Sentiment Analysis for Early Warnings and Disaster Management: a Survey

Raffaele Guarasci[1][0000−0002−0106−8635]

Institute for High Performance Computing and Networking (ICAR))
National Research Council of Italy (CNR)
raffaele.guarasci@icar.cnr.it

Abstract. Sentiment Analysis has played a crucial role in determining the effectiveness of the response and recovery efforts. In this paper, we present a survey exploring the application of sentiment analysis techniques in disaster management, focusing on their role in early warnings and disaster prediction using data from social media and crowdsourcing platforms. Extracting and categorizing emotional cues from online content during crises could help to interpret public sentiment to enhance early warning systems and disaster response strategies by harnessing machine learning techniques, such as supervised classification and deep learning models. The survey reviews various methodologies and tools employed in recent studies, focusing on their contributions to real-time disaster monitoring, sentiment-based classification of disaster events, and proactive management of crises. Insights drawn from these studies underscore the potential of sentiment analysis in leveraging online data to inform decision-making and improve resilience against natural disasters and other emergent events.

Keywords: Early Warnings · Sentiment Analysis · Disaster Recovery.

1 Introduction

The enormous impact of social media and instant messaging platforms on disaster management and health emergencies is undeniable. Their utility ranges from early warnings to providing direct communication channels with real-time information and handling public panic [29, 26]. The preferential and direct information lane made possible by social media has dramatically expanded the possibilities of spreading crucial information about disaster preparedness (i.e., evacuation routes), hygiene prevention measures, or vaccination drives [19, 37]. In addition to the level of preventive communication, user-generated content also has tremendous value after the fact, providing more opportunities for authorities to be informed about the status of the disaster, the status of the situation, and the assistance needed. In addition, direct communication between users brings into communication even very distant realities that would otherwise have little visibility, allowing different entities to coordinate donations, aid, or assistance [51].

Note that this is not "passive" communication. The ability for users to provide feedback or updates not only of logistical utility but also on the emotional and mental state of the situation allows for fine-grained emergency management like never before. A relevant aspect that should not be underestimated is, in fact, mental health support, made possible even in crisis conditions through online resources and support groups [69].

Among the most widely used social networks, the central role so far has been played by Twitter (now X) by its peculiar nature of allowing the phenomenon to be observed in a continuous conversational exchange over some time. The active user engagement, combined with the compactness of tweets limited to the number of characters, leads to a social network topology in which network properties enhance and speed up the spread of information, including locations, connections, personal opinions, and emotions. Starting from these motivations, the literature has mainly focused on this platform in recent years, especially in conveying emotions [65, 28]. In particular, a trend that has proven very fruitful is to use sentiment analysis on data extracted from Twitter at various stages of disaster management [32]. Sentiment analysis meant as both the polarity analysis and emotions detection, has now been established as one of the most successful tasks in Natural Language Processing (NLP) because of its scalability across different contexts and domains [52] and its adaptability to a wide variety of approaches, ranging from rule-based ones [20] to the latest deep learning approaches using most recent Neural Language Models (NLMs) [33, 23, 24].

However, there are also many open issues related to the use of online platforms. The main issue that pervades the current landscape is that of fake news and the spread of misinformation. Therefore, techniques that use social media for disaster management must necessarily implement strategies to address the scourge of misinformation and the problems related to the privacy of sensitive data [7, 66]. Furthermore, it is a newborn field, with many open issues that require integration from different perspectives, i.e., social aspects, crisis occasions, and systemic dependencies, to correctly estimate the correlation between social media contents and disaster management.

Given these premises and the growing interest in the community in the possibilities offered by NLP and sentiment analysis to handle, predict, and prevent disaster events, this paper aims to provide a comprehensive overview of existing approaches developed so far in this field, with a particular focus on natural disaster-related events. Existing resources extracted from social media, approaches developed, and open issues yet to be addressed are presented and discussed.

The paper is organized as follows: First, a theoretical background is described in order to clarify the assumptions underlying this work. Early warning systems are introduced in the subsection 2.1. After that, a brief overview of what sentiment analysis is provided in section 2.2. Next, the classification of existing works that have dealt with the relationship between sentiment analysis with early warnings and disaster management is presented (section 2. In detail, existing resources, whether created *ad hoc* or gathered through crowdsourcing, are described in subsection 3.1;. In contrast, subsection 3.2 presents approaches developed for this task, starting from early ones relying on rule-based strategies to the newest models based on deep learning. Finally, in the section 4, a summary on the current status of research in this topic, the benefits, limitations, and possible future directions is provided.

