2025 IEEE International Conference on Information Reuse and Integration and Data Science (IRI) | 979-8-3315-9944-7/25/$31.00 ©2025 IEEE | DOI: 10.1109/IRI66576.2025.00019

Media Impact Index for Disaster Vulnerability Assessment: A Thematic Classification and Vulnerability Indexing Framework

Jainil Anilkumar Patel*, Mohammadreza Akbari Lor*, Shu-Ching Chen†, *Fellow, IEEE*,
Mei-Ling Shyu*, *Fellow, IEEE*, Steven Luis‡
*School of Science and Engineering, University of Missouri-Kansas City, MO, USA
†Data Science and Analytics Innovation Center, University of Missouri-Kansas City, MO, USA
‡Knight Foundation School of Computing and Information Sciences, Florida International University, FL, USA
Emails: {japmyy, ma7fy, s.chen, shyum}@umkc.edu, luiss@fiu.edu

***Abstract*—This paper proposes a data-driven framework for quantifying disaster vulnerability using social media analytics, repurposing a previously collected Twitter dataset originally intended for evacuation behavior analysis. After refining the dataset to isolate signals of distress and need, a category based classification strategy is introduced in which thematic dictionaries guide the grouping of Tweets based on the semantic similarity of their embeddings. Focusing on Hurricane Dorian, a compound disaster during the COVID-19 pandemic characterized by high distress and negative sentiment, a weighted amplification factor is incorporated that prioritizes Tweet categories based on the immediacy of impact on human life, while normalizing by Tweet volume and population density. The resulting *Media Impact Index (MII)* is calculated at the Census Block Group (CBG) level for the United States. To demonstrate the cross-cultural flexibility of the pipeline, the same methodology is applied to Typhoon Hagibis in Japan, with a comparable vulnerability index generated at the district level. The findings suggest that the proposed framework can provide emergency management agencies with a scalable and adaptable tool for identifying and prioritizing vulnerable regions in diverse types of disasters and sociocultural contexts.**

***Index Terms*—Social Media Analysis, Natural Language Processing, Vulnerability Index, Semantic Similarity**

I. Introduction

Natural disasters frequently evolve into compound events, extending impacts beyond the primary affected areas and creating secondary vulnerability zones that traditional assessment methods often miss [1]. Although physical damage evaluation identifies immediate impact zones, it frequently overlooks regions experiencing resource shortages, infrastructure damage, and socioeconomic disruption despite minimal physical damage [2] [3] [4].

Social networks have become a critical communication channel during disasters, with affected people sharing updates in real-time, seeking help, and expressing emotional responses to evolving conditions [5]. This digital footprint offers unique opportunities to detect emerging vulnerability patterns, particularly in secondary impact zones. Processing social media data presents significant challenges - information velocity, volume, and veracity issues complicate the extraction of actionable intelligence [6]. Content geolocation remains particularly challenging when mapping disaster impacts spatially [7]. However, sentiment analysis of disaster-related content provides valuable insights into community conditions that might otherwise remain undetected [8].

This study presents a generalizable framework for deriving spatial vulnerability indices from geolocated social media data. While this approach initially is applied to a US-based dataset for Hurricane Dorian, its application is extended to Typhoon Hagibis in Japan, demonstrating its adaptability across both geographic and cultural contexts. Key contributions include:

1) A multi-stage filtering and cleaning pipeline for extracting disaster-relevant social media content;
2) A thematic categorization mechanism using embedding-based semantic similarity and curated dictionaries;
3) A sentiment-driven amplification metric named Media Impact Index (MII), normalized by message frequency and local population density.

Together, these components enable the generation of vulnerability indices that are sensitive not only to content sentiment, but also to thematic urgency and demographic context. The framework provides a scalable tool for emergency planners and humanitarian agencies seeking timely and culturally adaptive insights into community needs during crises.

The rest of this paper is organized as follows. Section II reviews related work. The proposed framework is introduced in Section III, and the experimental results are presented in Section IV. Finally, Section V concludes this paper.

