sustainability

Systematic Review

A Sustainable Way Forward: Systematic Review of Transformer Technology in Social-Media-Based Disaster Analytics

Fahim Sufi

School of Public Health and Preventive Medicine, Monash University, Melbourne, VIC 3004, Australia; research@fahimsufi.com or fahim.sufi@monash.edu

Abstract: Transformer technologies, like generative pre-trained transformers (GPTs) and bidirectional encoder representations from transformers (BERT) are increasingly utilized for understanding diverse social media content. Despite their popularity, there is a notable absence of a systematic literature review on their application in disaster analytics. This study investigates the utilization of transformer-based technology in analyzing social media data for disaster and emergency crisis events. Leveraging a systematic review methodology, 114 related works were collated from popular databases like Web of Science and Scopus. After deduplication and following the exclusion criteria, 53 scholarly articles were analyzed, revealing insights into the geographical distribution of research efforts, trends in publication output over time, publication venues, primary research domains, and prevalently used technology. The results show a significant increase in publications since 2020, with a predominant focus on computer science, followed by engineering and decision sciences. The results emphasize that within the realm of social-media-based disaster analytics, BERT was utilized in 29 papers, BERT-based methods were employed in 28 papers, and GPT-based approaches were featured in 4 papers, indicating their predominant usage in the field. Additionally, this study presents a novel classification scheme consisting of 10 distinct categories that thoroughly categorize all existing scholarly works on disaster monitoring. However, the study acknowledges limitations related to sycophantic behavior and hallucinations in GPT-based systems and raises ethical considerations and privacy concerns associated with the use of social media data. To address these issues, it proposes strategies for enhancing model robustness, refining data validation techniques, and integrating human oversight mechanisms.

Keywords: GPTs in disaster analytics; LLM for emergency situations; transformer for crisis; social media analytics; BERT; language models; systematic literature review; algorithms

check for updates

Citation: Sufi, F. A Sustainable Way Forward: Systematic Review of Transformer Technology in Social-Media-Based Disaster Analytics. *Sustainability* **2024**, *16*, 2742. https://doi.org/10.3390/su16072742

Academic Editor: Jun (Justin) Li

Received: 26 February 2024
Revised: 20 March 2024
Accepted: 25 March 2024
Published: 26 March 2024

1. Introduction

Traditionally, scholarly investigations into natural and anthropogenic disasters, such as landslides [1,2], have relied upon machine learning (ML) and artificial intelligence (AI) algorithms, which necessitated access to meticulously curated and validated data sources (e.g., [3–5]) that were often challenging to procure. The widespread adoption of social media platforms by billions of users has emerged as an alternative and sustainable avenue for acquiring data pertaining to disaster, crisis, and emergency scenarios. A multitude of active social media users regularly disseminate real-time updates concerning critical crisis events, including earthquakes, landslides, floods, shootings, wildfires, and even pandemics. Owing to the dearth of officially curated disaster data repositories, recent scholarly endeavors have turned to harnessing social media posts to identify and evaluate natural calamities such as landslides, floods, earthquakes, typhoons, wildfires, and others [6–8]. Moreover, the dissemination, evaluation, and repercussions of the COVID-19 pandemic have been elucidated through the innovative application of ML and AI algorithms to live social media content, exemplified by platforms such as Twitter [9,10]. As shown in [9,10], through AI-driven methods researchers could identify and categorize the vast array of discussions surrounding the pandemic, enabling policymakers to tailor responses more effectively to specific contexts and needs. These analytics facilitate a comprehensive understanding of global perspectives and multilingual expressions related to COVID-19, ensuring that disaster response strategies are informed by a diverse range of voices and concerns, ultimately leading to more effective and responsive public health communication and intervention strategies.

With the advent of generative pre-trained transformers (GPTs), large language models (LLMs), and transformer technology, complex social media posts can now be comprehended by machines, with invaluable insights [11–13]. Transformers are a class of deep learning models that excel in capturing long-range dependencies in sequential data, making them particularly effective in language tasks. By parsing through vast volumes of social media data with remarkable speed and accuracy, these new technologies enable researchers to discern emerging trends, identify key information, and gauge public sentiment, thereby facilitating more informed decision making in times of crisis. Leveraging the capabilities of GPT, transformer, and LLM technologies empowers analysts to extract actionable intelligence from the chaotic landscape of social media, enhancing our understanding and response to critical events in a sustainable manner.

The gaps in the existing literature are threefold. Firstly, there is a scarcity of comprehensive reviews that systematically collate and analyze the use of transformer technologies in disaster analytics, hindering the development of a cohesive understanding of this field. Secondly, a clear categorization of how transformer-based technologies could assist in disaster analytics is completely missing. Lastly, critical discussions on the limitations, ethical considerations, and future directions of these technologies in disaster analytics are notably sparse. By systematically reviewing and categorizing existing research, this study aims to fill these gaps, providing a foundation for future research to build upon and ensuring that the exploration of these technologies in disaster analytics progresses in a balanced and ethically considerate manner. By articulating the current state of research, this study aims to contribute to the field by (1) synthesizing existing applications of transformer technologies in disaster analytics, (2) developing a comprehensive categorization of how existing research applied transformer technologies on social media-based disaster analytics, (3) identifying research trends and insights, and (4) proposing avenues for future research that address current limitations, particularly in enhancing the ethical use of social media data and improving the robustness and adaptability of transformer technologies in diverse disaster contexts.

In short, this study answers a critical research question on "how do transformer technologies contribute to disaster analytics, and what are the implications for categorizing their applications, addressing limitations, within the context of social media-based disaster analysis". Thus, this investigation systematically examined the extant literature employing innovative transformer-based technology for the critical analysis of disaster and emergency crisis events. Through the meticulous configuration of pertinent keywords within prominent academic databases, namely Web of Science and Scopus [14,15], a total of 114 relevant publications were retrieved. Employing the preferred reporting items for systematic reviews and meta-analyses (PRISMA) methodology, 53 highly pertinent academic studies were identified for comprehensive review and analysis. This meticulous review and analysis facilitated the development of a novel classification scheme, categorizing the aforementioned 53 publications into 10 distinct categories. Additionally, it is noteworthy that beyond the creation of this innovative classification scheme for existing literature concerning the utilization of transformer technology in social-media-based disaster analytics, this study presents a methodological and systematic review of the subject of employing contemporary AI-based tools such as Litmaps [16,17].

In the subsequent section (namely, Section 2), a succinct contextual backdrop on social-media-based disaster analytics is proffered, alongside an explication of the systematic literature review methodology employed in this research endeavor. Section 3 elucidates the outcomes of the literature review, accompanied by an extensive discourse on the comprehensive categorization of extant scholarly works. Furthermore, a bibliographic scrutiny is furnished pertaining to the resultant corpus of existing publications.

2. Materials and Methods

A brief contextual background of both transformer-related algorithms and social-media-based disaster analytics is necessary for clearly articulating the systematic literature methodology within this study.

2.1. Algorithms in Social-Media-Based Disaster Analytics

The evolution of algorithms leading to the development of GPT can be traced through significant milestones in natural language processing (NLP) and deep learning. Beginning with traditional techniques like support vector machines (SVMs) and TF-IDF for classification and feature extraction, respectively, in the late 1990s and early 2000s, the field advanced with algorithms like TextRank and latent dirichlet allocation (LDA) for text summarization and topic modeling around 2004. In the era of deep learning, convolutional neural networks (CNNs) emerged as powerful tools for various tasks, including NLP, around 2012. Subsequently, the introduction of transformer architecture revolutionized sequence modeling, paving the way for models like BERT, which improved natural language understanding, around 2018. Variants such as ALBERT, RoBERTa, and Distil-BERT aimed at enhancing efficiency and performance, while MobileBERT addressed resource constraints from 2019 onwards. GPT, based on the transformer architecture, represents the pinnacle of this evolution, leveraging pre-training on vast corpora to achieve remarkable capabilities in natural language generation and understanding; this started in 2018. Table 1 shows the details of these algorithms and Figure 1 shows the evolution.

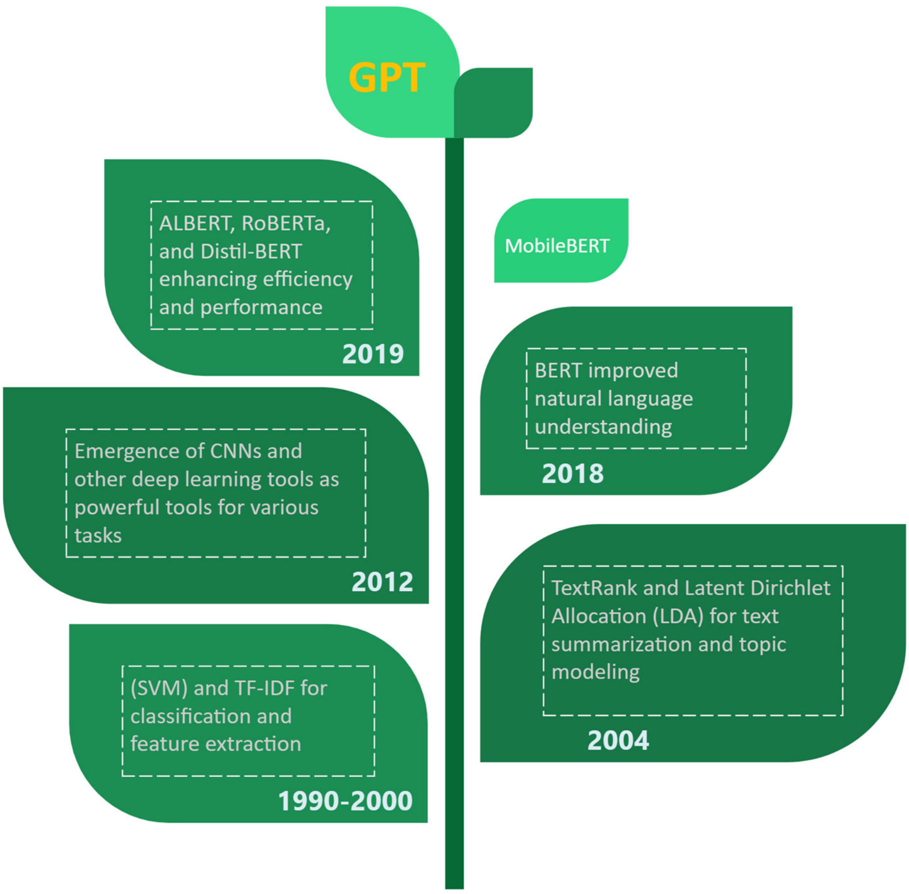

Figure 1. Evolution of algorithms leading toward GPT technologies.

Table 1. Algorithm categories and algorithms investigated in this review.

Algorithm Category	Abbreviation	Short Description
GPT-based	GPT	Generative Pre-trained Transformer—A model architecture used primarily for natural language understanding and generation.
BERT	BERT	Bidirectional Encoder Representations from Transformers—A model architecture for natural language understanding tasks.
BERT Based	ALBERT	A Lite BERT-based Model—An improved version of BERT with reduced parameter size and faster training.
	RoBERTa	Robustly optimized BERT approach—A variant of BERT with modifications to improve performance and training efficiency.
	DistilBERT	Distilled BERT—A smaller, faster version of BERT designed for resource-constrained environments.
	MobileBERT	A BERT variant optimized for mobile devices with reduced parameters and computational requirements.
Deep Learning	CNN	Convolutional Neural Network—A deep learning architecture commonly used for image recognition and natural language processing tasks.
	AVEDL	Average Voting Ensemble Deep Learning Model—A model ensemble technique combining multiple deep learning architectures for improved performance.
	Transformer	Transformer—A deep learning architecture known for its effectiveness in sequence-to-sequence tasks such as language translation and text summarization.
	T5	T5, or Text-to-Text Transfer Transformer, is a versatile machine learning model designed to convert all natural language processing tasks into a unified text-to-text framework.
Other	SVM	Support Vector Machine—A supervised learning model used for classification and regression analysis.
	TF-IDF	Term Frequency-Inverse Document Frequency—A numerical statistic used to evaluate the importance of a word in a document corpus.
	TextRank	TextRank—An algorithm for automatic text summarization based on graph-based ranking.
	LDA	Latent Dirichlet Allocation—A generative statistical model used for topic modeling in text corpora.

2.2. Social-Media-Based Disaster Analytics

Figure 2 illustrates a conceptual model depicting the utilization of GPT, LLM, and transformer technology to gain critical insights into an earthquake event that occurred in Japan. As depicted in Figure 2, social media users expressed their concerns regarding the earthquake through various forms of posts, including live updates, expressions of support for victims, and calls for assistance from those directly affected. Systems equipped with GPT, LLM, or transformer technology have the capability to automatically analyze millions of these messages, enabling the detection of disaster or crisis events, determination of event location, and assessment of event severity, as elucidated in Figure 2. Essential inquiries regarding these disasters, such as their location, impact severity, causation, and assistance requirements, can all be addressed using this advanced technology, as depicted in Figure 2. The responses to fundamental questions regarding who, what, where, when, why, and how, illustrated in Figure 3, have the potential to significantly alleviate the adverse effects of disaster, crisis, and emergency events.