2 Theoretical Background

In this section, early warning systems and sentiment analysis are briefly described, with a particular focus on approaches developed for data extracted from social media.

2.1 Systems for Early Warnings

Traditionally, the primary purpose of natural disaster warnings was to inform the public about the likelihood of hazards (e.g., floods, heavy rainstorms, or other weather alerts) from a scientific perspective. With the pervasive spread of the Internet over mobile networks and technological progress in weather forecasting, the role of warnings has changed, including social aspects, leading to the development of Early Warning Systems (EWSs). EWSs, unlike traditional warnings, do not focus solely on providing information; they include strategies to forecast and respond to these warnings and undertake public actions.

Following the guidelines of the United Nations Office for Disaster Risk Reduction (UNDRR)[1], EWSs are officially composed of two essential elements: end-to-end and people-centered. The end-to-end feature indicates the capability that EWSs must possess to undertake public actions. In the literature, different components have been identified [43], as shown in the figure 1. Such classification has not garnered unanimous opinions; other scholars believe that each component of the EWS should be autonomous and self-sufficient to ensure proper organization and an effective response [18].

The second feature that distinguishes EWSs from traditional systems is their people-centric orientation. Strong public involvement is a central aspect ensuring rapid response of EWSs. It shifts the focus from the "official" stance of competent authorities to the involved users and local communities, who are directly engaged in all phases of the emergency. Different studies have highlighted that only through this close synergistic collaboration between global and local players can the strategies implemented by EWSs be effective [5].

2.2 Sentiment Analysis

Sentiment analysis, also known as opinion mining [13], is a branch of NLP aimed at detecting the emotional attitude of a speaker towards a given topic. The main goal of sentiment analysis is detecting the hidden subjective expression in

[1] https//www.undrr.org/

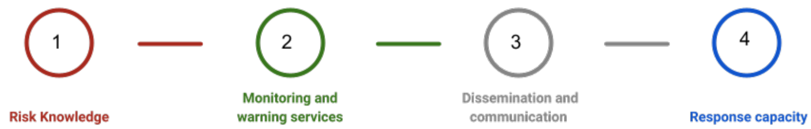

Fig. 1. Components composing end-to-end EWSs

the text. In order to detect subjective content, various features must be taken into account. This subjective expression can be formulated as "opinions" and annotated or predicted using a polarity score (usually ranging from negative -3 to positive +3) or using emotions (following emotion classifications already well-known in the literature).

It has a long history in Theoretical and Computational Linguistics; early works focus on identifying viewpoints [61], evaluating the semantic meaning of specific part-of-speech [25], categorizing degrees of subjectivity [62], and review consumer opinions [14].

Concerning the topic of this work, sentiment analysis has been used for the crisis domain, particularly using tweets as an early warning system for detecting earthquakes. In [53], an approach relying on keywords and word counts has been proposed, while in [50], machine learning algorithms have been developed to classify crisis-related tweets. Some hybrid approaches have combined official data with social media ones [12]. Other works have exploited different emotions classifications [59], while [54] has based its analysis on stylometry features

In recent years, sentiment analysis has gained significant attention in different domains, particularly healthcare[4]. There has been a growing focus on analyzing patient opinions, especially in attempting to detect early distress that could arise due to post-pandemic stress [44]. Several datasets have been released, and a wide range of models have been pre-trained or fine-tuned specifically for sentiment analysis purposes. The COVID-19 pandemic has dramatically accelerated the development of techniques and resources for sentiment analysis focused on emergency management.

For the sake of completeness, recently, techniques derived from sentiment analysis have been combined with other NLP well-known tasks such as acceptability judgments [57, 8] or information extraction [21, 40], to improve performances and evaluate the goodness of state-of-the-art LLMs[56, 41].

3 Classification of Work

In this section, resources and approaches developed for the purpose of dealing with disaster management or EWSs using sentiment analysis techniques are described. The resources are presented chronologically and follow the taxonomy most used in the literature to group disasters by types (Figure 2). Notice that this is a partial representation since it includes only the types of disasters processed through sentiment analysis on online platforms. In particular, given the preliminary nature of this work, only natural disasters will be covered since they have been the ones that have generated the most interest to date in automated analysis using sentiment analysis techniques.