II. Related Work

Recent advances in disaster Tweet classification have demonstrated the viability of machine learning for vulnerability assessment, although significant gaps remain in cross-cultural analysis and real-time processing. Existing approaches primarily focus on three key areas:

Disaster Relevance Classification: Baseline methods using SVM with PMI-weighted n-grams were established for the identification of hurricane-related Tweets [9]. Subsequent work in [10] demonstrated BERT embeddings' superiority

2835-5776/25/$31.00 ©2025 IEEE
DOI 10.1109/IRI66576.2025.00019

over traditional word embeddings (GloVe, fastText) for disaster Tweet classification, achieving 83% accuracy through contextual semantic understanding.

Real-Time Processing: IIASA's social media intelligence tool [11] enables near real-time monitoring through multi-platform data integration and adaptive machine learning pipelines. In [12], it was advanced through multimodal classification of 74K Tweets into six humanitarian categories, although with hourly update latency.

Cross-Cultural Adaptation: Recent work has established robust frameworks for multilingual disaster analysis. [13] achieved 82% F1-score for cross-lingual classification across 10 languages using Multilingual BERT with Manifold Mixup regularization, demonstrating zero-shot generalization to unseen languages. This approach effectively handles morphological variations in Japanese and regional dialects in English through shared semantic spaces. Hybrid models that combined CNNs with self-training strategies were developed, improving the classification accuracy by 14% for low-resource languages through iterative domain adaptation [14].

Information Integration and Multimodal Analytics: An early architecture for a Business Continuity Information Network (BCIN) was developed, emphasizing the importance of semantic interoperability and distributed systems for rapid disaster recovery [3]. Their framework laid the foundational work for integrating heterogeneous disaster-related information sources. A multimodal deep learning approach was developed using Multiple Correspondence Analysis to fuse visual, textual, and metadata features for disaster event classification, showcasing the value of cross-modal information fusion to improve disaster response accuracy [15].

III. Methodology

A. Social Media Data Collection Methodology

Disaster-related social media data from two primary sources was utilized to analyze public sentiment during hurricanes:

Core Dataset: Florida Hurricane Tweets: The primary analysis has around 40,000 geotagged Twitter posts sourced from a previously curated dataset used for now-casting mobility during disaster [16]. The dataset includes:

- Timestamped posts from hurricane-affected regions;
- Geolocation metadata for spatial mapping;
- Engagement indicators such as reTweets, likes, and replies;
- Hashtags and Tweet content in English.

Supplementary Japanese Dataset: A supplementary dataset consisting of 2.98 million Japanese-language Tweets was incorporated to provide comparative linguistic and cultural insights. Key distinctions in processing included:

- Language-specific sentiment analysis models;
- Language-specific embedding generation for classification;
- Culturally contextualized keyword sets;
- different source of population data.

B. Extensions and Challenges

Although this study focuses on Twitter data, which conveniently provides geolocation information, the proposed framework can be extended to other social media platforms. One key challenge, however, is that platforms other than Twitter do not readily supply geolocation metadata. Researchers must rely on text parsing and natural language processing (NLP) techniques to extract location cues and map them to standardized geographic identifiers such as ZIP codes. For Hurricane Dorian, readily available geolocation data was leveraged, allowing for a more precise regional analysis.

C. Data Cleaning and Preprocessing

A systematic cleaning and preprocessing protocol was employed to ensure the integrity and quality of the textual data for analysis. This process encompassed several key steps:

1) **Removal of ReTweet Indicators and Mentions:** ReTweet markers and user mentions were eliminated to focus on the primary content of the Tweets.
2) **Text Normalization:** The text was converted to lowercase, and Unicode inconsistencies were rectified to standardize the dataset.
3) **Elimination of URLs and Emojis:** All URLs and emojis were removed to prevent potential noise in the data.
4) **ASCII Character Retention:** Non-ASCII characters were discarded to maintain consistency in text encoding.
5) **Length Filtering:** Tweets containing fewer than four words after cleaning were excluded to ensure sufficient contextual information for analysis.