Researchers in [18] demonstrated how BERT and multi-layer perceptron (MLP) technologies have been directly applied to enhance disaster response outcomes. Focusing on the DKI Jakarta flood disaster in early 2020, the research utilized BERT for classifying tweets related to flooding incidents and MLP to process geospatial features, achieving an accuracy of 82% without stemming and with stop-word removal. This approach not only enabled the effective categorization of tweets into "flooded" and "not flooded" but also facilitated the visualization of classified tweets on a two-dimensional interactive map, thereby providing critical insights for disaster response and situational awareness. This novel application of transformer technologies underscores their potential in leveraging social media data for timely and accurate disaster response and management [18].

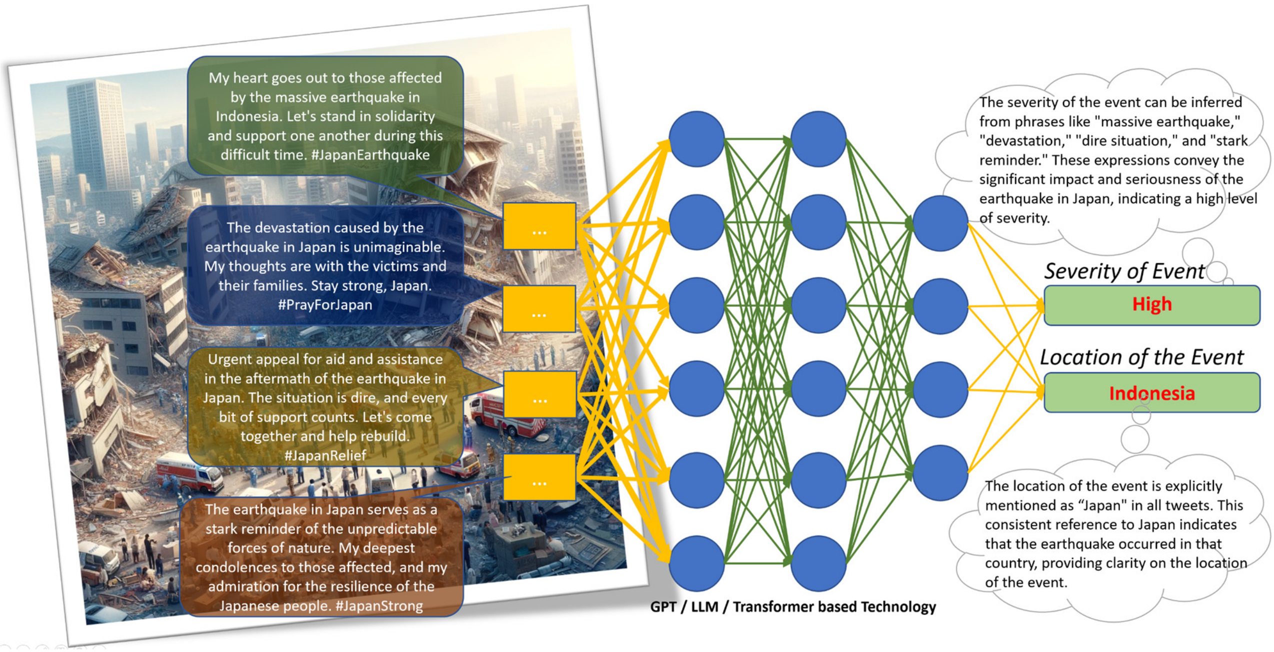

Figure 2. Use of transformer technology in social-media-based disaster event detection and classification.

Figure 3. Use of transformer technology in social-media-based disaster analytics for answering specific questions.

In another study, researchers introduce a novel average voting ensemble deep learning model (AVEDL model) that combines pre-trained transformer-based models like BERT, DistilBERT, and RoBERTa [19]. This model aims to classify emotions from COVID-19-related emergency calls and social media data, showcasing the direct impact of transformer and GPT technologies on disaster response outcomes. By achieving an accuracy of 86.46%

and a macro-average F1-score of 85.20%, the AVEDL model outperforms standard deep learning and machine learning models in detecting emotions from textual data during the pandemic. This approach demonstrates the effectiveness of leveraging advanced NLP techniques to support mental health care and emergency response efforts by understanding and addressing the public's emotional state during a crisis [19].

2.3. Systematic Literature Review

This study used major databases like Scopus and Web of Science to acquire existing scholarly works that used transformers on social-media-based disaster analytics. Earlier work in [14,15] has suggested using these data sources (i.e., Scopus and Web of Science) as primary sources. These benchmark studies on systematic literature review (i.e., [14,15]) suggested against using Google Scholar as a primary source within a systematic study. Despite its potential utility for exploratory analysis, Google Scholar's lack of clear inclusion criteria and limitations in executing advanced query-based searches, essential for systematic literature review, have led previous scholars to advise against its use for such purposes [15]. Section 2.1 outlines the necessity of utilizing a variety of keywords to denote algorithm types, such as GPT, transformer, and LLM. Additionally, it emphasizes the importance of including multiple disaster scenarios like flood, earthquake, cyclone, and landslide, as well as relevant social media platforms such as social media and Twitter, as discussed in Section 2.2. These keywords are combined using "AND" and "OR" logic to precisely define the scenario described in Section 2.2. Consequently, an advanced query approach is employed for this research.

Table 2 shows the advanced query used for capturing most relevant research works from databases like Scopus and Web of Science. Data sources and platforms that do not support advanced query were explicitly excluded (e.g., Google Scholar [15]). As seen from Table 2, 79 articles were found in Scopus and another 33 papers were found in Web of Science. For acquiring a more comprehensive list of the existing literature, AI-based tools like Litmaps were also utilized as alternative registers. By typing the DOIs of published papers found in both Scopus and Web of Science, any works that cited these seed works were identified in Litmaps. An additional two articles were suggested by Litmaps, as shown in Figure 4. Figure 4 represents the PRISMA flow diagram and Table 3 shows the detailed inclusion and exclusion criteria used for this study.

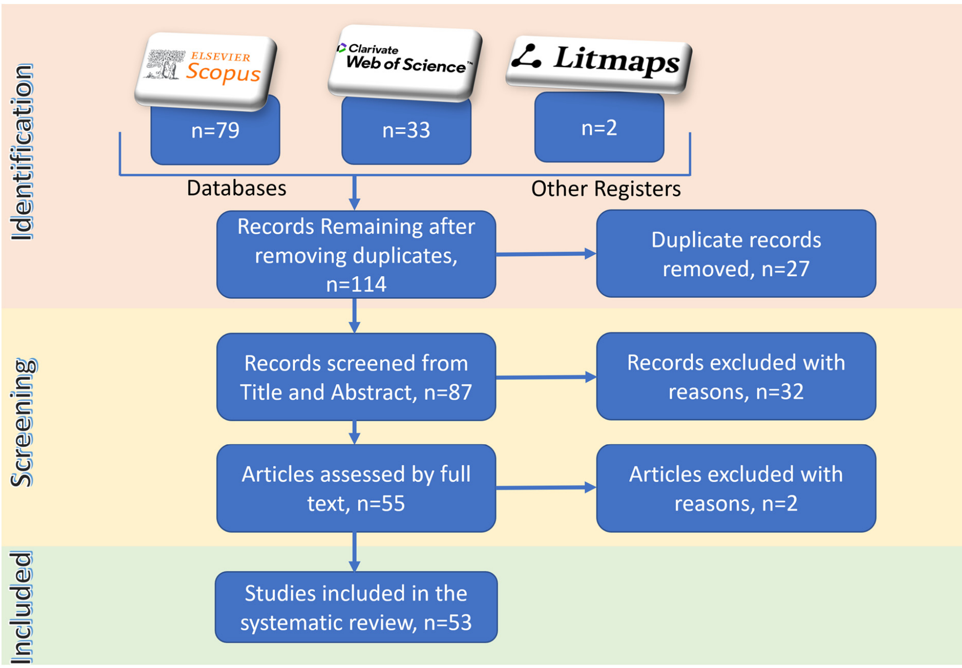

Figure 4. Flowchart showing how PRISMA methodology was applied in conducting this systematic literature review.

Table 2. Use of advanced query for database search.

Database Name	Advanced Query	Results Returned
Scopus	(TITLE-ABS-KEY ("Transformer") OR TITLE-ABS-KEY ("GPT") OR TITLE-ABS-KEY ("LLM")) AND (TITLE-ABS-KEY ("Disaster") OR TITLE-ABS-KEY ("Landslide") OR TITLE-ABS-KEY ("Flood") OR TITLE-ABS-KEY ("Earthquake") OR TITLE-ABS-KEY ("Cyclone") OR TITLE-ABS-KEY ("Typhoon")) AND (TITLE-ABS-KEY ("Twitter") OR TITLE-ABS-KEY ("Social media")) AND (LIMIT-TO (LANGUAGE, "English"))	79
Web of Science	(ALL = (Transformer) OR ALL = (GPT) OR ALL = (LLM)) AND (ALL = (Disaster) OR ALL = (Landslide) OR ALL = (Flood) OR ALL = (Earthquake) OR ALL = (Cyclone) OR ALL = (Typhoon)) AND (ALL = (Twitter) OR ALL = (Social media))	33

Table 3. Inclusion and exclusion criteria for acquiring existing scholarly works.

Category	Criteria
Inclusion	• Advanced query = (Transformer ∨ GPT ∨ LLM) ∧ (Disaster ∨ Landslide ∨ Flood ∨ Earthquake ∨ Cyclone ∨ Typhoon) ∧ (Twitter ∨ Social Media) • Peer-reviewed reviews, peer-reviewed original research articles, studies stored in pre-print servers • Papers indexed in popular peer-reviewed sources (i.e., Scopus, Web of Science) • Papers focusing into research scope "Transformer technologies used in social-media-based disaster analytics" • Studies available in English language • Studies available in full text
Exclusion	• Papers not in English language • Tutorial papers • Short papers less than 4 pages • Poster papers, editorials, abstracts (i.e., lacking detailed information)

As seen from Figure 4, from a total of 114 articles found from databases (Scopus and Web of Science) and the register (i.e., Litmaps), 27 duplicates were identified. After removing these 27 articles, 87 articles were screened via critical analysis of their titles and abstracts. Any articles that did not explicitly cover the focus of study (i.e., "use of transformer technology in social-media based disaster analytics") were screened out. As shown in Figure 4, 32 articles were initially screened out before obtaining the full texts. Finally, full texts were downloaded for 55 studies. After critical review of all of these 55 articles, 2 articles were excluded as they did not use any transformer technology to analyze disaster situations.

This rigorous systematic literature review process is reproducible, as the precise queries employed within the respective data sources are delineated in Table 2. Notably, these queries were executed on 15 February 2024 and, consequently, any studies published subsequent to this date was not incorporated into the study. Furthermore, inclusion criteria encompassed solely peer-reviewed journal and conference papers in the English language, while unpublished studies failing to meet the stringent standards of peer review were excluded from this systematic literature review.

Following the systematic literature review, a comprehensive categorization scheme was synthesized following the thorough examination of the 53 existing literacy works. Each research work was meticulously analyzed to identify the common themes, objectives, and methodologies employed. These themes were then organized into distinct categories based on their overarching goals and contributions to the field of disaster management and response. The resulting category scheme aimed to provide a comprehensive overview of the diverse research efforts in utilizing advanced technologies for analyzing social media data in disaster and crisis contexts, offering insights into the breadth and depth of the research landscape. The next section details the results of this categorization scheme.

In summary, this study conducted a systematic review of the literature on using transformer-based technologies for disaster analytics on social media. It employed advanced query strategies in major databases like Scopus and Web of Science, excluding Google Scholar due to its limitations. The search process emphasized diverse keywords and disaster scenarios, resulting in 114 identified articles, of which 55 were subjected to thorough analysis. Two articles were excluded for not using transformer technology. Reproducibility was ensured by detailing search queries, limited to literature before 15 February 2024, and restricted to peer-reviewed English publications. A categorization scheme was developed based on common themes, offering insights into research efforts in leveraging technology for disaster management via social media analytics.

3. Results

Using the above methodology of systematic literature review, 53 existing works specifically focusing on using transformer-based technologies on social-media-driven disaster analytics were found. When categorizing studies into broader categories by reading the literature, the key themes, methodologies, research goals, and outcomes mentioned across these documents were identified. Then, these studies were grouped based on similarities in these aspects, forming categories that reflected the overarching topics or research areas within the literature. The category scheme serves as a valuable organizational framework, allowing researchers and practitioners to navigate the complex and diverse landscape of GPT/LLM/transformer technology applications in disaster and crisis analysis on social media. By systematically categorizing research works into distinct themes and objectives, the scheme facilitates a better understanding of the various dimensions of research in this domain. Additionally, it provides a basis for comparative analysis, enabling researchers to identify trends, gaps, and emerging areas of interest for further investigation and development. As shown in Figure 5, the following 10 categories comprehensively classify the existing scholarly works by critically looking at the areas of disasters, algorithms used, technologies used, benefits, and disadvantages. Tables 4–13 categorize all of the 53 published papers into the 10 proposed broader categories.

Figure 5. Categorization of the existing studies into six different broader categories.

Individual breakdowns from all of the 53 existing works are provided in Appendix A (Tables A1–A9). These tables offer a numerical summary of transformer technology applications in disaster management on social media, showcasing accuracy percentages, precision, recall, F1 scores, and other quantitative metrics across various domains. For instance, the accuracy levels for disaster event detection and classification range from 82% to 98%, sentiment analysis achieves precision scores of up to 97.63%, recall scores exceed 96.64%, and tweet classification attains F1 scores as high as 97.16%. These metrics underscore the effectiveness of different algorithms in addressing diverse challenges in disaster management through social media analytics, while providing insights into their practical implications and potential limitations.

