Machine Learning with Applications 24 (2026) 100882

Contents lists available at ScienceDirect

Machine Learning with Applications

journal homepage: www.elsevier.com/locate/mlwa

DisasterSense: Event classification in social media using RAKE-based pattern mining and dictionary-driven named entity recognition

Muhammad Salman Tiwana [a], Ali Braytee [a], Faezeh Karimi [a], Madhushi Bandara [a], Tony Jan [b], Mukesh Prasad [a,*]

[a] School of Computer Science, Faculty of Engineering and Technology, University of Technology Sydney, Australia
[b] Centre for Artificial Intelligence and Optimization, DCT, Torrens University

ARTICLE INFO

Keywords:
Social Media
Machine Learning
Deep Learning
Data Mining
Natural Language Processing
Name Entity Recognition

ABSTRACT

Natural disasters have a major effect on the Earth and cause severe repercussions. The use of social networks is a valuable method for gathering information and insight during such situations. Natural language processing techniques that use deep learning and machine learning models show strong potential for categorising disaster-related data. To enhance the classification and recognition of social media information related to natural and man-made disasters, this research introduces a novel approach, *DisasterSense*. Many existing detection tools misclassify man-made disasters as natural because social media texts often lack clear linguistic patterns. In this study, Dictionary-Driven Named Entity Recognition was applied to create a word-based disaster dictionary for categorising natural and man-made disasters, which was then combined with pattern mining to identify recurrent themes, patterns, and relationships among frequently used phrases. Using frequent patterns extracted with the Rapid Automatic Keyword Extraction (RAKE) model, an unsupervised technique for identifying keywords and phrases from text, a new pattern-based disaster dictionary was developed to enhance classification accuracy through pattern comparison. The effectiveness of the proposed approach was evaluated by comparing it with supervised and unsupervised models. This research aims to help emergency response organisations respond more rapidly and mitigate the impacts of disasters. Experiments on disaster-related data demonstrated that our approach achieved the highest accuracy of 0.789, outperforming all other well-known classification methods evaluated in this study.

1. Introduction

Social media platforms such as Twitter, Facebook, and Weibo generate continuous streams of real-time content that reflect public experiences, concerns, and behaviours (Cresci et al., 2015). During disaster events, these platforms become crucial channels for rapid information sharing, where individuals post updates, images, and requests for assistance. This abundance of user-generated information has positioned social media as an important resource in disaster informatics, enabling researchers and emergency agencies to enhance situational awareness and support timely decision-making (Widyastuti, 2021). Automated text classification, therefore, plays a central role in extracting relevant disaster information from large, unstructured datasets.

Despite extensive research, several limitations persist in disaster-related text classification. Many machine learning and deep learning models depend heavily on large labelled datasets, which are rarely available during the early stages of disasters (Sharma et al., 2022). Although hashtags are sometimes used as surrogate labels, they are often noisy, inconsistent, and insufficient for reliable classification. Traditional NLP approaches, such as n-grams, lexicons, and part-of-speech tagging, also struggle with the informal, multilingual, and irregular linguistic characteristics of social media content, reducing model robustness and generalisability (Anand et al., 2023). Furthermore, most existing studies primarily focus on sub-categorising natural disasters such as floods, earthquakes, hurricanes, or bushfires, while far fewer address the broader and operationally critical task of distinguishing between natural and man-made disasters (Bikku et al., 2022; Tounsi & Temimi, 2023). This gap limits the practical applicability of current models in real-world emergency management, where both categories must be identified rapidly to support effective response. A common

* Corresponding author at: School of Computer Science, Faculty of Engineering and Technology, University of Technology Sydney, Australia.
E-mail addresses: MuhammadSalman.Tiwana@uts.edu.au (M.S. Tiwana), ali.braytee@uts.edu.au (A. Braytee), Faezeh.karimi@uts.edu.au (F. Karimi), madhushi.bandara@uts.edu.au (M. Bandara), tony.jan@torrens.edu.au (T. Jan), Mukesh.prasad@uts.edu.au (M. Prasad).

https://doi.org/10.1016/j.mlwa.2026.100882

Available online 17 March 2026

challenge in disaster classification arises when events contain features of both natural and man-made causes. For instance, consider a report stating: *"The cargo ship sank during the storm after the crew failed to secure the loading bay."* Although the storm is a natural hazard, the primary cause of the incident is human negligence. Traditional classifiers may incorrectly label this as a natural disaster because of the presence of keywords such as *storm* or *extreme weather*, overlooking the human-induced component. This highlights the need for more context-aware methods capable of distinguishing natural events from man-made ones, even when both appear in the same narrative (Dominguez-Péry et al., 2021).