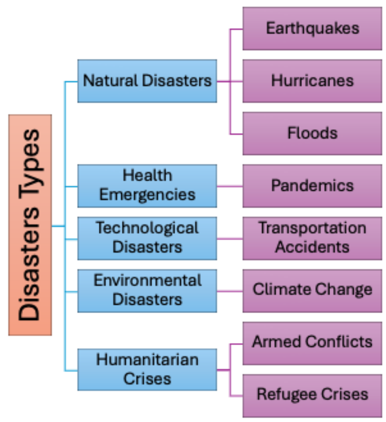

Fig. 2. Classification of disasters

3.1 Resources

Several disaster datasets have been collected to understand how people communicate on social media during disaster scenarios. In recent years, the leading source of this kind of data collection has been Twitter (now X). It is due to several reasons, internal and external. Among the internal reasons, there is the nature of instant communication of the platform, which lends itself well to describing or debating a topic in real time. Concerning external reasons, the availability of freely accessible APIs has provided a valuable way to aggregate such tweets into homogeneous datasets useful for research purposes. An overview of the resources is shown in the table 1.

Concerning natural disasters, the datasets are most likely about real-time event descriptions from those directly involved. All datasets listed below were collected from Twitter and merged into a series of collections available online [2]

Among the best-known datasets, the first resource in chronological order is the Pakistan Earthquake 2013 dataset that captures tweets during the event. A similar resource, the Chile Earthquake 2014 dataset, contains approximately

[2] \https://crisisnlp.qcri.org/

368,000 tweets, split into three subsets ready to be processed using automatic techniques. Covering extensive floods in India, the India Floods 2014 dataset comprises over 5 million tweets, offering a valuable resource for machine learning approaches since its size. A similar criterion and a division into the train, test, and dev splits were followed for the collection of datasets related to Typhoon Hagupit 2014, Nepal Earthquake 2015, and Cyclone 2015, consisting of 626,000, 4 million, and 500,000 tweets, respectively. A smaller dataset was created for the Italy Earthquake 2016, comprising over 70,000 tweets. Other resources about earthquakes are Iraq-Iran Earthquake 2017 dataset, with over 200,000 tweets, and Mexico Earthquake 2017 dataset, comprising 3.8 million tweets, is divided into training, validation, and testing data, enabling advancements in machine learning techniques for disaster scenarios. Two big datasets were collected to study hurricanes: The Hurricane Harvey 2017 dataset, with 6.6 million tweets, and the Hurricane Maria 2017 dataset, with nearly 3 million tweets, meticulously distributed to facilitate in-depth model analysis. Another natural disaster for which data were collected using Twitter is the California Wildfires 2017 dataset, which contains 455,000 tweets. This dataset provides opportunities to explore machine learning in disaster management through a balanced allocation strategy. Finally, we mention the Sri Lanka Floods 2017 dataset, with around 41,000 tweets. Despite its modest size, it is strategically partitioned to support effective model development and accurate evaluation.

Type	**Title**	**Year**	**Size**
Earthquakes	Pakistan	2013	not specified
	Chile	2014	368,000 tweets
	Nepal	2015	4 million tweets
	Italy	2016	70,000 tweets
	Iraq-Iran	2017	200,000 tweets
	Mexico		3.8 million tweets
Hurricanes Typhoons Cyclones	Typhoon Hagupit	2014	626,000 tweets
	Cyclone	2015	500,000 tweets
	Hurricane Maria	2017	3 million tweets
	Hurricane Harvey		6.6 million tweets
Floods	India	2014	5 million tweets
	Sri Lanka	2017	41,000 tweets
Wildfires	California	2017	455,000 tweets

Table 1. List of main datasets for natural disasters

Notice that, similar to what has happened in other areas of NLP, these and other datasets have been merged to create larger benchmark datasets to optimize performance in training increasingly efficient but increasingly resource-intensive NLMs. In the case of natural disasters, the most well-known aggregated datasets are CrisisNLP [3], Crisislex [46], and the multimodal resource CrisisMMD [1].

[3] http://crisisnlp.qcri.org/

3.2 Approaches

Because of the overwhelming flow of information spread during disasters on social media, it is crucial to identify proper disaster-related content quickly. Different strategies have been proposed to automate disaster discovery and classification on Twitter (X).

Early approaches were based on a set of preprocessing techniques, ranging from classical support vector machines (SVM) or k-nearest neighbors (KNN) and random forest [55]. The combination of different techniques has been another route taken by some studies. For instance, in [1], Random Forest and AdaBoost have been used together to identify disaster-related tweets for damage assessment using word features and statistics. The approach has been tested on real datasets and is showing good performance.