3.1. Disaster Event Detection and Classification

Within these classification groups, studies that portray techniques for detecting and classifying various types of disasters, such as floods, wildfires, earthquakes, etc., from social media data are included. Thus, 22 of the 53 studies found during this literature review were included [18,20–40] (as shown in Table 4 and detailed in Appendix A, Table A1). As seen from Table 4, models within this category demonstrate a commendable ability to discern disaster-related content within the vast expanse of social media data, often yielding high accuracy rates in classification tasks. However, this effectiveness is juxtaposed with various challenges and limitations, ranging from the stemming process potentially removing crucial features to the computational demands necessitated by complex model architectures.

Table 4. Summarization of algorithm categories for "Disaster Event Detection and Classification".

Algorithm Category	References	Generic Advantages	Generic Disadvantages
BERT	[18,20–22,27,28,30,31,33–37,40]	High accuracy in classifying tweets related to specific disasters.	Stemming process may remove important features, noise from broad search terms, reliance on keyword position, challenges with imbalanced data, and processing informal social media text.
BERT-based (e.g., RoBERTa, DistilBERT)	[23–25,29,32,38,39]	Superior performance in various tasks, including textual and visual analysis.	Substantial computational and memory resources, potential hardware limitations, complexity in integrating multimodal data, and challenges with real-time processing and scalability.
GPT-based	[26]	Enhanced performance in sentiment analysis on large-scale datasets, including text, images, and audio.	Complexity in implementation and optimization.

3.2. Sentiment Analysis and Public Perception

This category includes studies focusing on analyzing public sentiment and perceptions towards disasters or crisis situations. The research works in [19,41–45] fall within this category, as shown in Table 5 (detailed in Appendix A, Table A2). Notably, models such as BERT demonstrate remarkable accuracy in classifying sentiments towards COVID-19 vaccines and reporting symptoms, leveraging contextual embeddings for nuanced understanding [19,42]. However, challenges such as lower performance exhibited by fixed embeddings, as evidenced by Word2Vec, underscore the importance of employing adaptable models to capture the contextual nuances effectively [42].

Table 5. Summarization of algorithm categories for "Sentiment Analysis and Public Perception".

Algorithm Category	References	Generic Advantage	Generic Disadvantage
Deep Learning	[19,41,44]	High accuracy in detecting beliefs, opinions, and emotions, showcasing effective analysis.	Difficulty in collecting and labeling diverse data due to variations in human dialect and speech. Limited application without extensive preprocessing and NLP understanding.
BERT	[43]	Accurate analysis of public sentiment and effective in topic modeling and sentiment analysis.	Focuses more on sentiment analysis rather than direct disaster response strategies. Potential challenges include processing vast datasets and identifying nuanced sentiment accurately.
BERT-based (e.g., RoBERTa, Word2Vec, LDA)	[42,45]	High accuracy in classifying sentiments and emotions, improving disaster response.	Word2Vec showed lower performance compared to BERT, indicating fixed embeddings may not capture contextual nuances effectively. Potential limitations include the limited availability of specific dataset details.

3.3. Information Summarization and Retrieval

This classification includes [28,33,34,46–48], which are research works aimed at summarizing and retrieving crisis-relevant information from social media data to enhance situational awareness and decision making. Table 6 presents this detailed categorization scheme (Detailed in Appendix A, Table A3). Models such as T5 for summarization [46] and GPT-3 for document retrieval demonstrate promising capabilities in facilitating effective summarization of crisis-relevant information from social media and online news sources, thereby enhancing situational awareness during crisis events. However, challenges such as the complexity of handling multilingual data and the potential for reduced accuracy in cross-lingual information retrieval and summarization underscore the need for careful consideration of contextual nuances and methodological refinement in employing transformer-based models for this purpose.

Table 6. Summarization of algorithm categories for "Information Summarization and Retrieval".

Algorithm Category	References	Generic Advantage	Generic Disadvantage
BERT	[28,33,34,48]	Effective summarization and classification of disaster-related information from social media.	Challenges in verifying the authenticity of user-generated content, potential limitations in adapting to new or unforeseen disaster types. Complexity in integrating multiple algorithms and scalability.
GPT	[46,47]	High comprehensiveness in summarizing crisis-relevant information from social media and online news. Rapid deployment due to few-shot learning.	High redundancy ratio in generated summaries, complexity in handling multilingual data, and potential for reduced accuracy in cross-lingual information retrieval and summarization.
Other (e.g., SVM, TF-IDF, TextRank, LDA)	[33,48]	Effective summarization and topic classification in specific disaster-related contexts.	Challenges in scalability, real-time processing, and handling multilingual data. Limited capability in areas with few Twitter activities and reliance on geotagged tweets.

3.4. Location Identification and Description Extraction

Studies in this category try to extract location descriptions and identify geographical references from disaster-related social media messages [49–51]. Table 7 summarizes the categorization of location identification and description extraction. The detailed review for each paper is provided in Appendix A, Table A4. BERT with BiLSTM-CRF achieves high accuracy in recognizing toponyms, aiding location identification in disaster communications [49,51]. However, its practical application lacks direct examples in disaster management contexts.

Integrating geo-knowledge with GPT models like ChatGPT and GPT-4 significantly improves location description extraction accuracy from social media, outperforming traditional NER approaches by over 40% [50]. Yet, its effectiveness depends on the availability and quality of geo-knowledge, posing challenges in generalization across regions and disaster types.

Table 7. Summarization of algorithm categories for "Location Identification and Description Extraction".

Algorithm Category	References	Generic Advantage	Generic Disadvantage
BERT	[49,51]	High accuracy in location identification, crisis tweet classification, and extraction of location descriptions from social media messages.	Challenges include potential lack of direct disaster management application examples and handling diverse data quality.
GPT-based	[50]	Significant improvement in accuracy of location extraction from social media messages by leveraging GPT models.	Effectiveness contingent on availability and quality of geo-knowledge about common forms of location descriptions.

3.5. Tweet Prioritization and Useful Information Extraction

Research works in this category demonstrate methods for prioritizing tweets and extracting useful information, such as rescue requests or informative tweets, to aid in disaster response efforts. Within this systematic literature review at least nine studies were categorized within this classification [29,31,52–58]. This study critically analyzed 53 published papers and identified 9 belonging into this category (as shown in Table 8). The detailed paper-by-paper comparison is provided in Appendix A, Table A5. These technologies, including BERT, RoBERTa, and multi-task learning approaches, show promise in accurately categorizing tweets, predicting flood levels, and enhancing database classifications [29,52,53]. However, challenges such as model complexity, computational costs, and the need for extensive training datasets underscore the ongoing efforts required to optimize their effectiveness in real-world disaster management scenarios. In particular, flash flood classification employs FF-BERT, a multi-label text classification model, to enhance existing databases [57]. Despite improvements, it exhibits relatively low prediction performance for minority labels compared to the baseline model.

Table 8. Summarization of algorithm categories for "Tweet prioritization and useful information extraction".

Algorithm Category	References	Generic Advantage	Generic Disadvantage
BERT-based	[29,52–55]	Achieves high accuracy and performance in various disaster-related tasks.	Complexity of models, computational costs, reliance on extensive datasets for training, potential challenges in handling large volumes of social media data, and processing informal text.
BERT	[29,31,56,57]	Effective in tweet classification, semantic similarity, and multi-label text classification.	May require integration of methods to assess tweet credibility, challenges with imbalanced data, and processing informal social media text.
Other (e.g., Deep Learning)	[54,58]	Achieves high accuracy and precision in predicting disaster-related events and identifying informative tweets.	Complexity of model implementation, reliance on large datasets for training, and computational overhead.

3.6. Multimodal Data Analysis

This category represents research works utilizing both textual and visual data from social media for more comprehensive disaster analysis and response. The existing studies [45,59–63] are grouped under this category, as demonstrated in Table 9. The detailed summary for papers in the "Multimodal Data Analysis" category is provided in Appendix A, Table A6.

Table 9. Summarization of algorithm categories for "Multimodal Data Analysis".

Algorithm Category	References	Generic Advantage	Generic Disadvantage
BERT-based	[45,59–63]	High accuracy and performance in disaster-related tasks, including flood detection and disaster management.	Complexity of integrating and optimizing multimodal data inputs, potential limitations in single modality analyses, challenges with feature extraction generalizability.
Other (e.g., Multimodal, LDA)	[59,60,63]	Effective fusion of textual and visual data, leading to more accurate informative tweet classification.	Complexity of the multimodal analysis process, reliance on extensive data preprocessing and manual labeling for accurate model training.

3.7. Multilingual and Cross-Lingual Disaster Analysis

These are studies addressing disaster analysis and classification across different languages or in multilingual contexts such as the works demonstrated in [55,64] (Table 10). The detailed summaries for both [55,64] are located in Appendix A, Table A7.

Table 10. Summarization of algorithm categories for "Multilingual and Cross-lingual Disaster Analysis".

Algorithm Category	References	Generic Advantage	Generic Disadvantage
BERT-Based	[55,64]	Achieved top performance in tweet prioritization and outperformed median performance for information type classification using pre-trained language models.	Not specified, but complexity and integration of GNN with transformer models might introduce computational overhead.

3.8. Emotion and Sentiment Identification

Studies in this group attempt to identify the emotions and sentiments expressed in social media posts related to disasters or crises. For example, [19] falls under this category, as shown in Table 11.

Table 11. Comparison of transformer technology used on social media for "Emotion and Sentiment Identification".

Ref.	Area of Disaster	Algorithm Used	GPT/Transformer Technology	Benefit	Disadvantage
[19]	Emotion Identification During COVID-19	Average Voting Ensemble Deep Learning Model (AVEDL Model) Incorporating BERT, DistilBERT, RoBERTa	BERT, DistilBERT, RoBERTa	Achieved high accuracy (86.46%) and macro-average F1-score (85.20%) in classifying emotions from COVID-19-related social media and emergency response calls, showcasing effective emotion analysis in pandemic conditions	The model's performance is contingent on the quality and size of the dataset, and its application is limited without extensive preprocessing and understanding of NLP concepts for accurate emotion extraction

3.9. Performance Evaluation and Comparison of Models

This category groups research works focusing on evaluating the performance of various models and techniques in disaster-related tasks, often comparing different algorithms or approaches. The studies in [30,32,40,65–68] are categorized under this group. Table 12 clearly shows the relevant studies within this category. Detailed summaries for all papers in this category are provided in Appendix A, Table A8. This categorization compares the transformer technologies used for performance evaluation in disaster management scenarios. Technologies such as BERT, DistilBERT, and MobileBERT demonstrate high accuracy and efficiency in tasks like tweet classification and crisis event detection. Challenges include reliance on keyword position for disaster prediction, complexity in integrating multiple data sources, and potential overfitting due to the complexity of models like BERT. For example, the SMDKGG framework and transformers in [67] for metadata classification approaches achieve high precision, recall, and accuracy in generating knowledge graphs from disaster tweets. However, the integration and processing of multiple data sources and algorithms may be complex, requiring extensive computational resources and expertise in machine learning and natural language processing.

Table 12. Summarization of algorithm categories for "Performance Evaluation and Comparison of Models".

Algorithm Category	References	Generic Advantage	Generic Disadvantage
BERT	[30,40,68]	Improved accuracy in disaster prediction on Twitter by incorporating keyword position information into the BERT model; BERT achieved the highest accuracy in classifying disaster-related tweets; high accuracy and low memory usage for flood prediction using Twitter data	Relies heavily on the keyword position, which may not always accurately reflect the context or importance of a tweet; increased complexity of the BERT architecture may lead to overfitting and requires careful adjustment

Table 12. *Cont.*

Algorithm Category	References	Generic Advantage	Generic Disadvantage
BERT-based	[32,65,66]	Precision of 0.81, recall of 0.76, and F-score of 0.78 for BERT in providing appropriate guidelines; high accuracy in multilingual tweet classification for disaster response; high accuracy and efficiency in detecting and classifying crisis-related events on social media, leveraging advanced transformer technology and optimized feature selection	Limited test data representing diverse crisis scenarios and issues with handling massive datasets due to token limitations; not explicitly mentioned, but potential issues could include data sparsity and language-specific challenges; challenges could include computational demands for processing and analyzing large-scale social media data in real-time and adapting to diverse and evolving crisis scenarios
Other	[67]	High precision (96.19%), recall (98.33%), and accuracy (97.26%) in generating knowledge graphs from disaster tweets, utilizing a comprehensive metadata-driven approach and diverse knowledge sources for enriched auxiliary knowledge	Complexity in the integration and processing of multiple data sources and algorithms, requiring extensive computational resources and expertise in machine learning and natural language processing

3.10. Practical Applications and System Development

These are studies describing the development and deployment of practical systems or tools for disaster support, such as chatbots, crisis communication platforms, etc. The four studies that fall within this category are shown in Table 13. The detailed paper-wise summaries are provided in Appendix A, Table A9.

Table 13. Summarization of algorithm categories for "Practical Applications and System Development".