This structured data cleaning approach enhances the reliability of subsequent analyses by providing a consistent and high-quality textual dataset.

D. Thematic Category Construction and Classification

Semantic categories that capture various disaster-related themes are first constructed. These categories are defined by domain-specific terms generated via large language model (LLM) prompting with human filtering. For instance, the *fatalities* category includes terms such as *died*, *fatalities*, *body count*, *missing and presumed dead*, and *memorial*, while the *damage* category contains terms like *destroyed*, *devastated*, *rubble*, *reconstruction*, and *structural damage*. The full list of categories is given in III-G

These carefully curated categories not only capture the semantic diversity of disaster events but also serve as the basis for downstream analyses.

E. Filtering and Classification

Since the original dataset contains a large number of mobility-related Tweets that are not necessarily relevant to disaster events (e.g., promotional or generic content), a tiered filtering approach was employed to retain only disaster-relevant posts. Initially, Tweets were evaluated using raw keyword matching against the category terms defined in III-D. If no match was found, lemmatization and stemming-based

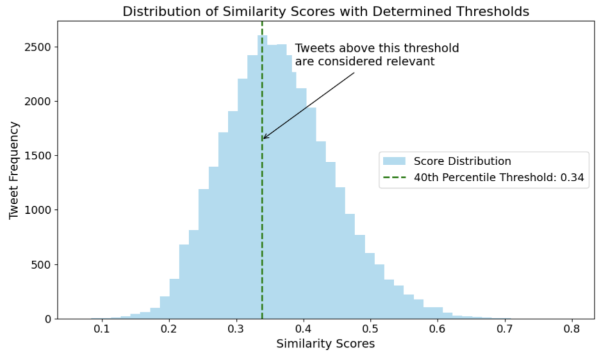


Fig. 1: Threshold to determine relevant posts

matching were applied to improve recall. Finally, for Tweets that still lacked a match, semantic similarity analysis, as described in III-E1, was performed. This multi-stage filtering was designed to prioritize direct lexical matches, which are often more reliable indicators of relevance, before relying on semantic inference.

1) Embedding-Based Unsupervised Classification Using LLMs

Unsupervised classification of the remaining Tweets is performed by leveraging high-quality embeddings generated from the `SentenceTransformer` model `all-mpnet-base-v2`[1]. Rather than relying on traditional approaches such as bag-of-words or static embeddings (e.g., GloVe) that depend on simple word matching or averaged representations, contextual embeddings are utilized for a more nuanced classification and in filtering the unrelated posts.

For each Tweet, an embedding e_t is generated. Similarly, each term in a category B is encoded to an embedding, resulting in a set $\{e_{B1}, e_{B2}, \ldots, e_{Bk}\}$ for category B. Instead of averaging keyword embeddings per category, individual term embeddings were maintained to maximize category-specific precision. The cosine similarity between the Tweet embedding and each term embedding within the category is computed. The maximum similarity score among these terms is selected as the representative similarity score for category B:

$$\text{sim}(t, B) = \max_{1 \leq i \leq k} \cos\left(e_t, e_{Bi}\right) \tag{1}$$

where $\cos(\cdot, \cdot)$ denotes the cosine similarity. k denotes the number of terms inside the category.

A similarity threshold is determined from the distribution of these scores by empirical study to filter unrelated posts. The 40th percentile of the similarity score distribution was used as the threshold (as shown in Figure 1), instead of using the mean, median, or mode. Tweets exceeding the threshold are flagged as likely disaster-related and are considered in further analysis.

To validate the category definitions and assess their suitability as proper classes for categorization, the inter- and intra-similarities of category terms need to be analyzed. For this, the cosine similarity scores taken over all pairs of terms for each category pair are averaged. In Figure 2a, a generally strong degree of semantic separability between categories is highlighted, with most inter-category similarity scores shown to be significantly lower than intra-category values. This indicates that the categories are capturing coherent and distinct semantic themes, suggesting effective term grouping.