Currently, the typical pipeline involves machine learning and NLP approaches. Models are trained on labeled datasets, and NLP algorithms discriminate between disaster-related and non-disaster-related social media content. The significant advantage of such an approach is the ability to provide real-time information. However, several challenges arise from the informal language used on social media and the presence of emotional content.

[11, 45] were among the earliest studies underscoring the potential of processing social media data, from texts to multimedia content, to detect emergencies in urban environments. After that, other features extracted from social media have been taken into account (i.e., geotagging information and the number of interactions between users). [60] presents the Hurricane Sandy Twitter(X) Corpus, providing 6.5 million geotagged tweets for disaster research and response, supporting collaboration among practitioners and organizations. Furthermore, [30] has exploited geotagged tweets to assess social media engagement in disaster-affected regions, focusing on marginalized groups.

Machine learning algorithms' importance in managing spatiotemporal data has also been highlighted in [34]. In fact, during events such as the Hurricane Sandy's power outage, the only source left to provide timely data turned out to be Twitter [48].

In [63], methods for identifying related tweets from disaster-related microblogging platforms have been presented systematically, introducing pre-trained models and broader datasets. Social media's role in disaster response has also been examined in the survey proposed by [10].

Since the rise of NLMs, different word embeddings were considered to improve performance in retrieving pertinent tweets during disasters [42]: Word2vec, Glove, BERT, and crisis-specific embeddings.

In [60] a novel approach to identifying informative tweets during disasters by combining RoBERTa for text analysis and VGG-16 for image analysis, has been introduced. This hybrid method has shown promising results, with ample room for improvement by increasing the size of social media data considered. has proposed a different approach citechoi2021local, introducing a novel local event detection scheme for social networks. The approach outperforms baselines regarding precision, recall, and F-measure scores by analyzing non-geo-tagged documents with geographical data and constructing a weighted keyword graph based on social network characteristics. Notice that there is an ongoing debate of features and classification techniques based on event type, detection task, and method [35], from which derives the importance of evaluating detection techniques using public benchmarks. A comprehensive analysis of worldwide disasters, including different metrics and a discussion on evaluation strategies and detection quality assessment, can be found in [47]

Other works that have exploited similar techniques include [39], which achieve impressive accuracies between 96.31 and 97.20 for various disaster events. A similar work, [38], has proposed a neural-based approach to identify situational tweets during disasters, combining the Roberta model and a feature-based method. It has tested the Typhoon Hagupit da,taset introdua cing cross-attention multimodal (CAMM) approach, achieving an accuracy of 87.60.

Recently, with the growing interest in explainable AI, several works have tapped into this trend. In the context of human-centric explainable AI, [31] has explored machine and deep learning use during disaster events, focusing on explainable event detection and semantics-based approaches.

Moving to works only based on sentiments, [67] has developed a prediction algorithm to categorize disasters, particularly hurricanes, better using social media data. The methodology has considered hurricanes Harvey and Irma, exploring the correlation between tweets and hurricane severity. An open issue is determining the optimal data volume and tweet threshold to enhance prediction accuracy. In particular, the sentiment conveyed by analyzed tweets has been used as a primary factor. Subsequent work has used sentiment analysis and Naïve Bayes classification [16].

In addition to predictive or classification tasks, an effective contribution from sentiment analysis is monitoring public reactions and enhancing disaster response and communication strategies, even if the correct interpretation of emotionally charged language is still challenging and often ambiguous. Nevertheless, a real-time understanding of public sentiment significantly impacts the timing and quality of interventions as needed [6].

Moreover, fine-grained sentiment analysis can also lead to access to otherwise latent information. For instance, [64] has shown that different ethnic groups can react differently during crises, influenced by various factors that impact their safety. [27] has introduced the possibility to classify dangerous events in social media using factors such as sentiment, scenario, and action types, endorsing a shared approach to detect such events for public safety.

Additional work that has exploited sentiment analysis for classification purposes is proposed by [68] during multi-hazard disasters like the 2011 Tohoku earthquake, which started a tsunami and the Fukushima nuclear disaster. The most effective way to discriminate between emergency-related tweets and other ones has been sentiment analysis using support vector machines.