Algorithm Category	References	Generic Advantage	Generic Disadvantage
BERT	[68]	High accuracy and low memory usage for flood prediction using Twitter data	Not mentioned explicitly, but complexity and potential overfitting can be inferred as disadvantages
BERT-based	[45,69]	Effective in disaster detection and flood event detection, improving decision making	Limited by potential algorithm efficiency exploration and dataset availability
Other	[70]	Real-time support, high accuracy, and low memory usage for flood prediction using Twitter data	Limited by data specificity, complexity, and potential overfitting inferred as disadvantages

4. Discussion

The validation process in this study, following the PRISMA methodology, involved rigorous screening of articles based on inclusion and exclusion criteria (detailed earlier in Table 3), ensuring relevance to the study's focus on transformer technology in social media disaster analytics. In this section, prevailing research trends and discernments (as exemplified in Figures 6–10) are elucidated. Additionally, the potential directions for future research aimed at mitigating the existing constraints are described. Thus, the research objectives delineated earlier in the introduction section have been deliberated upon.

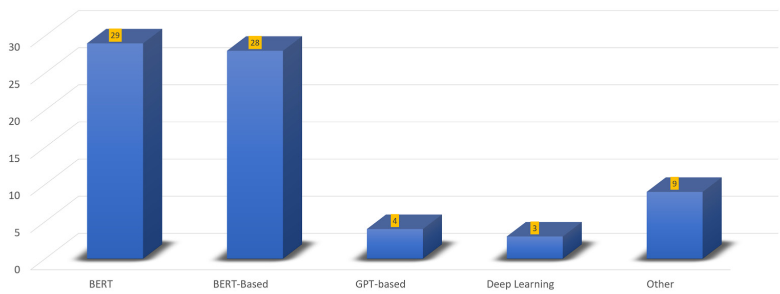

Figure 6. Five different categories of algorithms prominently being used within the surveyed scholarly works.

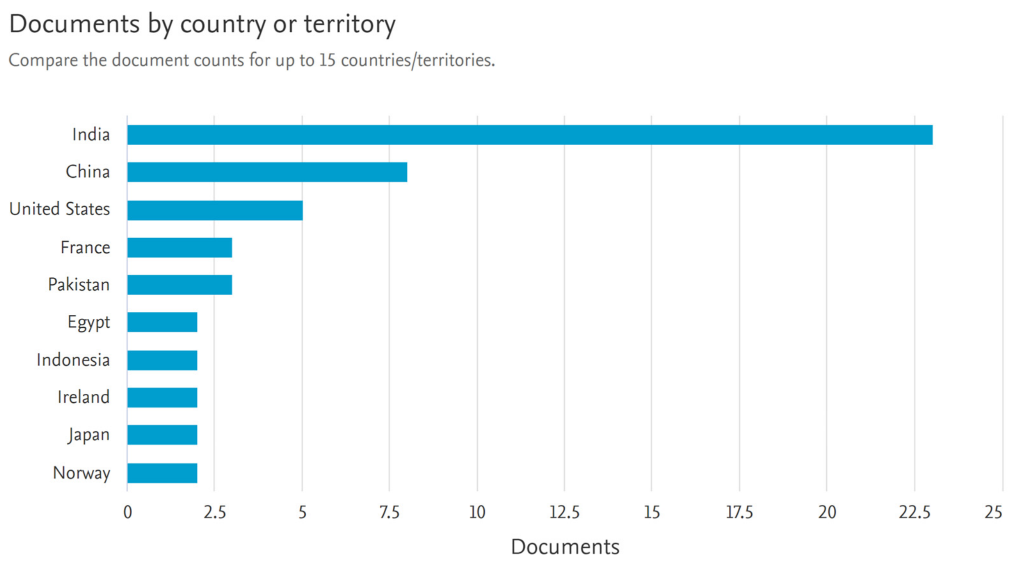

Figure 7. Number of publications by country.

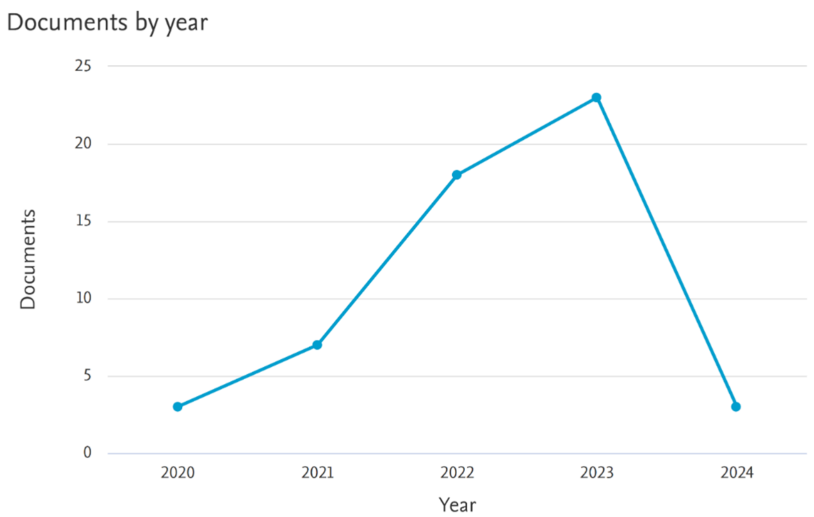

Figure 8. Number of publications by year.

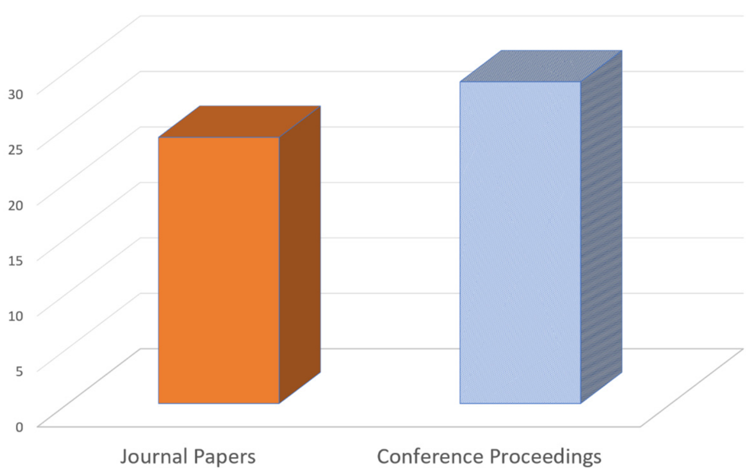

Figure 9. Number of publications by publication venue (i.e., journals and conferences).

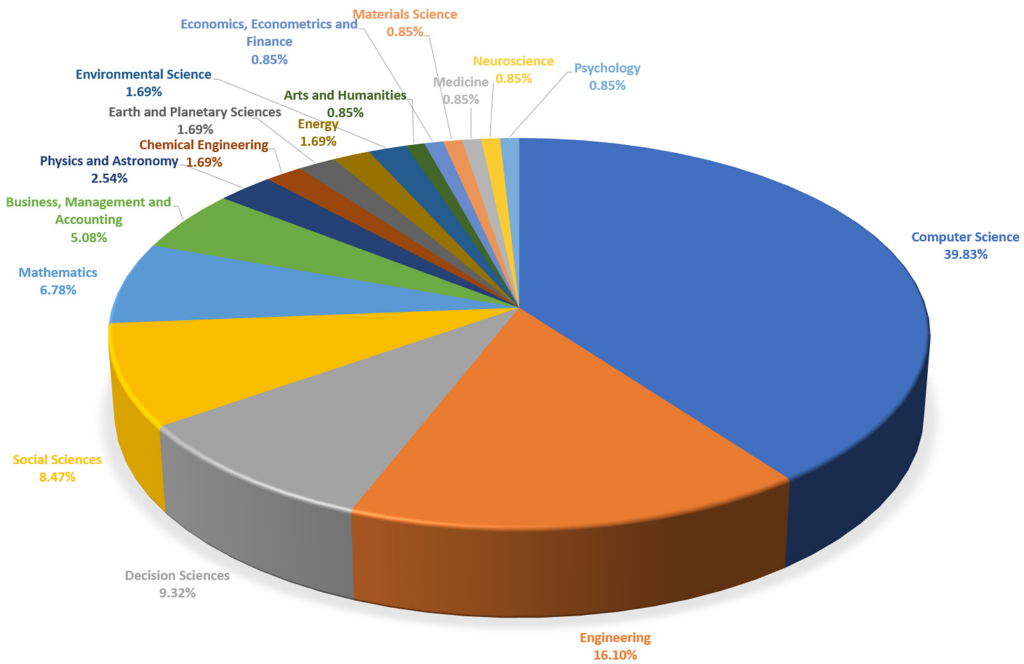

Figure 10. Number of publications by area of research.

4.1. Critical Analysis and Interpretation

Through meticulous examination spanning from Tables 4–13, coupled with an in-depth scrutiny of the algorithmic methodologies employed across these extant studies pertaining to diverse facets of disaster analytics, it was discerned that BERT emerged as the preeminent choice among algorithms, with subsequent prevalence observed in BERT-based variants such as DistilBERT, RoBERTa, and MobileBERT, among others. This trend is corroborated by the data presented in Figure 6, wherein 29 scholarly works explicitly incorporated BERT, while an additional 28 works referenced BERT-based algorithms. Furthermore, four studies leveraged GPT, three delved into deep learning paradigms, and nine opted for assorted other algorithmic frameworks. The specific classifications of these algorithms within overarching categories, as illustrated in Figure 6, were earlier delineated in Table 1. BERT and its derivatives, being the most established technologies as of 2018 and 2019 (as demonstrated earlier in Figure 1), have garnered the highest level of popularity, in contrast to emerging technologies such as GPT.

The 53 scholarly articles revealed various insight in terms of number of documents per country (i.e., origin of research), number of publications per year, the particular domain of research, the publication venues, and others. Firstly, the origin of research, as shown in Figure 7, provides insights into the geographical distribution of research efforts, revealing which countries are actively contributing to this field and potentially identifying regions where more research support or collaboration may be needed. Secondly, examining the number of publications per year (as shown in Figure 8) allows for the identification of trends and shifts in research focus over time, highlighting emerging areas of interest or declining topics.

As depicted in Figure 8, the utilization of GPT, LLM, and transformer technology in the analysis of disaster-related social media posts emerges as a prominent subject, evident from the upward trajectory in yearly publications observed since 2020. As seen from Figure 9, within the 53 articles investigated, 24 were from journals and 29 were from conference proceedings. Identifying publication venues helps researchers identify the key journals or conferences where significant contributions are being made, providing guidance for future publication strategies and networking opportunities within the academic community. Finally, categorizing publications by the specific domains of research enables researchers to understand the breadth and depth of research within different subfields, facilitating targeted exploration and knowledge synthesis. The primary research domains identified in this systematic literature review are computer science, accounting for 39.83% of the studies, followed by engineering with 16.10%, decision sciences with 9.32%, social sciences with 8.47%, and mathematics with 6.78% (as seen in Figure 10).

Overall, leveraging bibliometric data enhances understanding of the research landscape and informs strategic decision making for both individual researchers and the broader scholarly community.

4.2. Limitations of Social-Media-Based Disaster Analytics Using Transformer

The application of GPT-based disaster monitoring systems is susceptible to the limitations associated with sycophantic behavior and hallucinations. Sycophantic behavior may manifest when the system overly praises or exaggerates the severity of a situation based on the tone or content of social media posts, potentially leading to inaccurate assessments and responses [71]. Additionally, hallucinations may occur when the model generates false or misleading information based on incomplete or ambiguous data, resulting in erroneous interpretations of disaster-related events and subsequent actions [72]. These limitations underscore the importance of thorough validation and contextual understanding in utilizing GPT technology for disaster monitoring.

The utilization of social media data for disaster analytics, while invaluable for real-time insights and response optimization, is fraught with ethical considerations and potential privacy concerns. A significant ethical issue arises from the prevalence of fake users and accounts [73], which can distort analyses and lead to misinformation being spread during critical times, complicating disaster response efforts. Additionally, the proliferation of fake news on social media platforms exacerbates the challenge of distinguishing between reliable and misleading information, making it crucial for disaster analytics tools to incorporate robust verification mechanisms [74]. Privacy concerns are also paramount, as the collection and analysis of social media data might inadvertently expose sensitive personal information, risking breaches of individual privacy and violating user consent [75]. Ethical considerations extend to ensuring that the data used does not perpetuate biases or inequalities, given that social media users do not represent the entire population affected by a disaster. Hence, disaster analytics initiatives must prioritize ethical guidelines that include transparency, consent, privacy protection, and the critical evaluation of data sources to mitigate the impact of fake content, all while respecting the diverse voices within affected communities.

4.3. Addressing the Challenges and Future Research Avenues

To mitigate the limitations of social-media-based disaster analytics using transformer technology, such as sycophantic behavior, hallucinations, and ethical concerns, a multi-faceted strategy is essential. First, incorporating multi-modal data analysis that combines text with geographical and temporal data can reduce reliance on potentially biased social media narratives, enhancing the accuracy of disaster assessments. Implementing cross-validation techniques with authoritative external data sources, such as satellite imagery and official disaster reports, can help identify and correct hallucinations or exaggerated claims.

To address ethical and privacy concerns, deploying anonymization techniques and differential privacy measures ensures sensitive personal information is protected, while establishing a robust ethical framework for data use that includes transparent consent processes and bias mitigation algorithms. Furthermore, leveraging community feedback mechanisms can enhance the verification process, allowing for the correction of misinformation and the inclusion of diverse perspectives. This comprehensive approach not only addresses the current limitations but also strengthens the system's capacity to adapt to future challenges, ensuring that disaster analytics remains a reliable and ethical tool for crisis response and management.