To validate the categorizations and assess whether they effectively capture the Tweets into distinguishable categories, the inter- and intra-similarities of the categorized Tweets need to be analyzed.

Let the set of categories be defined as:

$$C = \{b_1, b_2, \ldots, b_N\} \tag{2}$$

where each category b_i contains a set of n_i embedding vectors:

$$b_i = \left\{\mathbf{e}_1^{(i)}, \mathbf{e}_2^{(i)}, \ldots, \mathbf{e}_{n_i}^{(i)}\right\}, \quad \mathbf{e}_k^{(i)} \in \mathbb{R}^d \tag{3}$$

where d is the embedding dimension. The internal similarity of a category b_i is defined as the mean of the maximum cosine similarity that each embedding has with all other embeddings in the same category:

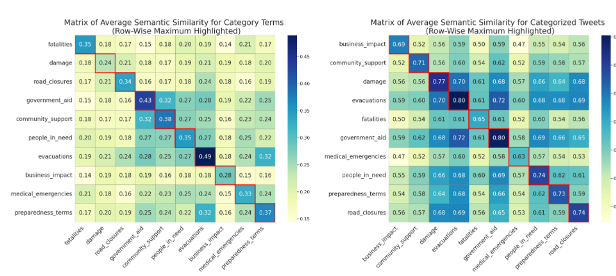
$$S_{\text{internal}}(b_i) = \frac{1}{n_i} \sum_{k=1}^{n_i} \max_{\substack{j=1 \\ j \neq k}}^{n_i} \cos(\mathbf{e}_k^{(i)}, \mathbf{e}_j^{(i)}) \tag{4}$$

The directional cross-category similarity from category b_i to b_j, where $i \neq j$, is computed as:

$$S_{\text{cross}}(b_i, b_j) = \frac{1}{n_j} \sum_{l=1}^{n_j} \max_{k=1}^{n_i} \cos(\mathbf{e}_k^{(i)}, \mathbf{e}_l^{(j)}) \tag{5}$$

The final similarity matrix $M \in \mathbb{R}^{10 \times 10}$ is constructed as:

$$M_{ij} = \begin{cases} S_{\text{internal}}(b_i) & \text{if } i = j \\ \frac{S_{\text{cross}}(b_i, b_j) + S_{\text{cross}}(b_j, b_i)}{2} & \text{if } i \neq j \end{cases} \tag{6}$$



(a) (b)

Fig. 2: The intra-category and cross-category average similarity scores with the highest in each row highlighted. (a) demonstrates the average similarity scores for category terms while (b) demonstrates the average similarity scores for categorized Tweets

As can be seen in Figure 2b, the categorization creates distinct groupings same as for category terms. Overall, the

[1]https://huggingface.co/sentence-transformers/all-mpnet-base-v2

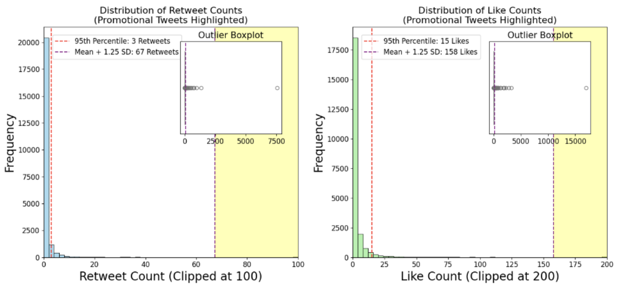


Fig. 3: Histogram of retweets and likes. The yellow-shaded regions represent Tweets above the calculated threshold, which were considered promotional.