A geospatial sentiment analysis framework was successfully applied to social platforms (Twitter and Flickr) during Hurricane Sandy and the Nepal Earthquake. The reliability of such data has been confirmed by their compliance with official reports, which, however, have a higher latency [2]. Diachronic studies carried out during Hurricane Irma have revealed that sentiment analysis can effectively track records of changing emotional states and regional variations during disasters [58]. Finally, from a technical point of view, [15] has presented a hybrid sentiment analysis model based on RoBERTa's attention-based text embeddings and GRU's dependency capture, achieving high accuracy across various sentiment analysis tasks.

4 Conclusions and Future Perspectives

This paper offers a pleiminary review of recent applications of sentiment analysis and similar techniques borrowed from NLP to emergency coordination and response, emphasizing the use of social media datasets for detecting events and classifying texts in natural disasters.

Several datasets were created collecting tweets for such kinds of disasters and machine learning models like SVM and random forest were initially tested obtaining accurate results. Subsequently, deep learning models, including LSTM, BERT, RoBERTa, and various ensemble techniques, were employed to improve outcomes. The varied content on social media allows for addressing event-related challenges such as announcements and resource mobilization. Text classification and sentiment analysis are pivotal in identifying disaster scenarios, with deep learning models demonstrating promising levels of accuracy.

Concerning future developments, we plan to extend the survey by including all existing types of natural disasters, starting from those that have less impact in terms of scale (e.g., transport accidents) but may offer valuable insights for new lines of research. Concerning technical improvements, we will try to explore the new possibilities offered by Quantum NLP [22, 36], a new sub-field that is proving particularly effective in text classification and sentiment analysis tasks [3, 17, 9]

Finally, we plan to introduce a fine-grained grouping of disasters, following the distinction already proposed in the literature by [49], distinguishing between preparedness and early warning, impact and response, mitigation, risk, and vulnerability modeling.

References

1. Alam, F., Ofli, F., Imran, M.: Crisismmd: Multimodal twitter datasets from natural disasters. In: Proceedings of the international AAAI conference on web and social media. vol. 12 (2018)
2. Alfarrarjeh, A., Agrawal, S., Kim, S.H., Shahabi, C.: Geo-spatial multimedia sentiment analysis in disasters. In: 2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA). pp. 193–202. IEEE (2017)
3. Alomari, A., Kumar, S.A.: Hybrid classical-quantum neural network for improving space weather detection and early warning alerts. In: 2023 IEEE Cognitive Communications for Aerospace Applications Workshop (CCAAW). pp. 1–6. IEEE (2023)