5. Conclusions

The systematic literature review on the use of GPT technology for social-media-based disaster analysis yields significant insights into the current state of research in this burgeoning field. By meticulously analyzing 114 articles obtained through rigorous database searches following the PRISMA process, 53 highly relevant scholarly works were identified, shedding light on various aspects such as the geographical distribution of research, publication trends, and thematic domains. The developed novel classification scheme of 10 distinct categories (as shown in Figure 5) provides a structured understanding of the research landscape, facilitating the identification of trends, gaps, and emerging areas. Figures 6–10 demonstrate analytical insights into current body of knowledge within the scope of this study. Despite the potential, challenges such as sycophantic behavior, hallucinations, ethical considerations, and privacy concerns underscore the necessity for enhanced model robustness, refined data validation techniques, and the integration of human oversight mechanisms. Addressing these challenges through future innovations in transformer technologies could significantly enhance disaster analytics' effectiveness, reliability, and ethical integrity, offering promising avenues for advancing disaster management practices in our increasingly interconnected and data-driven world.

This research holds substantial relevance for multiple stakeholders. Policymakers and disaster management authorities can leverage the findings to enhance their understanding of social media's role in disaster monitoring and response, informing more effective strategies and interventions. Academic researchers benefit from the categorized literature, which serves as a foundational resource for further investigation and exploration of specific topics within the realm of GPT technology and disaster analysis. Furthermore, practitioners in fields such as natural disaster response, crisis communication, and humanitarian aid can derive practical insights from the reviewed literature to improve their operational capabilities and decision-making processes.

Future innovations in transformer technologies are poised to significantly enhance disaster analytics through improved data synthesis, real-time information processing, and ethical data use. Enhanced models could offer more accurate, nuanced interpretations of complex social media data, leading to better identification of emerging crises and more targeted disaster response efforts. Additionally, advancements in ethical AI and privacy-preserving techniques will address concerns around data sensitivity and bias, ensuring that disaster analytics tools not only become more powerful but also more responsible and inclusive, ultimately leading to more effective and equitable disaster management and response strategies. Additionally, interdisciplinary collaborations between computer scientists, social scientists, and domain experts may yield innovative approaches for harnessing GPT technology's potential while ensuring the reliability and ethical integrity of disaster analysis efforts. By integrating the technical expertise of computer scientists with the contextual insights of social scientists and the specialized knowledge of domain experts, these collaborations can significantly enhance model accuracy, relevance, and ethical integrity. This comprehensive approach ensures that AI-driven disaster response systems are not only technically proficient but also socially responsible and aligned with the nuanced needs of affected communities. Such advancements hold promise for advancing the field and facilitating more resilient and responsive disaster management practices in an increasingly interconnected world.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Available upon request.

Acknowledgments: Special appreciation is reserved for the COEUS Institute, Maine, US (https://coeus.institute/ accessed on 3 February 2024), where the author serves as the Chief Technology Officer, for fostering an environment conducive to research excellence. The institute's commitment to advancing knowledge in AI and pattern discovery has significantly enriched the author's endeavors in conducting this systematic review.

Conflicts of Interest: The author declares no conflicts of interest.

Appendix A

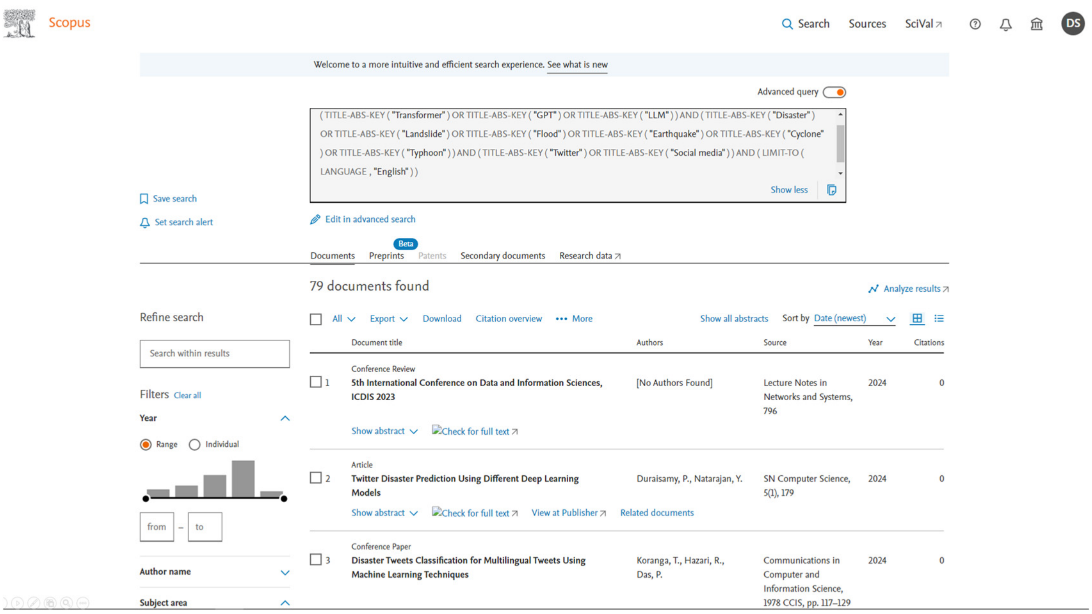

Figure A1. Advanced query result returned 79 documents from Scopus.

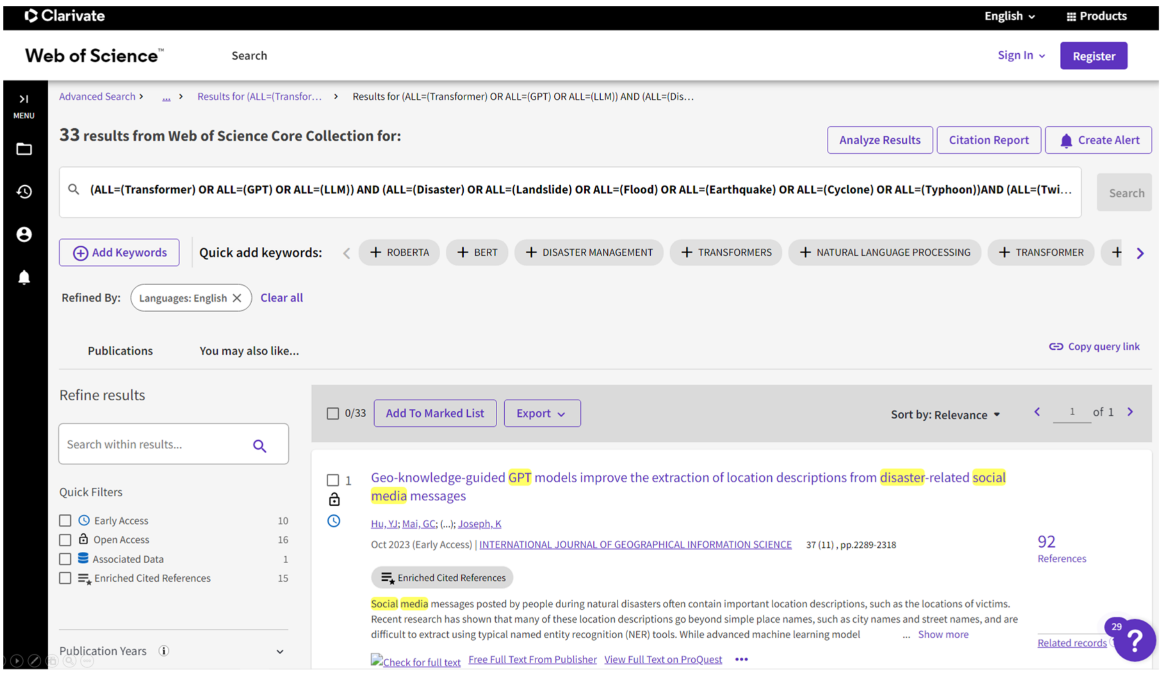

Figure A2. Advanced query result returned 33 documents from Web of Science.

Table A1. Comparison of transformer technology used on social media for "Disaster Event Detection and Classification".

Ref.	Area of Disaster	Algorithm Used	GPT/Transformer Technology	Benefit	Disadvantage
[18]	Flood (DKI Jakarta, Indonesia)	BERT-MLP	BERT	High accuracy (82%) in classifying tweets related to flood events using geospatial data.	Stemming process may remove important features, affecting accuracy.
[20]	Wildfires in the Western United States (2020)	BERTopic, NER	BERT	Real-time estimation of wildfire situations through social media analysis for decision support.	Potential noise from broad search terms, inaccuracies in user location data, and single-topic document assumption.
[21]	Twitter-Based Disaster Prediction	Improved BERT model, LSTM, GRU	BERT	Demonstrates superior accuracy in disaster prediction on Twitter by analyzing patterns associated with various types of disasters. Outperforms traditional models like LSTM and GRU in predicting disaster-related tweets.	Not specified.
[22]	Epidemics, Social Unrest, and Disasters	Enhanced BERT model, GloVe for feature extraction, LSTM for classification	BERT	Superior in terms of accuracy, precision, recall, and F1-score, capturing tweet semantics to accurately identify trend-related tweets.	Not specified.
[23]	Detection and Classification of Natural Disasters from Social Media	FNN, CNN, BLSTM, BERT, DistilBERT, Albert, RoBERTa, DeBERTa	BERT and its variants	Demonstrated the effectiveness of deep learning methods in accurately detecting and classifying disaster-related information from tweets, with preprocessing and bias mitigation enhancing performance.	The complexity of models and the need for extensive preprocessing and bias mitigation to handle the diverse and noisy nature of social media data.
[24]	Various Natural Disasters (Wildfires, Hurricanes, Earthquakes, Floods)	RoBERTa for text analysis, Vision Transformer for image understanding, Bi-LSTM for sequence processing, and attention mechanism for context awareness	RoBERTa, Vision Transformer	Superior performance with accuracy levels ranging from 94% to 98%, effective combination of textual and visual inputs through multimodal fusion.	Requires substantial computational and memory resources, potential hardware limitations.
[25]	Various Natural Disasters (Earthquakes, Floods, Hurricanes, Fires)	DenseNet, BERT	BERT for textual features, DenseNet for image features	Achieved an accuracy of 85.33% in classifying social media data into useful and non-useful categories for disaster response, outperforming state-of-the-art techniques.	Not specified, but potential issues could include the complexity of integrating multimodal data and the need for substantial computational resources.

Table A1. *Cont.*

Ref.	Area of Disaster	Algorithm Used	GPT/Transformer Technology	Benefit	Disadvantage
[26]	COVID-19 Pandemic Sentiment Analysis	RETN model, BERT-GRU, BERT-biLSTM, nature-inspired optimization techniques	GPT-2, GPT-3	Enhanced performance in sentiment analysis on large-scale datasets, including text, images, and audio.	Complexity in implementation and optimization.
[27]	Cyclone-Related Tweets	BERT, machine learning, and deep learning classifiers	BERT	BERT model achieves better results than other ML and DL models even on small labelled datasets.	Tweets often contain ambiguity and informal language that are hard for the machine to understand
[28]	Various (e.g., Earthquakes, Typhoons)	RACLC for classification, RACLTS for summarization	BERTweet (a variant of BERT for Twitter data)	High performance in disaster tweet classification and summarization. Provides interpretability through rationale extraction, enhancing trust in model decisions.	Potential limitations include the need for extensive training data and the challenge of adapting to new or unforeseen disaster types.
[29]	Various (e.g., Shootings, Hurricanes, Floods)	Transformer-based multi-task learning (MTL)	BERT, DistilBERT, ALBERT, ELECTRA	Demonstrates superior performance in classifying and prioritizing disaster-related tweets using a multi-task learning approach. Allows for effective handling of large volumes of data during crises.	Challenges include handling the high variability of disaster-related data and the computational demands of processing large datasets in real-time.
[30]	Disaster Detection on Twitter	BERT (Bidirectional Encoder Representations from Transformers) with keyword position information	BERT	Improved accuracy in disaster prediction on Twitter by incorporating keyword position information into the BERT model.	Relies heavily on the keyword position, which may not always accurately reflect the context or importance of a tweet.
[31]	General Disaster Management	Various BERT-based models (default BERT, BERT + NL, BERT + LSTM, BERT + CNN)	BERT	Effective in classifying disaster-related tweets by using balanced datasets and preprocessing techniques.	Challenges with imbalanced data and processing informal social media text.
[32]	Various Crises and Emergencies Detected via Social Media	MobileBERT for feature extraction, SSA improved with MRFO for feature selection	MobileBERT	High accuracy and efficiency in detecting and classifying crisis-related events on social media, leveraging advanced transformer technology and optimized feature selection.	Challenges include computational demands for processing and analyzing large-scale social media data in real-time and adapting to diverse and evolving crisis scenarios.
[33]	Flood-Related Volunteered Geographic Information (VGI)	BERT with TF-IDF, TextRank, MMR, LDA	BERT	Provides an ensemble approach combining BERT with traditional NLP methods for enhanced topic classification accuracy in flood-related microblogs.	Complexity in integrating multiple algorithms and potential challenges in scalability and real-time processing.