To address these challenges, this study proposes a multi-stage dictionary-driven classification framework. The approach begins by constructing a disaster dictionary using n-gram, bi-gram, and bag-of-words features extracted from pre-processed social media text. This dictionary is then expanded through dictionary-based named entity recognition (DB-NER), which automatically identifies disaster-related entities without requiring large, annotated datasets. To capture deeper contextual patterns, Rapid Automatic Keyword Extraction (RAKE) is applied to discover frequently occurring lexical structures, forming a complementary pattern-based dictionary. By integrating lexical, entity-level, and pattern-level information, the proposed model offers a scalable and data-efficient method for accurately distinguishing between natural and man-made disasters in dynamic social media environments.

The significance of this research lies in both its community impact and scientific advancement. From an emergency-response perspective, the framework enables rapid automated classification of disaster-related content, improving situational awareness and helping authorities identify urgent needs such as medical assistance, shelter, food, or infrastructure support based on real-time social media communication (Savale et al., 2024). From a scientific standpoint, the study addresses key challenges in disaster informatics by reducing dependence on manual data annotation, introducing an automated text-labelling strategy, and enhancing the efficiency and scalability of disaster information processing (Ro et al., 2024).

The major contributions of this work are as follows:

- Dictionary-driven framework capable of distinguishing natural and man-made disasters without requiring large, annotated datasets.
- DB-NER method enhances contextual understanding through automatic extraction of disaster-related entities.
- By incorporating RAKE-driven pattern mining, the approach extracts frequent lexical patterns that improve semantic representation.
- The dictionary-driven approach enhances robustness when processing noisy text and remains applicable to a broad range of textual data beyond social media.

The remainder of this paper is organised as follows. Section 2 reviews related literature on disaster text classification and dictionary-based NLP techniques. Section 3 describes the proposed methodology, including pre-processing, dictionary construction, DB-NER expansion, and pattern mining. Section 4 presents the classification algorithm, while Section 5 reports the experimental setup, evaluation metrics, and results. Section 6 concludes the paper and outlines directions for future research.

2. Literature review

The use of social media for real-time disaster prediction and classification, focusing particularly on techniques in natural language processing (NLP), machine learning (ML), and deep learning (DL). Twitter has been identified as a rapid source of situational awareness during disasters, with NLP pipelines developed to categorise tweets into actionable groups (Suwaileh et al., 2022). In the domain of machine learning, classifiers such as Support Vector Machines (SVM) and Random Forests have been employed to differentiate between informative and non-informative posts, often using feature extraction methods like Bag-of-Words (BoW) and Term Frequency-Inverse Document Frequency (TF-IDF) (Subramani et al., 2018). Recent progress in deep learning has drawn attention to models such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and BERT. LSTM networks, for instance, have been used to capture the temporal context of tweet sequences, leading to improved accuracy in disaster detection (Koshy & Elango, 2023). Similarly, BERT has been applied for fine-grained classification of disaster-related information, outperforming traditional machine learning techniques (Adesokan et al., 2023). Furthermore, frequent pattern mining has been used to extract key indicators associated with disasters, such as strong winds and heavy rainfall, helping distinguish relevant content from unrelated information (Yang et al., 2024).

2.1. Leveraging social media in crisis management

The frequency of natural and man-made disasters has increased, and they result in large losses in terms of both money and lives on a worldwide scale. Since it might be difficult to obtain pertinent information from disaster areas promptly, timely, data-driven decision-making is essential to effective disaster response. A viable remedy is crowdsourcing, especially via social media platforms, which can support disaster response operations by supplying real-time, human-generated data. During disasters, social media sites like Twitter provide real-time updates, but it can be difficult to analyse their unstructured data. This is addressed by classifying and extracting pertinent disaster-related data using machine learning techniques, such as deep learning and data mining. To enhance disaster management and response operations, techniques such as Sequential Pattern Mining and models like CRF and LSTM are used to find and categorise relevant content (Shidik et al., 2024).