4. Amato, F., Coppolino, L., Cozzolino, G., Mazzeo, G., Moscato, F., Nardone, R.: Enhancing random forest classification with nlp in dameh: A system for data management in ehealth domain. Neurocomputing **444**, 79–91 (2021)
5. Baudoin, M.A., Wolde-Georgis, T.: Disaster risk reduction efforts in the greater horn of africa. International Journal of Disaster Risk Science **6**, 49–61 (2015)
6. Beigi, G., Hu, X., Maciejewski, R., Liu, H.: An overview of sentiment analysis in social media and its applications in disaster relief. Sentiment analysis and ontology engineering: An environment of computational intelligence pp. 313–340 (2016)
7. Bioglio, L., Pensa, R.G.: Analysis and classification of privacy-sensitive content in social media posts. EPJ Data Science **11**(1), 12 (2022)
8. Bonetti, F., Leonardelli, E., Trotta, D., Guarasci, R., Tonelli, S.: Work hard, play hard: Collecting acceptability annotations through a 3D game. In: Calzolari, N., Béchet, F., Blache, P., Choukri, K., Cieri, C., Declerck, T., Goggi, S., Isahara, H., Maegaard, B., Mariani, J., Mazo, H., Odijk, J., Piperidis, S. (eds.) Proceedings of the Thirteenth Language Resources and Evaluation Conference. pp. 1740–1750. European Language Resources Association, Marseille, France (Jun 2022), `https://aclanthology.org/2022.lrec-1.185`
9. Buonaiuto, G., Guarasci, R., Minutolo, A., De Pietro, G., Esposito, M.: Quantum transfer learning for acceptability judgements. Quantum Machine Intelligence **6**(1), 13 (2024)
10. Costa, D., Peixoto, J., Jesus, T., Portugal, P., Vasques, F., Rangel, E., Peixoto, M.: A survey of emergencies management systems in smart cities. IEEE Access **10**, 61843–61872 (2022)
11. D'Andrea, E., Ducange, P., Lazzerini, B., Marcelloni, F.: Real-time detection of traffic from twitter stream analysis. IEEE Transactions on Intelligent Transportation Systems **16**(4), 2269–2283 (2015)
12. De Albuquerque, J.P., Herfort, B., Brenning, A., Zipf, A.: A geographic approach for combining social media and authoritative data towards identifying useful information for disaster management. International journal of geographical information science **29**(4), 667–689 (2015)
13. Elia, A., Pelosi, S., Maisto, A., Guarasci, R.: Towards a lexicon-grammar based framework for NLP: an opinion mining application. In: Mitkov, R., Angelova, G., Bontcheva, K. (eds.) Proceedings of the International Conference Recent Advances in Natural Language Processing. pp. 160–167. INCOMA Ltd. Shoumen, BULGARIA, Hissar, Bulgaria (Sep 2015), `https://aclanthology.org/R15-1023`
14. Elton, E.J., Gruber, M.J., Busse, J.A.: Do investors care about sentiment? Journal of Business **71**(4), 477–500 (1998)
15. Fan, C., Esparza, M., Dargin, J., Wu, F., Oztekin, B., Mostafavi, A.: Spatial biases in crowdsourced data: social media content attention concentrates on populous areas in disasters. Computers, Environment and Urban Systems **83**, 101514 (2020)
16. Fan, C., Wu, F., Mostafavi, A.: A hybrid machine learning pipeline for automated mapping of events and locations from social media in disasters. IEEE Access **8**, 10478–10490 (2020)
17. Ganguly, S., Morapakula, S.N., Coronado, L.M.P.: Quantum natural language processing based sentiment analysis using lambeq toolkit. In: 2022 Second International Conference on Power, Control and Computing Technologies (ICPC2T). pp. 1–6. IEEE (2022)
18. Garcia, C., Fearnley, C.J.: Evaluating critical links in early warning systems for natural hazards. Natural Hazards and Disaster Risk Reduction pp. 53–67 (2016)

19. Ghosh, S., Ghosh, K., Ganguly, D., Chakraborty, T., Jones, G.J., Moens, M.F., Imran, M.: Exploitation of social media for emergency relief and preparedness: Recent research and trends. Information Systems Frontiers **20**, 901–907 (2018)
20. Guarasci, R., Catelli, R., Esposito, M.: Classifying deceptive reviews for the cultural heritage domain: A lexicon-based approach for the italian language. Expert Systems with Applications **252**, 124131 (2024)
21. Guarasci, R., Damiano, E., Minutolo, A., Esposito, M., De Pietro, G.: Lexicon-grammar based open information extraction from natural language sentences in italian. Expert Systems with Applications **143**, 112954 (2020)
22. Guarasci, R., De Pietro, G., Esposito, M.: Quantum natural language processing: Challenges and opportunities. Applied sciences **12**(11), 5651 (2022)
23. Guarasci, R., Silvestri, S., De Pietro, G., Fujita, H., Esposito, M.: Bert syntactic transfer: A computational experiment on italian, french and english languages. Computer Speech & Language **71**, 101261 (2022)
24. Guarasci, R., Silvestri, S., De Pietro, G., Fujita, H., Esposito, M.: Assessing bert's ability to learn italian syntax: A study on null-subject and agreement phenomena. Journal of Ambient Intelligence and Humanized Computing **14**(1), 289–303 (2023)
25. Hatzivassiloglou, V., McKeown, K.R.: Predicting the semantic orientation of adjectives. In: Proceedings of the 35th annual meeting of the Association for Computational Linguistics and 8th Conference of the European Chapter of the Association for Computational Linguistics. pp. 174–181 (1997)
26. Imran, M., Castillo, C., Diaz, F., Vieweg, S.: Processing social media messages in mass emergency: A survey. ACM Computing Surveys (CSUR) **47**(4), 1–38 (2015)
27. Jamil, M.L., Pais, S., Cordeiro, J.: Detection of dangerous events on social media: a critical review. Social Network Analysis and Mining **12**(1), 154 (2022)
28. Ji, X., Chun, S.A., Wei, Z., Geller, J.: Twitter sentiment classification for measuring public health concerns. Social Network Analysis and Mining **5**, 1–25 (2015)
29. Jiang, C., Chen, Y., Liu, K.R.: Evolutionary dynamics of information diffusion over social networks. IEEE transactions on signal processing **62**(17), 4573–4586 (2014)
30. Khosla, P., Basu, M., Ghosh, K., Ghosh, S.: Microblog retrieval for post-disaster relief: applying and comparing neural ir models. arXiv preprint arXiv:1707.06112 (2017)
31. Kolajo, T., Daramola, O.: Human-centric and semantics-based explainable event detection: a survey. Artificial Intelligence Review **56**, 119–158 (2023)
32. Kostkova, P., Szomszor, M., St. Louis, C.: # swineflu: The use of twitter as an early warning and risk communication tool in the 2009 swine flu pandemic. ACM Transactions on Management Information Systems (TMIS) **5**(2), 1–25 (2014)
33. Kumar, N.P., Srinivasan, K., Ramesh, D.: Analyzing public sentiment towards llm: A twitter-based sentiment analysis. In: 2023 International Conference on the Confluence of Advancements in Robotics, Vision and Interdisciplinary Technology Management (IC-RVITM). pp. 1–8. IEEE (2023)
34. Kyrkou, C., Kolios, P., Theocharides, T., Polycarpou, M.: Machine learning for emergency management: A survey and future outlook. Proceedings of the IEEE **111**(1), 19–41 (2022)
35. Lamsal, R.: Design and analysis of a large-scale covid-19 tweets dataset. Applied Intelligence **51**, 2790–2804 (2021)
36. Lorenz, R., Pearson, A., Meichanetzidis, K., Kartsaklis, D., Coecke, B.: Qnlp in practice: Running compositional models of meaning on a quantum computer. Journal of Artificial Intelligence Research **76**, 1305–1342 (2023)