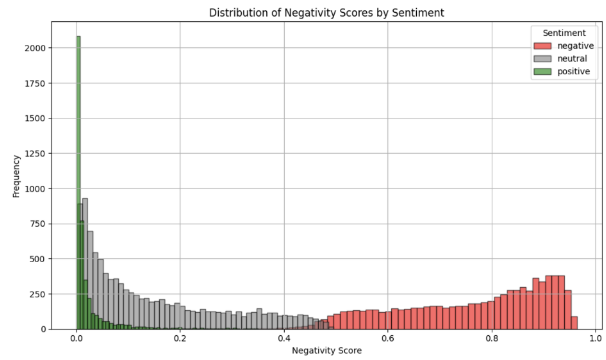


Fig. 4: Distribution of sentiment in the filtered Tweets

structure demonstrates that the categories are both internally cohesive and externally separable, a desirable property for robust unsupervised learning and downstream analysis.

2) Promotional Tweet Removal

To ensure that the dataset primarily contained informative disaster-related content, Tweets suspected to be promotional were removed. Promotional Tweets are typically characterized by abnormally high engagement metrics, such as retweet and like counts. Statistical thresholds based on the mean and standard deviation of these engagement metrics were used. Specifically, Tweets exceeding $1.25 \times z$-score in either retweets (above 67) or likes (above 158) were flagged as promotional. Figure 3 shows the distribution of retweet and like counts, clipped at the smaller counts for visualization purposes.

F. Sentiment Analysis and Vulnerability Filtering

To analyze emotional content, sentiment analysis was applied using the HuggingFace `pipeline` with the `cardiffnlp/twitter-roberta-base-sentiment-latest` model for English Tweets [17]. For each post, the sentiment label (positive, neutral, or negative) and the negativity score were recorded. Notably, even neutral Tweets often carried a non-zero negativity score, capturing subtle distress signals as shown in Figure 4.

A detailed review showed that positive Tweets were often unrelated to the hurricane, often consisting of ads or irrelevant content, while neutral and negative Tweets more accurately reflected distress, warnings, and help requests. Consequently, positive Tweets were excluded from further analysis to better capture vulnerability. Each post in the final dataset includes a unique ID, sentiment label, and negativity score.

G. Index Generation, Weighting, and Geospatial Mapping

To quantify regional vulnerability using sentiment analysis, a weighted indexing approach is developed that integrates emotional tone with geospatial population data. Each Tweet is associated with a geographic area defined by a Twitter-provided bounding box shapefile.

The dataset was first restricted to posts originating from Florida, and Census Block Group (CBG) identifiers were standardized to 12-digit strings, which were then merged with CBG-level population counts from SafeGraph [18].

To reflect the relative urgency of disaster-related content, weights are assigned to disaster categories using a predefined priority list:

```
PRIORITY_ORDER = [
    "medical_emergencies", % Highest
    "fatalities",
    "people_in_need",
    "evacuations",
    "damage",
    "road_closures",
    "business_impact",
    "preparedness_terms",
    "community_support",
    "government_aid"        % Lowest
]
```

The above ranking is based on immediacy of impact on human life and response urgency. Medical emergencies and fatalities demand immediate attention. People in need, evacuations, and damage reflect escalating but less direct threats. Road closures, business impact, and preparedness are operational or preventive. Community support and government aid are long-term and less time-sensitive.

Each category is given a weight inversely proportional to its priority rank: $w = \frac{1}{i}$. The support of each Tweet t is defined as the weight w_t corresponding to its category where i is the index of the category in `PRIORITY_ORDER`. This ensures a smooth decline in influence as urgency decreases. For each shapefile polygon s, the support values of all Tweets located within it are summed:

$$\text{support}_s = \sum_{t \in T_s} w_t \tag{7}$$

where T_s is the set of Tweets associated with shapefile s, and w_t is the category weight of Tweet t. Let C_s be the set of CBGs intersecting shapefile s, and let P_j be the population of CBG j. The adjusted support assigned to CBG j is:

$$\text{adjusted support}_j = \text{support}_s \times \frac{P_j}{\sum_{k \in \mathcal{C}_s} P_k} \tag{8}$$