For example, a woman was saved during Storm Harvey after requesting assistance over Twitter when the normal emergency line was unavailable (Battistoli et al., 2018). Numerous studies have examined social media's use in disaster management, emphasising its potential as an effective instrument for community involvement, communication, and information sharing. According to a recently published study (Bhuvana & Aram, 2019), they examined Twitter usage during the three major hurricanes that hit the United States, Isaac, Sandy, and Harvey. Social media is a useful tool for communication and for supplying real-time information for rescue and emergency response efforts. A methodology for classifying social media usage and issues in disaster resilience research and management with a particular focus on Twitter use during three significant hurricanes in the US. Based on their findings, social media data has four main benefits: it can be used as an efficient platform for communication, it can provide emergency response with ground truth information, it can provide insights into people's feelings, and it may facilitate predictive modelling (Lam et al., 2023). Expanding on these findings, researchers conducted a bibliometric analysis of social media applications in disaster management. Their study found three primary research clusters: information and knowledge exchange, post-disaster recovery procedures, and Twitter in disaster management. This report offers a thorough summary of the developing issues in research on social media and catastrophe management (Fauzi, 2023). The use of social media throughout the four stages of disaster risk management: prevention and mitigation, preparedness, relief and response, and recovery. This manually highlights the value of social media as an information tool and the need for dynamic communication during emergencies (Khan et al., 2022). Together, these studies highlight the diverse ways social media may help with disaster management, from promoting quick information sharing and raising situational awareness to aiding in long-term recovery and fostering community resilience.

2.2. Natural language processing

One area of artificial intelligence that focuses on how computers and human language interact is called natural language processing, or NLP (Kang et al., 2020). It entails creating models and algorithms that allow robots to comprehend, interpret, and produce meaningful and practical human language. Text categorisation, sentiment analysis, machine translation, and named entity recognition are important NLP activities that aid computers in effectively processing vast amounts of natural language data (Fanni et al., 2023). Large datasets, commonly known as "big data," are a major component of NLP, and a sizable amount of them come from online social media networks. For NLP tasks like text analysis, sentiment analysis, and named entity recognition, the wealth of user-generated content on social media offers rich and varied data (Hagiwara, 2021). Numerous research projects aimed at applying NLP techniques to social media content, including tweets, posts, and comments, have been prompted by the growing interest in the relationship between NLP and social media. Additionally, an approach suggested (Ketmaneechairat & Maliyaem, 2020), utilising Conditional Random Fields (CRF) and bidirectional Long Short-Term Memory (LSTM) models to extract named items from unstructured text and classify them into predetermined classes. A dataset of 1,000 words taken from natural disaster-related Twitter tweets and divided into six output categories was used in the trials. CRF, CRF-optimised, and a mix of LSTM and CRF were the three evaluation scenarios that were tested. Meanwhile, an experimental investigation was especially concerned with assessing Named Entity Recognisers' (NER) effectiveness when it came to obtaining location data from tweets (Lingad et al., 2013). The purpose of this study was to measure the effectiveness of current NER systems in locating and extracting locations that are referenced in tweets. Accurately identifying locations in social media text is essential across many domains, particularly disaster response, where real-time geolocation information can provide critical insights (Lovera & Cardinale, 2025). As NLP tools for social media analysis continue to advance, they offer expanding opportunities for both researchers and practitioners, enabling applications that range from location-based services to rapid crisis detection. Overall, existing research underscores the growing importance of NLP in disaster management. The studies (Albladi et al., 2025; Gardazi et al., 2025) demonstrate the effectiveness of NLP techniques in extracting actionable information from unstructured social media data. Methods such as named entity recognition, topic modelling, and sentiment analysis have shown considerable promise in supporting real-time crisis response. NLP advances, particularly in machine learning and deep learning, continue to improve the capacity to analyse and use disaster-related content for better informed decision making despite obstacles such as data noise and ambiguity.