37. Luna, S., Pennock, M.J.: Social media applications and emergency management: A literature review and research agenda. International journal of disaster risk reduction **28**, 565–577 (2018)
38. Madichetty, S.: Identification of medical resource tweets using majority voting-based ensemble during disaster. Social Network Analysis and Mining **10**, 1–18 (2020)
39. Madichetty, S., Muthukumarasamy, S.: Detection of situational information from twitter during disaster using deep learning models. Sādhanā **45**, 1–13 (2020)
40. Marulli, F., Pota, M., Esposito, M., Maisto, A., Guarasci, R.: Tuning syntaxnet for pos tagging italian sentences. In: Advances on P2P, Parallel, Grid, Cloud and Internet Computing: Proceedings of the 12th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC-2017). pp. 314–324. Springer (2018)
41. Miah, M.S.U., Kabir, M.M., Sarwar, T.B., Safran, M., Alfarhood, S., Mridha, M.: A multimodal approach to cross-lingual sentiment analysis with ensemble of transformer and llm. Scientific Reports **14**(1), 9603 (2024)
42. Mredula, M., Dey, N., Rahman, M., Mahmud, I., Cho, Y.Z.: A review on the trends in event detection by analyzing social media platforms' data. Sensors **22**(12), 4531 (2022)
43. Mukhtar, R.: Review of national multi-hazard early warning system plan of pakistan in context with sendai framework for disaster risk reduction. Procedia engineering **212**, 206–213 (2018)
44. Nematian, F., Edalattalab, F.: Identifying factors associated with anxiety and fear among adolescent girls during the covid-19 and post-covid-19 period in iran: A systematic review study. Journal of Pediatrics Review pp. 0–0
45. Nurwidyantoro, A., Winarko, E.: Event detection in social media: a survey. In: International Conference on ICT for Smart Society. pp. 1–5. IEEE (2013)
46. Olteanu, A., Castillo, C., Diaz, F., Vieweg, S.: Crisislex: A lexicon for collecting and filtering microblogged communications in crises. In: Proceedings of the international AAAI conference on web and social media. vol. 8, pp. 376–385 (2014)
47. Pavani, T., Malla, S.: A review of deep learning techniques for disaster management in social media: trends and challenges. The European Physical Journal Special Topics pp. 1–23 (2024)
48. Pourebrahim, N., Sultana, S., Edwards, J., Gochanour, A., Mohanty, S.: Understanding communication dynamics on twitter during natural disasters: A case study of hurricane sandy. International journal of disaster risk reduction **37**, 101176 (2019)
49. Qadir, J., Ali, A., ur Rasool, R., Zwitter, A., Sathiaseelan, A., Crowcroft, J.: Crisis analytics: big data-driven crisis response. Journal of International Humanitarian Action **1**, 1–21 (2016)
50. Ragini, J.R., Anand, P.R.: An empirical analysis and classification of crisis related tweets. In: 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC). pp. 1–4. IEEE (2016)
51. Reuter, C., Kaufhold, M.A.: Fifteen years of social media in emergencies: a retrospective review and future directions for crisis informatics. Journal of contingencies and crisis management **26**(1), 41–57 (2018)
52. Sahayak, V., Shete, V., Pathan, A.: Sentiment analysis on twitter data. International Journal of Innovative Research in Advanced Engineering (IJIRAE) **2**(1), 178–183 (2015)