This ensures the total support from each shapefile is distributed across overlapping CBGs in proportion to their population. For each CBG j, the mean negativity score from all Tweets whose shapefiles intersect that CBG is computed:

$$\text{negativity score mean}_j = \frac{1}{|T_j|} \sum_{t \in T_j} \text{negativity score}_t \tag{9}$$

where T_j is the set of Tweets whose shapefiles intersect CBG j. The Media Impact Index (MII) for each CBG j is defined as:

$$\text{MII}_j = \frac{\text{negativity score mean}_j \times \text{adjusted support}_j}{P_j} \tag{10}$$

This formulation integrates emotional intensity, disaster-topic urgency, and population-adjusted geospatial distribution to estimate vulnerability across regions with high fidelity.

IV. Experimental Results

To demonstrate the effectiveness of the proposed pipeline, two distinct events from 2019 were selected: Hurricane Dorian in Florida and Typhoon Hagibis in Tokyo. These cases were chosen to highlight the pipeline's ability to analyze diverse types of disasters and its cross-cultural applicability across different geographic and social contexts.

A. Hurricane Dorian in US

1) Social Media Analysis

As mentioned in section III-E, the categorization's first step involves direct keyword matching (raw and enhanced) to categorize a subset of Tweets. The results are as follows:

- Total Entries: 22,639
- Categorized Entries: 3,181 (14.05%) — Tweets that matched directly with some category keywords.
- Uncategorized ('Other') Entries: 19,458 (85.95%) — Tweets with no direct keyword match.

Although simple and explainable, the keyword matching method fails to classify a significant majority of the dataset, emphasizing the need for a more robust approach capable of handling semantic variability. Therefore, as discussed in section III-E1, semantic similarity is employed to further filter and categorize Tweets (depicted in Figure 5). Each Tweet that passes the thresholding defined in III-E1 is assigned:

- The most similar category based on cosine similarity.
- The corresponding similarity score.

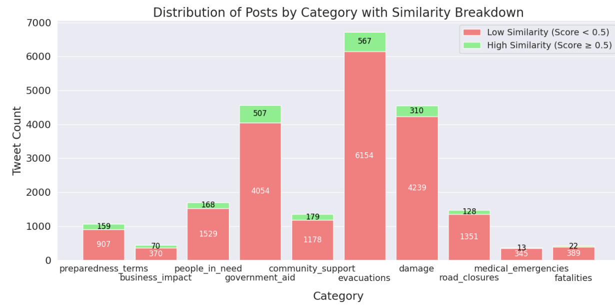


Fig. 5: Class Distribution after Categorization

2) Geographic Post Distribution

To examine the spatial distribution of disaster-related Tweets, a choropleth map was generated showing the density of geolocated posts with negative sentiment (Figure 6a). Each square represents a Twitter-derived shapefile, colored by the number of negative Tweets. Because some shapefiles span multiple CBGs, Tweets may be assigned to several CBGs, introducing redundancy but capturing geolocation uncertainty. Tweet concentrations are highest in major cities such as Tallahassee, Kissimmee, and Miami—underscoring the need for population-based normalization in the MII calculations.

