterey, CA, USA, October 8–12, 2018, Proceedings, Part I 17*. Springer, 2018, pp. 617–633.

[7] F. Alam, F. Ofli, and M. Imran, "Crisismmd: Multimodal twitter datasets from natural disasters," in *Proceedings of the international AAAI conference on web and social media*, vol. 12, no. 1, 2018.

[8] F. Ofli, F. Alam, and M. Imran, "Analysis of social media data using multimodal deep learning for disaster response," *arXiv preprint arXiv:2004.11838*, 2020.

[9] I. Sirbu, T. Sosea, C. Caragea, D. Caragea, and T. Rebedea, "Multimodal semi-supervised learning for disaster tweet classification," in *Proceedings of the 29th International Conference on Computational Linguistics*, 2022, pp. 2711–2723.

[10] T. Sosea, I. Sirbu, C. Caragea, D. Caragea, and T. Rebedea, "Using the image-text relationship to improve multimodal disaster tweet classification," in *The 18th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2021)*, 2021.

[11] Z. Zou, H. Gan, Q. Huang, T. Cai, and K. Cao, "Disaster image classification by fusing multimodal social media data," *ISPRS International Journal of Geo-Information*, vol. 10, no. 10, p. 636, 2021.

[12] R. Pranesh, "Exploring multimodal features and fusion strategies for analyzing disaster tweets," in *Proceedings of the Eighth Workshop on Noisy User-generated Text (W-NUT 2022)*, 2022, pp. 62–68.

[13] S. Bidari, "Categorization of disaster related tweets using multimodal approach," Ph.D. dissertation, Pulchowk Campus, 2021.

[14] M. Abavisani, L. Wu, S. Hu, J. Tetreault, and A. Jaimes, "Multimodal categorization of crisis events in social media," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 14 679–14 689.

[15] M. Rezk, N. Elmadany, R. K. Hamad, and E. F. Badran, "Categorizing crises from social media feeds via multimodal channel attention," *IEEE Access*, 2023.

[16] S. Kotha, S. Haridasan, A. Rattani, A. Bowen, G. Rimmington, and A. Dutta, "Multimodal combination of text and image tweets for disaster response assessment." International Workshop on Data-driven Resilience Research, 2022.

[17] T. Liang, G. Lin, M. Wan, T. Li, G. Ma, and F. Lv, "Expanding large pre-trained unimodal models with multimodal information injection for image-text multimodal classification," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 15 492–15 501.

[18] G. Verma, V. Vinay, R. A. Rossi, and S. Kumar, "Robustness of fusion-based multimodal classifiers to cross-modal content dilutions," *arXiv preprint arXiv:2211.02646*, 2022.

[19] M. Karimiziarani, "Social media analytics in disaster response: A comprehensive review," *arXiv preprint arXiv:2307.04046*, 2023.

[20] S. Kamoji, N. D. Jayaprakash, C. B. Biju, and G. A. Benoy, "Data analytics for disaster management response using sentiments," in *2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS)*. IEEE, 2023, pp. 1690–1695.

[21] D. Maldeniya, M. De Choudhury, D. Garcia, and D. M. Romero, "Pulling through together: social media response trajectories in disaster-stricken communities," *Journal of Computational Social Science*, pp. 1–52, 2023.

[22] R. Lamsal and T. V. Kumar, "Twitter-based disaster response using machine learning," *International Journal of Social Ecology and Sustainable Development (IJSESD)*, vol. 14, no. 1, pp. 1–18, 2023.