Table A1. *Cont.*

Ref.	Area of Disaster	Algorithm Used	GPT/Transformer Technology	Benefit	Disadvantage
[34]	Electricity Infrastructure	BERT	BERT for text classification	Capable of sensing the temporal evolutions and geographic differences of electricity infrastructure conditions through social media analysis.	Limited capability in areas with few Twitter activities, reliance on geotagged tweets, which are a small portion of total tweets.
[35]	Transportation Disaster Detection and Classification in Nigeria	BERT with AdamW optimizer	BERT	Improved accuracy in identifying and classifying transportation disaster tweets with an accuracy of 82%, outperforming existing algorithms.	Relies on named entity recognition (NER) for location identification, which may not be effective if users do not specify their location accurately.
[36]	Disaster Prediction from Tweets	GloVe embeddings for word representation, BERT for classification	BERT	Achieved 87% accuracy in classifying tweets related to disasters, showing BERT's superiority over traditional models like LSTM, random forest, decision trees, naive Bayes.	Requires significant preprocessing and understanding of NLP concepts to implement effectively.
[37]	Natural Disaster Tweet Classification	CNN with BERT embedding	BERT	Achieved high accuracy (97.16%), precision (97.63%), recall (96.64%), and F1-score (97.13%) in classifying natural disaster tweets.	Requires complex preprocessing and might overfit after certain epochs, indicating a need for careful model training and validation setup.
[38]	Analysis and Classification of Disaster Tweets from a Metaphorical Perspective	BERT, RoBERTa, DistilBERT	BERT, RoBERTa, DistilBERT	Demonstrated improved performance in classifying disaster-related tweets, including those with metaphorical contexts, highlighting the models' ability to capture metaphorical text representations effectively.	The study did not specifically address the computational efficiency or potential limitations in processing metaphorical content across diverse disaster types and languages.
[39]	Various (e.g., Earthquakes, Floods, Shootings)	Transformer-based model with multitask learning approach, including a fine-tuned encoder based on RoBERTa and transformer layers as a task adapter	RoBERTa	Significant improvements in classifying and prioritizing tweets in emergency situations by leveraging entities, event descriptions, and hashtags. This approach benefits from the adaptability of Transformers to handle noisy social media data.	The complexity of the model requires substantial computational resources for training and fine-tuning. The effectiveness of the model is dependent on the quality and representation of the input data, including the preprocessing of hashtags and the augmentation with event metadata.
[40]	Various (e.g., Earthquakes, Floods)	BERT, GRU, LSTM	BERT	BERT achieved the highest accuracy (96.2%) in classifying disaster-related tweets, indicating its effectiveness in understanding and categorizing disaster information from social media.	Increased complexity of the BERT architecture may lead to overfitting and requires careful adjustment.

Table A2. Comparison of transformer technology used on social media for "Sentiment Analysis and Public Perception".

Ref.	Area of Disaster	Algorithm Used	GPT/Transformer Technology	Benefit	Disadvantage
[41]	General Natural Disasters	Transformer Network	Transformer	Accurate analysis of public sentiment towards natural disasters	Focuses more on sentiment analysis rather than direct disaster response strategies
[42]	COVID-19 Vaccine Sentiment Analysis and Symptom Reporting from Tweets	BERT, Word2Vec	BERT	High accuracy in classifying sentiments towards COVID-19 vaccines and reporting of symptoms, leveraging contextual embeddings for nuanced understanding	Word2Vec showed lower performance compared to BERT, indicating fixed embeddings may not capture contextual nuances effectively
[43]	Climate Change	LDA, BERT	BERT for sentiment analysis	Effective in topic modeling and sentiment analysis with high precision (91.35%), recall (89.65%), and accuracy (93.50%)	Not explicitly mentioned, but potential challenges include processing vast datasets and identifying nuanced sentiment accurately
[44]	Public Opinions on Climate Change on Twitter	Convolutional Neural Network (CNN)	BERT pre-trained model	High accuracy in detecting believers (98%) and deniers (90%) of climate change, useful for smart city governance	Difficulty in collecting and labeling diverse Twitter data due to variations in human dialect and speech
[19]	Emotion Identification During COVID-19	Average Voting Ensemble Deep Learning model (AVEDL Model) incorporating BERT, DistilBERT, RoBERTa	BERT, DistilBERT, RoBERTa	Achieved high accuracy (86.46%) and macro-average F1-score (85.20%) in classifying emotions from COVID-19-related social media and emergency response calls, showcasing effective emotion analysis in pandemic conditions	The model's performance is contingent on the quality and size of the dataset, and its application is limited without extensive preprocessing and understanding of NLP concepts for accurate emotion extraction
[45]	Floods	RoBERTa, VADER, LSTM, CLIP	RoBERTa (Transformer and CLIP models)	Enhanced flood event detection through social media analysis, improving disaster response	Limited by the availability of specific dataset details

Table A3. Comparison of transformer technology used on social media for "Information Summarization and Retrieval".

Ref.	Area of Disaster	Algorithm Used	GPT/Transformer Technology	Benefit	Disadvantage
[46]	General Crisis Situations	Cross-lingual method for retrieving and summarizing crisis-relevant information from social media, utilizing multilingual transformer embeddings for summarization (T5)	T5 for summarization	Enables effective summarization of crisis-relevant information across multiple languages, enhancing situational awareness	Complexity in handling multilingual data and the potential for reduced accuracy in cross-lingual information retrieval and summarization
[47]	Crisis Event Social Media Summarization	NeuralSearchX for document retrieval, GPT-3 for summarization	GPT-3	High comprehensiveness in generated summaries of emergency events from social media and online news, rapid deployment due to few-shot learning	High redundancy ratio in the generated summaries, indicating potential information repetition
[48]	Various Disasters	SVM, BART	BART for summarization	Effective summarization of disaster-related tweets, differentiation between authoritative and user-generated content	Challenges in verifying the authenticity of information from user-generated content
[28]	Various (e.g., Earthquakes, Typhoons)	RACLC for classification, RACLTS for summarization	BERTweet (a variant of BERT for Twitter data)	High performance in disaster tweet classification and summarization. Provides interpretability through rationale extraction, enhancing trust in model decisions	The document does not explicitly list disadvantages, but potential limitations could include the need for extensive training data and the challenge of adapting to new or unforeseen disaster types
[33]	Flood-Related Volunteered Geographic Information (VGI)	BERT with TF-IDF, TextRank, MMR, LDA	BERT	Provides an ensemble approach combining BERT with traditional NLP methods for enhanced topic classification accuracy in flood-related microblogs	Complexity in integrating multiple algorithms and potential challenges in scalability and real-time processing
[34]	Electricity Infrastructure	BERT	BERT for text classification	Capable of sensing the temporal evolutions and geographic differences of electricity infrastructure conditions through social media analysis	Limited capability in areas with few Twitter activities, reliance on geotagged tweets, which are a small portion of total tweets

Table A4. Comparison of transformer technology used on social media for "Location Identification and Description Extraction".

Ref.	Area of Disaster	Algorithm Used	GPT/Transformer Technology	Benefit	Disadvantage
[49]	Location Identification from Textual Data	BERT-BiLSTM-CRF	BERT	High accuracy in recognizing toponyms, enhancing location identification in disaster communications	Focus on technical aspects of toponym recognition without direct disaster management application examples
[50]	Extraction of Location from Disaster-Related Social Media Posts	Geo-knowledge-guided approach, fusion of geo-knowledge and GPT models	GPT models such as ChatGPT and GPT-4	Significantly improves the accuracy of extracting location descriptions from social media messages by over 40% compared to NER approaches, requiring only a small set of examples encoding geo-knowledge	The approach's effectiveness is contingent on the availability and quality of geo-knowledge about common forms of location descriptions, which may vary by region and disaster type
[51]	Crisis Communication	BERT, CNN, MLP, LSTM, Bi-LSTM	BERT	BERT outperforms traditional and other deep learning models in crisis tweet classification	Not explicitly mentioned, but potential challenges include handling diverse data quality and the dynamic nature of social media language

Table A5. Comparison of transformer technology used on social media for "Tweet Prioritization and Useful Information Extraction".

Ref.	Area of Disaster	Algorithm Used	GPT/Transformer Technology	Benefit	Disadvantage
[52]	Earthquake Disasters	Feature-based, BLSTM-based, BERT-based, RoBERTa-based	BERT, BLSTM, RoBERTa	Provides methods to calculate usefulness ratings of tweets with behavioral facilitation information, with BERT achieving the best accuracy	Limited to tweets, may require integration of methods to assess tweet credibility
[53]	Harvesting Rescue Requests in Disaster Response from Social Media	BERT, GloVe, ELMo, RoBERTa, DistilBERT, ALBERT, XLNet, with classifiers like CNN, LSTM	BERT and its variations	Significantly increased accuracy in categorizing rescue-related tweets with the best model (a customized BERT-based model with a CNN classifier) achieving an F1-score of 0.919, which outperforms the baseline model by 10.6%	The complexity of models and computational costs, with the need for large and diverse training datasets to achieve high performance and stability
[54]	Flood Prediction from Twitter Data and Image Analysis	BMLP, SDAE, HHNN (hyperbolic Hopfield neural network), rule-based matching	BERT for text preprocessing	Achieved high accuracy (97%), precision (95%), recall (96%), and F1 score (96%) in predicting flood levels from Twitter data and images, effectively addressing semantic information loss and enhancing classification accuracy	Complexity in model implementation and reliance on extensive datasets for training. Requires sophisticated preprocessing to handle text and image data effectively

Table A5. *Cont.*

Ref.	Area of Disaster	Algorithm Used	GPT/Transformer Technology	Benefit	Disadvantage
[55]	Disaster-Related Multilingual Text Classification	GNoM (graph neural network enhanced language models)	BERT, mBERT, XLM-RoBERTa	Outperforms state-of-the-art models in disaster domain across monolingual, cross-lingual, and multilingual settings with improved F1 scores	Not explicitly mentioned, but complexity and integration of GNN with transformer models might introduce computational overhead
[56]	Uttarakhand Floods 2021	BERT, k-means, USE (universal sentence encoder)	BERT for tweet classification, USE for semantic similarity in clustering	Effective retrieval and prioritization of critical information from Twitter for emergency management, aiding timely disaster response	Not explicitly mentioned, but challenges may include processing vast amounts of social media data and the accuracy of critical information extraction
[57]	Flash Floods	FF-BERT: a multi-label text classification model	BERT (bidirectional encoder representations from transformers)	Enhances existing databases by classifying and categorizing information about flash flood events from web data	The primary limitation lies in its relatively low prediction performance for minority labels despite improvements over the baseline model
[29]	Various (e.g., Shootings, Hurricanes, Floods)	Transformer-based multi-task learning (MTL)	BERT, DistilBERT, ALBERT, ELECTRA	Demonstrates superior performance in classifying and prioritizing disaster-related tweets using a multi-task learning approach. Allows for effective handling of large volumes of data during crises	Challenges include handling the high variability of disaster-related data and the computational demands of processing large datasets in real-time.
[31]	General Disaster Management	BERT, BERT + NL, BERT + LSTM, BERT + CNN	BERT	Effective in classifying disaster-related tweets by using balanced datasets and preprocessing techniques	Challenges with imbalanced data and processing informal social media text.
[58]	Informative Tweet Prediction for Disasters	Deep learning architecture (transformer), semantic similarity models, logistic regression, glowworm optimization	Transformers	High precision in identifying informative tweets with the integration of disaster ontology and metadata classification	Complexity of the model and need for large datasets for effective training

Table A6. Comparison of transformer technology used on social media for "Multimodal Data Analysis".

Ref.	Area of Disaster	Algorithm Used	GPT/Transformer Technology	Benefit	Disadvantage
[59]	Flood Detection via Twitter Streams	Multimodal bitransformer model for text and image, pretrained Italian BERT model for text, VGGNet16 and ResNet152 for images	Not specified	Highest micro F1-score achieved with multimodal approach (0.859 for development set), demonstrating the effectiveness of combining textual and visual features	Specific performance metrics for individual modalities (text or image alone) were lower compared to the multimodal approach, indicating potential limitations in single modality analyses
[60]	Flood Detection via Twitter	Multimodal bitransformer model for text and image, pretrained Italian BERT model for text, VGGNet and ResNet for images	Not specified	High micro F1-score (0.859) achieved with multimodal approach, showing effectiveness in flood event detection combining textual and visual information	The complexity of integrating and optimizing multimodal data inputs for real-time analysis
[61]	General Disaster-Related Tweets on Social Media	Visual and linguistic double transformer fusion model (VLDT)	ALBERT for text, S-CBAM-VGG for visuals	Effective fusion of textual and visual data, leading to more accurate informative tweet classification	Potential challenges with feature extraction generalizability to new disaster types not in training data
[62]	Disaster Management	BERT	Vision Transformer (ViT)	Improved feature extraction for image processing, attention mechanism for relevance focus	CNN's limitation to understand feature relations, computing efficiency decreases with large kernels
[63]	Henan Heavy Storm (2021)	LDA, BERT, VGG-16	BERT for text classification; VGG-16 for image classification	High accuracy (0.93) in classifying natural disaster topics from social media data, enabling real-time understanding of disaster themes for informed decision making	Complexity of the multimodal analysis process and reliance on extensive data preprocessing and manual labeling for accurate model training
[45]	Floods	RoBERTa, VADER, LSTM, CLIP	RoBERTa (Transformer and CLIP models)	Enhanced flood event detection through social media analysis, improving disaster response.	Limited by the availability of specific dataset details.