3) Media Impact Index

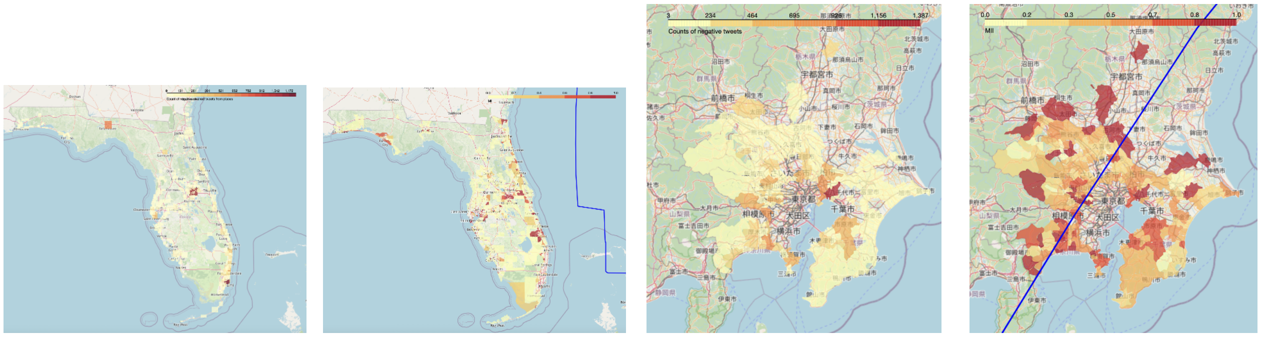
As can be seen in Figure 6b, MII is calculated for Florida regarding Hurricane Dorian. The vulnerability score is winsorized to reduce extreme vulnerability values from over shadowing other areas before normalization. While urban areas overall show more concentrated negative sentiment and display higher vulnerability due to higher Tweet count, the vulnerability is mostly concentrated in the east coast near the hurricane path, as was expected, highlighting the MII's ability to capture disaster impact through social media analysis.

B. Typhoon Hagibis in Japan

MII for Typhoon Hagibis in Japan was developed using 2.98 Million Japanese Tweets. Thematic categories were created with English keys and Japanese terms, translated via ChatGPT. To determine Tweet relevance, the `sbintuitions/sarashina-embedding-v1-1b`[2] model is used from HuggingFace for semantic similarity specific to the Japanese language. Tweets below the 40th percentile (score < 0.82) were filtered out. Promotional Tweets (N=3,652) were removed using a $1.25 \times z$-score threshold on retweets and likes. Direct word matching accounted for only 0.09% of matches, and the rest were matched semantically.

The model used for sentiment classification was `kit-nlp/bert-base-japanese-sentiment-irony`[3]. Only negative Tweets (N=72,216) were retained for vulnerability analysis (Figure 6c). District shapefiles were sourced from HDX[4], and population per district was

[2]https://huggingface.co/sbintuitions/sarashina-embedding-v1-1b
[3]https://huggingface.co/kit-nlp/bert-base-japanese-sentiment-irony
[4]https://data.humdata.org/dataset/cod-xa-jpn



(a) Dorian Negative Count (b) Dorian MII (c) Hagibis Negative Count (d) Hagibis MII

Fig. 6: MII displayed for Hurricane Dorian and Typhoon Hagibis in 2019. (a) The negative Tweet counts for Hurricane Dorian. (b) The calculated MII for it along with the hurricane path depicted with a blue line. (c) The negative Tweet counts for Typhoon Hagibis. (d) The calculated MII for it along with the Typhoon path depicted with a blue line.

computed from census-based 1 Km^2 density data. These components were integrated in the final district-level MII (Figure 6d).

V. Conclusion

This study demonstrates the potential of social media data for deriving valuable insights that can aid in effective decision-making during disaster events. By leveraging this data, it is possible to derive additional metrics, such as sentiment trends and geographic patterns, that provide a more comprehensive picture of disaster-related conditions. Moreover, the integration of LLMs presents an effective method for automating the generation of disaster-related queries and post classification. By embedding textual outputs from LLMs, it is possible to create semantic class representations that can be compared with social media content, enabling efficient and accurate categorization of posts. This approach facilitates real-time monitoring and response during disaster events, allowing for better-informed decisions. Additionally, the analysis of Tweet volume relative to population density reveals important indicators of distress in specific regions. Higher volumes of negative Tweets in proportion to the population within certain areas (CBGs) can signal acute distress. Factors contributing to this distress can be derived from the distribution of categories in each area. Identifying these indicators can help prioritize response efforts and resources where they are most needed.

Acknowledgment

This work is partially supported by NSF CNS-2301552.