2.3. Machine learning and deep learning

Machine learning (ML) has emerged as a potent instrument for classifying and examining natural disaster-related content on social media sites. Traditional manual monitoring techniques are no longer practical due to the enormous volume of real-time data provided through tweets, comments, and photos. To assist authorities in taking preventative action by showing how machine learning algorithms can scan vast amounts of tweets to find important disaster information like location, severity, and urgency. Machine learning can improve emergency responders' situational awareness and provide early warnings by identifying patterns in tweets. It does this by removing extraneous information to highlight important insights. To help with more efficient disaster response, this study uses five machine learning techniques to categorise tweets as either non-disaster or disaster-related. Three of these five machine learning algorithms function similarly; however, out of all of them, Logistic Regression has the best model accuracy (80.5%) (Islam et al., 2023). Likewise, supervised machine learning techniques were used to identify tweets regarding disasters. The linguistic and other statistical features that are specific to each tweet, such as its part of speech, user references, duration, quantity of hashtags, etc., are used by the majority of supervised machine learning techniques in this field (Habdank et al., 2017). One prominent work used Gated Recurrent Units (GRUs) and hybrid convolutional layers to categorise tweets about disasters. This study emphasises the significance of Natural Language Processing (NLP) in comprehending the subtleties of human language in social media situations (Pratama & Pardede, 2023). In the same way, tweets classification using convolutional neural networks (CNN) and BERT, highlights the potential of sophisticated machine learning techniques to glean insightful information from social media data (Girsang & Noveta, 2024). Additionally, using machine learning (ML) makes it easier to extract useful insights from unstructured data. To better map disaster occurrences, A study suggested a hybrid machine learning pipeline that incorporates NER to find areas cited in tweets (Fan et al., 2020). This method provides strategic planning for disaster response activities in addition to assisting with situational awareness in real time. Researchers and practitioners can improve readiness, reaction, and recovery efforts by using sophisticated algorithms and NLP approaches to glean insightful information from social media. As Deep learning techniques can use enormous datasets to enhance the accuracy and timeliness of forecasting disasters like hurricanes, floods, earthquakes, and wildfires, they have become indispensable in improving disaster management and prediction. These techniques allow for more accurate forecasts, quicker reaction times, and more efficient use of resources, which eventually saves lives and lowers financial losses. Deep learning models can simulate crises, identify patterns, and support real-time decision-making by using sensor data, social media feeds, and satellite imagery. A deep neural network (DNN) model was created to forecast typhoon-related building losses. The promise of deep learning in natural disaster risk assessment was demonstrated by their framework's increased accuracy when compared to conventional statistical models (Kim et al., 2023). Effective disaster planning depends on deep learning's capacity to interpret complex data and produce more accurate forecasts, as this study demonstrates. Building on this basis, a hybrid deep learning system that uses historical rainfall data to project flood occurrences by integrating convolutional neural networks (CNN) with long short-term memory networks. Improved flood risk management capabilities were shown by this strategy (Moishin et al., 2021). Combining several neural network topologies demonstrates how deep learning may be tailored to solve particular disaster prediction problems. Similarly, CNN was used to create flood sensitivity maps according to several environmental and geographic parameters. To evaluate flood risks, their model successfully examined intricate interactions between several factors (Khosravi et al., 2018). This application demonstrates how deep learning may be used to handle and interpret a variety of data sources to produce thorough risk evaluations. By examining continuous time series data from seismic stations, deep learning models were used for early earthquake identification and prediction, broadening the focus to include additional disaster types. They identified earthquake-prone areas using K-means clustering (Novianti et al., 2017). This study shows how deep learning can be used in early warning systems and real-time monitoring of seismic events. The potential of deep learning in post-disaster damage assessment was demonstrated in their research, and they used CNN-based analysis of satellite pictures to identify damage brought on by powerful hurricanes. This application demonstrates how deep learning may be used to quickly assess the aftermath of disasters in addition to predicting them, which is essential for efficient response and recovery efforts. Together, these studies demonstrate the wide range of uses of deep learning in disaster management, ranging from anticipating different types of natural catastrophes to evaluating damage and enhancing response tactics. Deep learning algorithms are probably going to become more and more important as the area develops, improving our capacity to anticipate, respond to, and lessen the effects of natural disasters.

Recent advances in machine learning and deep learning research further highlight the growing effectiveness of automated text analysis for classification tasks relevant to disaster informatics. Studies on COVID-19 misinformation detection have shown that sequential deep learning architectures such as BiGRU and LSTM outperform traditional models, with BiGRU achieving an accuracy of 91%, demonstrating strong robustness when processing noisy, short social media text (Bangyal et al., 2023). Similarly, research on Roman Urdu opinion mining reported that tree-based classifiers, particularly Random Tree, achieved extremely high accuracy (up to 99.78%), illustrating the efficiency of lightweight ML models for analysing text in low-resource languages, a challenge similar to multilingual disaster data (Shafqat et al., 2022). Work on neuroimaging classification further showed that hybrid approaches combining metaheuristic feature-selection (e.g., ACO) with classifiers such as SVM and KNN significantly enhance predictive accuracy, reinforcing the importance of optimised feature extraction in specialised text-classification domains (Ahmad et al., 2025). Additionally, deep learning–based diagnostic tools demonstrated that richer feature representations, such as pseudo amino acid composition and dipeptide composition, consistently outperform simpler feature sets, supporting the rationale for moving from basic word-based features to more complex pattern-based dictionaries in disaster text analysis (Ahmad et al., 2025). Finally, evidence from sentiment-analysis studies confirms that deep learning architectures offer superior performance over classical machine-learning models when handling linguistic variability and contextual dependencies in social media content, highlighting the value of pattern-level information for disaster communication analysis (Zamir et al., 2024).