53. Sakaki, T., Okazaki, M., Matsuo, Y.: Tweet analysis for real-time event detection and earthquake reporting system development. IEEE transactions on knowledge and Data Engineering **25**(4), 919–931 (2012)
54. Sen, A., Rudra, K., Ghosh, S.: Extracting situational awareness from microblogs during disaster events. In: 2015 7th international conference on communication systems and networks (Comsnets). pp. 1–6. IEEE (2015)
55. Sreenivasulu, M., Sridevi, M.: Comparative study of statistical features to detect the target event during disaster. Big Data Mining and Analytics **3**(2), 121–130 (2020)
56. Tran, V., Matsui, T.: Improving llm prompting with ensemble of instructions: A case study on sentiment analysis. In: JSAI International Symposium on Artificial Intelligence. pp. 299–305. Springer (2024)
57. Trotta, D., Guarasci, R., Leonardelli, E., Tonelli, S.: Monolingual and cross-lingual acceptability judgments with the Italian CoLA corpus. In: Moens, M.F., Huang, X., Specia, L., Yih, S.W.t. (eds.) Findings of the Association for Computational Linguistics: EMNLP 2021. pp. 2929–2940. Association for Computational Linguistics, Punta Cana, Dominican Republic (Nov 2021). https://doi.org/10.18653/v1/2021.findings-emnlp.250, `https://aclanthology.org/2021.findings-emnlp.250`
58. Vayansky, I., Kumar, S.A., Li, Z.: An evaluation of geotagged twitter data during hurricane irma using sentiment analysis and topic modeling for disaster resilience. In: 2019 IEEE international symposium on technology and society (ISTAS). pp. 1–6. IEEE (2019)
59. Vo, B.K.H., Collier, N.: Twitter emotion analysis in earthquake situations. Int. J. Comput. Linguistics Appl. **4**(1), 159–173 (2013)
60. Wang, H., Hovy, E., Dredze, M.: The hurricane sandy twitter corpus (2014)
61. Wiebe, J.: Learning subjective adjectives from corpora. In: Proceedings of the National Conference on Artificial Intelligence. pp. 735–740 (2001)
62. Wiebe, J., Bruce, R.F., O'Hara, T.P.: Development and use of a gold-standard data set for subjectivity classifications. In: Proceedings of the 37th annual meeting of the Association for Computational Linguistics. pp. 246–253 (1999)
63. Xiang, W., Wang, B.: A survey of event extraction from text. IEEE Access **7**, 173111–173137 (2019)
64. Xiong, S., Lv, H., Zhao, W., Ji, D.: Towards twitter sentiment classification by multi-level sentiment-enriched word embeddings. Neurocomputing **275**, 2459–2466 (2018)
65. Xue, J., Chen, J., Hu, R., Chen, C., Zheng, C., Su, Y., Zhu, T.: Twitter discussions and emotions about the covid-19 pandemic: Machine learning approach. Journal of medical Internet research **22**(11), e20550 (2020)
66. Yao, Y., Duan, J., Xu, K., Cai, Y., Sun, Z., Zhang, Y.: A survey on large language model (llm) security and privacy: The good, the bad, and the ugly. High-Confidence Computing p. 100211 (2024)
67. Zahera, H., Elgendy, I., Jalota, R., Sherif, M., Voorhees, E.: Fine-tuned bert model for multi-label tweets classification. In: TREC. pp. 1–7 (2019)
68. Zahra, K., Das, R.D., Ostermann, F.O., Purves, R.S.: Towards an automated information extraction model from twitter threads during disasters. In: 19th International Conference on Information Systems for Crisis Response and Management, ISCRAM 2022. pp. 637–653. ISCRAM (2022)
69. Zhou, J., Yang, S., Xiao, C., Chen, F.: Examination of community sentiment dynamics due to covid-19 pandemic: a case study from a state in australia. SN Computer Science **2**, 1–11 (2021)