Table A7. Comparison of transformer technology used on social media for "Multilingual and Cross-lingual Disaster Analysis".

Ref.	Area of Disaster	Algorithm Used	GPT/Transformer Technology	Benefit	Disadvantage
[64]	Wildfires, Earthquakes, Floods, Typhoons/Hurricanes, Bombings, Shootings	Fine-tuned RoBERTa-based encoder and Transformer blocks; bag of words for priority classification	RoBERTa, Transformer	Achieved top performance in tweet prioritization and surpassed median performance for information type classification by leveraging pre-trained language models and highlighting entities and hashtags	Not specified
[55]	Disaster-Related Multilingual Text Classification	GNoM (graph neural network enhanced language models)	BERT, mBERT, XLM-RoBERTa	Outperforms state-of-the-art models in disaster domain across monolingual, cross-lingual, and multilingual settings with improved F1 scores	Not explicitly mentioned, but complexity and integration of GNN with transformer models might introduce computational overhead

Table A8. Comparison of transformer technology used on social media for "Performance Evaluation and Comparison of Models".

Ref.	Area of Disaster	Algorithm Used	GPT/Transformer Technology	Benefit	Disadvantage
[65]	Various Disasters	BERT, DistilBERT, T5	Transformer-based question answering techniques	Precision of 0.81, recall of 0.76, and F-score of 0.78 for BERT in providing appropriate guidelines	Limited test data representing diverse crisis scenarios and issues with handling massive datasets due to token limitations
[66]	Various Disasters	Naive Bayes, Logistic Regression, Random Forest, SVM, KNN, Gradient Boosting, Decision Tree, LSTM, BiLSTM, CNN, BERT, DistilBERT	BERT and DistilBERT for tweet classification	High accuracy in multilingual tweet classification for disaster response	Not explicitly mentioned, but potential issues could include data sparsity and language-specific challenges
[30]	Disaster Detection on Twitter	BERT (Bidirectional Encoder Representations from Transformers) with Keyword Position Information	BERT	Improved accuracy in disaster prediction on Twitter by incorporating keyword position information into the BERT model	Relies heavily on the keyword position, which may not always accurately reflect the context or importance of a tweet
[32]	Various Crises and Emergencies Detected via Social Media	MobileBERT for Feature Extraction, SSA Improved with MRFO for Feature Selection	MobileBERT	High accuracy and efficiency in detecting and classifying crisis-related events on social media, leveraging advanced transformer technology and optimized feature selection	The document does not explicitly list disadvantages, but challenges could include computational demands for processing and analyzing large-scale social media data in real-time and adapting to diverse and evolving crisis scenarios
[67]	Various Natural Disasters	SMDKGG framework, AdaBoost classifier, STM, LOD Cloud, NELL, DBPedia, CYC, TSS, NGD, Chemical Reaction Optimization	Transformers for metadata classification	High precision (96.19%), recall (98.33%), and accuracy (97.26%) in generating knowledge graphs from disaster tweets, utilizing a comprehensive metadata-driven approach and diverse knowledge sources for enriched auxiliary knowledge	Complexity in the integration and processing of multiple data sources and algorithms, requiring extensive computational resources and expertise in machine learning and natural language processing

Table A8. *Cont.*

Ref.	Area of Disaster	Algorithm Used	GPT/Transformer Technology	Benefit	Disadvantage
[40]	Various (e.g., Earthquakes, Floods)	BERT, GRU, LSTM	BERT	BERT achieved the highest accuracy (96.2%) in classifying disaster-related tweets, indicating its effectiveness in understanding and categorizing disaster information from social media	Increased complexity of the BERT architecture may lead to overfitting and requires careful adjustment
[68]	Flood Forecasting	ODLFF-BDA, BERT, GRU, MLCNN, Equilibrium Optimizer (EO)	BERT for emotive contextual embedding from tweets	High accuracy and low memory usage for flood prediction using Twitter data	Not mentioned explicitly, but complexity and potential overfitting can be inferred as disadvantages

Table A9. Comparison of transformer technology used on social media for "Practical Applications and System Development".

Ref.	Area of Disaster	Algorithm Used	GPT/Transformer Technology	Benefit	Disadvantage
[70]	Disaster Support via Chatbot	Dual Intent Entity Transformer (DIET) for NLU, RASA for Conversation Management	Not specified directly, but DIET and RASA use underlying transformer models	Provides real-time disaster support and information dissemination in Portuguese, improving situational awareness and decision making	Limited by the specificity of its training data and potentially the depth of its knowledge base, requiring ongoing updates and expansions to remain effective in diverse disaster scenarios
[69]	Twitter Disaster Detection	BERT Variants (ELECTRA, Talking Head, TN-BERT), CNN, NN, LSTM, Bi-LSTM	Not specified	Good performance with F-scores between 76% and 80% and AUC between 86% and 90%, demonstrating effectiveness in disaster detection from Twitter data	Only marginally different performances among models, indicating a need for further exploration to identify the most efficient algorithm
[45]	Floods	RoBERTa, VADER, LSTM, CLIP	RoBERTa (transformer and CLIP models)	Enhanced flood event detection through social media analysis, improving disaster response	Limited by the availability of specific dataset details
[68]	Flood Forecasting	ODLFF-BDA, BERT, GRU, MLCNN, Equilibrium Optimizer (EO)	BERT for emotive contextual embedding from tweets	High accuracy and low memory usage for flood prediction using Twitter data	Not mentioned explicitly, but complexity and potential overfitting can be inferred as disadvantages

References

1. Wang, H.; Nie, D.; Tuo, X.; Zhong, Y. Research on crack monitoring at the trailing edge of landslides based on image processing. *Landslides* **2020**, *17*, 985–1007. [CrossRef]
2. Amatya, P.; Kirschbaum, D.; Stanley, T.; Tanyas, H. Landslide mapping using object-based image analysis and open source tools. *Eng. Geol.* **2021**, *282*, 106000. [CrossRef]
3. Rabby, Y.W.; Li, Y. Landslide inventory (2001–2017) of Chittagong hilly areas, Bangladesh. *Data* **2020**, *5*, 4. [CrossRef]
4. Sufi, F.K.; Alsulami, M. Knowledge Discovery of Global Landslides Using Automated Machine Learning Algorithms. *IEEE Access* **2021**, *9*, 131400–131419. [CrossRef]
5. Tamizi, A.; Young, I.R. A dataset of global tropical cyclone wind and surface wave measurements from buoy and satellite platforms. *Sci. Data* **2024**, *11*, 106. [CrossRef]
6. Sufi, F.K.; Khalil, I. Automated Disaster Monitoring From Social Media Posts Using AI-Based Location Intelligence and Sentiment Analysis. *IEEE Trans. Comput. Soc. Syst.* **2022**. [CrossRef]
7. Sufi, F.K. AI-SocialDisaster: An AI-based software for identifying and analyzing natural disasters from social media. *Softw. Impacts* **2022**, *13*, 100319. [CrossRef]
8. Sufi, F. A decision support system for extracting artificial intelligence-driven insights from live twitter feeds on natural disasters. *Decis. Anal. J.* **2022**, *5*, 100130. [CrossRef]
9. Sufi, F. A New Social Media Analytics Method for Identifying Factors Contributing to COVID-19 Discussion Topics. *Information* **2023**, *14*, 545. [CrossRef]
10. Sufi, F. Automatic identification and explanation of root causes on COVID-19 index anomalies. *MethodsX* **2023**, *10*, 101960. [CrossRef]
11. Poulsen, S.; Sarsa, S.; Prather, J.; Leinonen, J.; Becker, B.A.; Hellas, A.; Denny, P.; Reeves, B.N. Solving Proof Block Problems Using Large Language Models. In Proceedings of the SIGCSE 2024, Portland, OR, USA, 20–23 March 2024; Volume 7. [CrossRef]
12. Orrù, G.; Piarulli, A.; Conversano, C.; Gemignani, A. Human-like problem-solving abilities in large language models using ChatGPT. *Front. Artif. Intell.* **2023**, *6*, 1199350. [CrossRef]
13. Kieser, F.; Wulff, P.; Kuhn, J.; Küchemann, S. Educational data augmentation in physics education research using ChatGPT. *Phys. Rev. Phys. Educ. Res.* **2023**, *19*, 020150. [CrossRef]
14. Gusenbauer, M.; Haddaway, N.R. Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. *Res. Synth. Methods* **2020**, *11*, 181–217. [CrossRef]
15. Halevi, G.; Moed, H.; Bar-Ilan, J. Suitability of Google Scholar as a source of scientific information and as a source of data for scientific evaluation—Review of the Literature. *J. Inf.* **2017**, *11*, 823–834. [CrossRef]
16. Kaur, A.; Gulati, S.; Sharma, R.; Sinhababu, A.; Chakravarty, R. Visual citation navigation of open education resources using Litmaps. *Libr. Hi Tech News* **2022**, *39*, 7–11. [CrossRef]
17. Sufi, F. Generative Pre-Trained Transformer (GPT) in Research: A Systematic Review on Data Augmentation. *Information* **2024**, *15*, 99. [CrossRef]
18. Maulana, I.; Maharani, W. Disaster Tweet Classification Based on Geospatial Data Using the BERT-MLP Method. In Proceedings of the 2021 9th International Conference on Information and Communication Technology, ICoICT 2021, Yogyakarta, Indonesia, 3–5 August 2021; pp. 76–81. [CrossRef]
19. Nimmi, K.; Janet, B.; Selvan, A.K.; Sivakumaran, N. Pre-trained ensemble model for identification of emotion during COVID-19 based on emergency response support system dataset. *Appl. Soft Comput.* **2022**, *122*, 108842. [CrossRef]
20. Ma, Z.; Li, L.; Yuan, Y.; Baecher, G.B. Appraising Situational Awareness in Social Media Data for Wildfire Response. In Proceedings of the ASCE Inspire 2023: Infrastructure Innovation and Adaptation for a Sustainable and Resilient World-Selected Papers from ASCE Inspire 2023, Arlington, VA, USA, 16–18 November 2023; pp. 289–297.
21. Duraisamy, P.; Natarajan, Y. Twitter Disaster Prediction Using Different Deep Learning Models. *SN Comput. Sci.* **2024**, *5*, 179. [CrossRef]
22. Duraisamy, P.; Duraisamy, M.; Periyanayaki, M.; Natarajan, Y. Predicting Disaster Tweets using Enhanced BERT Model. In Proceedings of the 7th International Conference on Intelligent Computing and Control Systems, ICICCS 2023, Madurai, India, 17–19 May 2023; pp. 1745–1749. [CrossRef]
23. Fontalis, S.; Zamichos, A.; Tsourma, M.; Drosou, A.; Tzovaras, D. A Comparative Study of Deep Learning Methods for the Detection and Classification of Natural Disasters from Social Media. In Proceedings of the 12th International Conference on Pattern Recognition Applications and Methods, Lisbon, Portugal, 22–24 February 2023; pp. 320–327. [CrossRef]
24. JayaLakshmi, G.; Madhuri, A.; Vasudevan, D.; Thati, B.; Sirisha, U.; Praveen, S.P. Effective Disaster Management Through Transformer-Based Multimodal Tweet Classification. *Rev. D'intelligence Artif.* **2023**, *37*, 1263–1272. [CrossRef]
25. Kamoji, S.; Kalla, M.; Joshi, C. Fusion of Multimodal Textual and Visual Descriptors for Analyzing Disaster Response. In Proceedings of the 2023 5th International Conference on Smart Systems and Inventive Technology, ICSSIT 2023, Tirunelveli, India, 23–25 January 2023; pp. 1614–1619. [CrossRef]
26. Kour, H.; Gupta, M.K. AI Assisted Attention Mechanism for Hybrid Neural Model to Assess Online Attitudes About COVID-19. *Neural Process. Lett.* **2023**, *55*, 2265–2304. [CrossRef]