References

[1] United Nations Office for Disaster Risk Reduction, "Hazard definition & classification review: Technical report," UNDRR, Geneva, Switzerland, Tech. Rep., 2020.

[2] S.-C. Chen, M. Chen, N. Zhao, S. Hamid, K. Chatterjee, and M. Armella, "Florida public hurricane loss model: Research in multi-disciplinary system integration assisting government policy making," *Government Information Quarterly*, vol. 26, no. 2, pp. 285–294, 2009.

[3] K. Saleem, S. Luis, Y. Deng, S.-C. Chen, V. Hristidis, and T. Li, "Towards a business continuity information network for rapid disaster recovery," in *The international conference on Digital government research*, 2008, pp. 107–116.

[4] S. E. Chang, "Socioeconomic impacts of infrastructure disruptions," *Oxford Research Encyclopedia of Natural Hazard Science*, 10 2016.

[5] T. S. Joseph E. Trainor, "Social media in emergency management: Academic perspective," in *Critical Issues in Disaster Science and Management: A Dialogue Between Researchers and Practitioners*. FEMA, 2012, pp. 349–392.

[6] M. Imran, C. Castillo, F. Diaz, and S. Vieweg, "Processing social media messages in mass emergency: A survey," *ACM Computing Surveys*, vol. 47, no. 4, Jun. 2015.

[7] Z. Wang, X. ye, and M.-H. Tsou, "Spatial, temporal, and content analysis of twitter for wildfire hazards," *Natural Hazards*, vol. 83, 08 2016.

[8] D. Erokhin and N. Komendantova, "Social media data for disaster risk management and research," *International Journal of Disaster Risk Reduction*, vol. 114, p. 104980, 2024.

[9] K. Stowe, J. Anderson, M. Palmer, L. Palen, and K. Anderson, "Improving classification of Twitter behavior during hurricane events," in *The International Workshop on Natural Language Processing for Social Media*, Jul. 2018, pp. 67–75.

[10] A. K. Chanda, "Efficacy of bert embeddings on predicting disaster from twitter data," *ArXiv*, vol. abs/2108.10698, 2021.

[11] H. Hassani, N. Komendantova, E. Rovenskaya, and M. R. Yeganegi, "Leveraging social media intelligence for disaster risk management," IIASA Blog, 2024, accessed: 2025-04-11.

[12] M. Karimiziarani and H. Moradkhani, "Social response and disaster management: Insights from twitter data assimilation on hurricane ian," *International Journal of Disaster Risk Reduction*, vol. 95, p. 103865, 2023.

[13] J. R. Chowdhury, C. Caragea, and D. Caragea, "Cross-lingual disaster-related multi-label tweet classification with manifold mixup," in *Annual Meeting of the Association for Computational Linguistics: Student Research Workshop*, Online, Jul. 2020, pp. 292–298.

[14] H. Li, D. Caragea, C. Caragea, and N. Herndon, "Disaster response aided by tweet classification with a domain adaptation approach," *Journal of Contingencies and Crisis Management*, vol. 26, no. 1, pp. 16–27, 2018.

[15] S. Pouyanfar, Y. Tao, H. Tian, S.-C. Chen, and M.-L. Shyu, "Multimodal deep learning based on multiple correspondence analysis for disaster management," *World Wide Web*, vol. 22, no. 5, pp. 1893–1911, Sep 2019.

[16] R. Jiang, Z. Wang, Y. Tao, C. Yang, X. Song, R. Shibasaki, S.-C. Chen, and M.-L. Shyu, "Learning social meta-knowledge for nowcasting human mobility in disaster," in *ACM Web Conference*, 2023, p. 2655–2665.

[17] CardiffNLP, "Twitter roberta base sentiment model," https://huggingface.co/cardiffnlp/twitter-roberta-base-sentiment-latest, 2022, accessed: 2025-04-14.

[18] SafeGraph, "Open census data," https://docs.safegraph.com/docs/open-census-data, 2022, accessed: 2025-04-13.