27. Sharma, S.; Basu, S.; Kushwaha, N.K.; Kumar, A.N.; Dalela, P.K. Categorizing disaster tweets into actionable classes for disaster managers: An empirical analysis on cyclone data. In Proceedings of the 2021 International Conference on Electrical, Computer, Communications and Mechatronics Engineering, ICECCME 2021, Mauritius, Mauritius, 7–8 October 2021. [CrossRef]
28. Nguyen, T.H.; Rudra, K. Rationale Aware Contrastive Learning Based Approach to Classify and Summarize Crisis-Related Microblogs. In Proceedings of the 31st ACM International Conference on Information & Knowledge Management, Atlanta, GA, USA, 17–21 October 2022; pp. 1552–1562. [CrossRef]
29. Wang, C.; Lillis, D.; Nulty, P. Transformer-based Multi-task Learning for Disaster Tweet Categorisation Transformer-based Multi-task Learning for Disaster Tweet Categorisation. In Proceedings of the International ISCRAM Conference, Blacksburg, VA, USA, 23 May 2021; pp. 705–718. Available online: https://www.researchgate.net/publication/355367274 (accessed on 15 February 2024).
30. Wang, Z.; Zhu, T.; Mai, S. Disaster Detector on Twitter Using Bidirectional Encoder Representation from Transformers with Keyword Position Information. In Proceedings of the 2020 IEEE 2nd International Conference on Civil Aviation Safety and Information Technology, ICCASIT 2020, Weihai, China, 14–16 October 2020; pp. 474–477. [CrossRef]
31. Naaz, S.; Abedin, Z.U.; Rizvi, D.R. Sequence Classification of Tweets with Transfer Learning via BERT in the Field of Disaster Management. *EAI Endorsed Trans. Scalable Inf. Syst.* **2021**, *8*, e8. [CrossRef]
32. Dahou, A.; Mabrouk, A.; Ewees, A.A.; Gaheen, M.A.; Abd Elaziz, M. A social media event detection framework based on transformers and swarm optimization for public notification of crises and emergency management. *Technol. Forecast. Soc. Change* **2023**, *192*, 122546. [CrossRef]
33. Du, W.; Ge, C.; Yao, S.; Chen, N.; Xu, L. Applicability Analysis and Ensemble Application of BERT with TF-IDF, TextRank, MMR, and LDA for Topic Classification Based on Flood-Related VGI. *ISPRS Int. J. Geo-Inf.* **2023**, *12*, 240. [CrossRef]
34. Chen, Y.; Umana, A.; Yang, C.; Ji, W. Condition Sensing for Electricity Infrastructure in Disasters by Mining Public Topics from Social Media. In Proceedings of the International ISCRAM Conference, Blacksburg, VA, USA, 23 May 2021; pp. 598–608.
35. Prasad, R.; Udeme, A.U.; Misra, S.; Bisallah, H. Identification and classification of transportation disaster tweets using improved bidirectional encoder representations from transformers. *Int. J. Inf. Manag. Data Insights* **2023**, *3*, 100154. [CrossRef]
36. Ranade, A.; Telge, S.; Mate, Y. Predicting Disasters from Tweets Using GloVe Embeddings and BERT Layer Classification. In *International Advanced Computing Conference*; Communications in Computer and Information Science; Springer: Cham, Switzerland, 2022; pp. 492–503. [CrossRef]
37. Dharma, L.S.A.; Winarko, E. Classifying Natural Disaster Tweet using a Convolutional Neural Network and BERT Embedding. In Proceedings of the 2022 2nd International Conference on Information Technology and Education, ICIT and E 2022, Malang, Indonesia, 22 January 2022; pp. 23–30. [CrossRef]
38. Alcántara, T.; García-Vázquez, O.; Calvo, H.; Torres-León, J.A. Disaster Tweets: Analysis from the Metaphor Perspective and Classification Using LLM's. In *Mexican International Conference on Artificial Intelligence*; Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Springer: Cham, Switzerland, 2024; pp. 106–117. [CrossRef]
39. Boros, E.; Lejeune, G.; Coustaty, M.; Doucet, A. Adapting Transformers for De-tecting Emergency Events on Social Media. In Proceedings of the 14th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, IC3K-Proceedings, Valletta, Malta, 24–26 October 2022; pp. 300–306. [CrossRef]
40. Ullah, I.; Jamil, A.; Hassan, I.U.; Kim, B.S. Unveiling the Power of Deep Learning: A Comparative Study of LSTM, BERT, and GRU for Disaster Tweet Classification. *IEIE Trans. Smart Process. Comput.* **2023**, *12*, 526–534. [CrossRef]
41. Li, S.; Sun, X. Application of public emotion feature extraction algorithm based on social media communication in public opinion analysis of natural disasters. *PeerJ Comput. Sci.* **2023**, *9*, e1417. [CrossRef]
42. Bansal, A.; Jain, R.; Bedi, J. Detecting COVID-19 Vaccine Stance and Symptom Reporting from Tweets using Contextual Embeddings. In Proceedings of the CEUR Workshop, Kolkata, India, 9–13 December 2022; pp. 361–368. Available online: http://ceur-ws.org (accessed on 15 February 2024).
43. Uthirapathy, S.E.; Sandanam, D. Topic Modelling and Opinion Analysis on Climate Change Twitter Data Using LDA and BERT Model. *Procedia Comput. Sci.* **2022**, *218*, 908–917. [CrossRef]
44. Lydiri, M.; El Mourabit, Y.; El Habouz, Y.; Fakir, M. A performant deep learning model for sentiment analysis of climate change. *Soc. Netw. Anal. Min.* **2023**, *13*, 8. [CrossRef]
45. Bryan-Smith, L.; Godsall, J.; George, F.; Egode, K.; Dethlefs, N.; Parsons, D. Real-time social media sentiment analysis for rapid impact assessment of floods. *Comput. Geosci.* **2023**, *178*, 105405. [CrossRef]
46. Vitiugin, F.; Castillo, C. Cross-Lingual Query-Based Summarization of Crisis-Related Social Media: An Abstractive Approach Using Transformers. In Proceedings of the 33rd ACM Conference on Hypertext and Social Media, Barcelona, Spain, 28 June–1 July 2022; pp. 21–31. [CrossRef]
47. Pereira, J.; Fidalgo, R.; Nogueira, R. Crisis Event Social Media Summarization with GPT-3 and Neural Reranking. In Proceedings of the International ISCRAM Conference, Omaha, NE, USA, 28–31 May 2023; pp. 371–384. Available online: https://www.researchgate.net/publication/371038649 (accessed on 15 February 2024).
48. Sakhapara, A.; Pawade, D.; Dodhia, B.; Jain, J.; Bhosale, O.; Chakrawar, O. Summarization of Tweets Related to Disaster. In Proceedings of the International Conference on Recent Trends in Computing: ICRTC 2021; Lecture Notes in Networks and Systems. Springer: Singapore, 2022; pp. 651–665. [CrossRef]

49. Ma, K.; Tan, Y.J.; Xie, Z.; Qiu, Q.; Chen, S. Chinese toponym recognition with variant neural structures from social media messages based on BERT methods. *J. Geogr. Syst.* **2022**, *24*, 143–169. [CrossRef]
50. Hu, Y.; Mai, G.; Cundy, C.; Choi, K.; Lao, N.; Liu, W.; Lakhanpal, G.; Zhou, R.Z.; Joseph, K. Geo-knowledge-guided GPT models improve the extraction of location descriptions from disaster-related social media messages. *Int. J. Geogr. Inf. Sci.* **2023**, *37*, 2289–2318. [CrossRef]
51. Chandrakala, S.; Raj, S.A.A. Identifying the label of crisis related tweets using deep neural networks for aiding emergency planning. In Proceedings of the 2022 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems, ICSES 2022, Chennai, India, 15–16 July 2022. [CrossRef]
52. Yamamoto, F.; Kumamoto, T.; Suzuki, Y.; Nadamoto, A. Methods of Calculating Usefulness Ratings of Behavioral Facilitation Tweets in Disaster Situations. In Proceedings of the 11th International Symposium on Information and Communication Technology, Hanoi, Vietnam, 1–3 December 2022; pp. 88–95. [CrossRef]
53. Zhou, B.; Zou, L.; Mostafavi, A.; Lin, B.; Yang, M.; Gharaibeh, N.; Cai, H.; Abedin, J.; Mandal, D. VictimFinder: Harvesting rescue requests in disaster response from social media with BERT. *Comput. Environ. Urban Syst.* **2022**, *95*, 101824. [CrossRef]
54. Kamoji, S.; Kalla, M. Effective Flood prediction model based on Twitter Text and Image analysis using BMLP and SDAE-HHNN. *Eng. Appl. Artif. Intell.* **2023**, *123*, 106365. [CrossRef]
55. Ghosh, S.; Maji, S.; Desarkar, M.S. GNoM: Graph Neural Network Enhanced Language Models for Disaster Related Multilingual Text Classification. In Proceedings of the 14th ACM Web Science Conference 2022, Barcelona, Spain, 26–29 June 2022; pp. 55–65. [CrossRef]
56. Varshney, A.; Kapoor, Y.; Chawla, V.; Gaur, V. A Novel Framework for Assessing the Criticality of Retrieved Information. *Int. J. Comput. Digit. Syst.* **2022**, *11*, 1229–1244. [CrossRef]
57. Wilkho, R.S.; Chang, S.; Gharaibeh, N.G. FF-BERT: A BERT-based ensemble for automated classification of web-based text on flash flood events. *Adv. Eng. Inform.* **2024**, *59*, 102293. [CrossRef]
58. Arulmozhivarman, M.; Deepak, G. TPredDis: Most Informative Tweet Prediction for Disasters Using Semantic Intelligence and Learning Hybridizations. In *International Conference on Robotics, Control, Automation and Artificial Intelligence*; Lecture Notes in Electrical Engineering; Springer: Singapore, 2023; pp. 993–1002. [CrossRef]
59. Alam, F.; Hassan, Z.; Ahmad, K.; Gul, A.; Reiglar, M.; Conci, N.; Al-Fuqaha, A. Flood Detection via Twitter Streams using Textual and Visual Features. *arXiv* **2020**, arXiv:2011.14944. [CrossRef]
60. Wahid, J.A.; Shi, L.; Gao, Y.; Yang, B.; Wei, L.; Tao, Y.; Hussain, S.; Ayoub, M.; Yagoub, I. Topic2Labels: A framework to annotate and classify the social media data through LDA topics and deep learning models for crisis response. *Expert Syst. Appl.* **2022**, *195*, 116562. [CrossRef]
61. Zhou, J.; Wang, X.; Liu, N.; Liu, X.; Lv, J.; Li, X.; Zhang, H.; Cao, R. Visual and Linguistic Double Transformer Fusion Model for Multimodal Tweet Classification. In Proceedings of the 2023 International Joint Conference on Neural Networks, Gold Coast, Australia, 18–23 June 2023. [CrossRef]
62. Koshy, R.; Elango, S. Multimodal tweet classification in disaster response systems using transformer-based bidirectional attention model. *Neural Comput. Appl.* **2023**, *35*, 1607–1627. [CrossRef]
63. Zhang, M.; Huang, Q.; Liu, H. A Multimodal Data Analysis Approach to Social Media during Natural Disasters. *Sustainability* **2022**, *14*, 5536. [CrossRef]
64. Boros, E.; Nguyen, N.K.; Lejeune, G.; Coustaty, M.; Doucet, A. Transformer-based Methods with #Entities for Detecting Emergency Events on Social Media. In Proceedings of the 30th Text REtrieval Conference, TREC 2021-Proceedings, Online, 15–19 November 2021; Available online: http://trec.nist.gov (accessed on 15 February 2024).
65. Karam, E.; Hussein, W.; Gharib, T.F. Detecting needs of people in a crisis using Transformer-based question answering techniques. In Proceedings of the 2021 IEEE 10th International Conference on Intelligent Computing and Information Systems, ICICIS 2021, Cairo, Egypt, 5–7 December 2021; pp. 348–354. [CrossRef]
66. Koranga, T.; Hazari, R.; Das, P. Disaster Tweets Classification for Multilingual Tweets Using Machine Learning Techniques. In *International Conference on Computation Intelligence and Network Systems*; Communications in Computer and Information Science; Springer: Cham, Switzerland, 2024; pp. 117–129. [CrossRef]
67. Bhaveeasheshwar, E.; Deepak, G. SMDKGG: A Socially Aware Metadata Driven Knowledge Graph Generation for Disaster Tweets. In *International Conference on Applied Machine Learning and Data Analytics*; Communications in Computer and Information Science; Springer: Cham, Switzerland, 2023; pp. 64–77. [CrossRef]
68. Indra, G.; Duraipandian, N. Modeling of Optimal Deep Learning Based Flood Forecasting Model Using Twitter Data. *Intell. Autom. Soft Comput.* **2023**, *35*, 1455–1470. [CrossRef]
69. Balakrishnan, V.; Shi, Z.; Law, C.L.; Lim, R.; Teh, L.L.; Fan, Y.; Periasamy, J. A Comprehensive Analysis of Transformer-Deep Neural Network Models in Twitter Disaster Detection. *Mathematics* **2022**, *10*, 4664. [CrossRef]
70. Boné, J.; Ferreira, J.C.; Ribeiro, R.; Cadete, G. Disbot: A Portuguese disaster support dynamic knowledge chatbot. *Appl. Sci.* **2020**, *10*, 9082. [CrossRef]
71. Ranaldi, L.; Pucci, G. When Large Language Models contradict humans? Large Language Models' Sycophantic Behaviour. *arXiv* **2023**, arXiv:2311.09410. [CrossRef]
72. Ji, Z.; Lee, N.; Frieske, R.; Yu, T.; Su, D.; Xu, Y.; Ishii, E.; Bang, Y.J.; Madotto, A.; Fung, P. Survey of Hallucination in Natural Language Generation. *ACM Comput. Surv.* **2023**, *55*, 1–38. [CrossRef]

73. Sahoo, S.R.; Gupta, B.B. Real-time detection of fake account in twitter using machine-learning approach. In *Advances in Computational Intelligence and Communication Technology: Proceedings of CICT 2019*; Advances in Intelligent Systems and Computing; Springer: Singapore, 2021; pp. 149–159. [CrossRef]
74. Murayama, T.; Wakamiya, S.; Aramaki, E.; Kobayashi, R. Modeling the spread of fake news on Twitter. *PLoS ONE* **2021**, *16*, e0250419. [CrossRef]
75. Gustafson, D.L.; Woodworth, C.F. Methodological and ethical issues in research using social media: A metamethod of Human Papillomavirus vaccine studies. *BMC Med. Res. Methodol.* **2014**, *14*, 127. [CrossRef]