2.4. Pattern mining

Social media networks are essential for exchanging information in real time during catastrophic situations. Through the process of mining user-generated content, including tweets, posts, and hashtags, researchers can identify common patterns of behaviour, including shared information types, posting locations, and posting times (Zhang et al., 2022). This method also makes it possible to monitor false information, which can spread quickly in emergencies. Authorities can reduce fear and confusion by proactively countering rumours and providing accurate information, which becomes possible when patterns in misleading or incorrect content are effectively identified. Data mining techniques applied to historical disaster records further enhance the ability to detect risk factors and forecast potential future events. By analysing past disasters alongside weather trends, geological data, and other relevant variables, researchers can uncover key risk indicators that support the development of more reliable early warning systems. For instance, historical information on storm behaviour, weather patterns, and seismic activity can inform predictive models that estimate the likelihood of future events (Lensing et al., 2025). Pattern mining was used to analyse evacuation behaviours during hurricanes. They looked at patterns of behaviour by inhabitants, including when they chose to leave, the routes they followed, and their reactions to varying warning levels, using data from previous hurricanes. Emergency planners can create better evacuation routes, enhance public awareness campaigns, and lessen bottlenecks or congestion during high-stress events by knowing the decisions and actions that usually take place during evacuations. To find recurrent patterns of events that result in workplace accidents, such as dangerous behaviours, equipment malfunctions, and hazardous conditions, pattern mining was applied to occupational accident data in the Turkish manufacturing sector. This allowed for more focused safety enhancements and improved preventive measures (Mutlu et al., 2023). By identifying crucial alarm sequences that suggested possible system failures, a priority-aware pattern mining technique was created for alarm flood analysis in industrial settings, which improved decision-making and allowed operators to react quickly to operational hazards. To examine defect text records from maintenance logs and service reports, experts in the field of sustainability suggested a text-specific sequential pattern mining technique. This technique improved maintenance strategies and predictive maintenance models by integrating text mining with sequential pattern mining to find patterns in the order of equipment breakdowns and repair activities (Yuan et al., 2018). In contrast to other methods for pattern mining, RAKE is very adept at processing brief, chaotic, and unstructured messages, rendering it highly appropriate for social media content like tweets. It does not depend on predetermined n-gram sizes or sliding windows, which frequently constrain the adaptability of phrase extraction in other models. Moreover, RAKE is language-agnostic, enabling it to handle multilingual information and typographical errors with greater efficacy (Alzamzami & El Saddik, 2021). Collectively, these studies demonstrate the evolving and interconnected nature of pattern mining applications in disaster management, safety, and risk assessment, contributing to improved decision-making and emergency response strategies across multiple domains.

3. Methodology

3.1. Overview

Social media text often contains valuable information related to specific events or topics. When filtered and analysed, this content can be categorised into multiple relevant classes. From a research perspective, analysing textual data sourced from online platforms offers significant insights, particularly in domains such as disaster management.

In this study, disaster-related content collected from online sources was classified using a dictionary-driven approach. The objective was to categorise the data into two primary types: natural disasters and man-made disasters. Initially, a pre-processing phase was applied to clean the noisy data gathered from social media. After removing irrelevant or redundant content, several feature extraction techniques, including n-gram, bi-gram, and bag-of-words (BoW), were employed to identify informative terms, as shown in Fig. 1.

The most common and frequent terms extracted using these techniques were added to a custom-built disaster dictionary designed to support disaster type classification. Following this, Named Entity Recognition (NER) was applied to extract key entities, which were also incorporated into the dictionary to enhance its effectiveness.

The proposed methodology consisted of two main phases. In the first phase, word-based classification was performed using the disaster dictionary to categorise text data into natural or man-made disasters. To capture deeper contextual cues that may have been overlooked by single-word features, pattern mining was performed using the RAKE algorithm to identify multi-word patterns commonly appearing in disaster-related texts. These patterns were incorporated into the dictionary and were used in a second, pattern-based classification step, enabling a more nuanced interpretation of linguistic expressions typical of social media content (Shinde & Shah, 2018). These patterns were subsequently added to the disaster dictionary and were used in a pattern-based classification process aimed at improving the performance achieved in the first phase, as shown in Fig. 1. Finally, the results of the proposed *DisasterSense* classification approach, combining word-based and pattern-based methods, were compared against various models, each tested with different feature extraction techniques. This comparison was conducted to evaluate the overall effectiveness and accuracy of the proposed method (Wang et al., 2020).

The overall methodology of our approach was described in this section. Initially, the tweet data collected from social media underwent a pre-processing phase. This step was necessary because raw text obtained from online platforms typically contains substantial noise, which can negatively affect model training and evaluation. Pre-processing was therefore essential to refine the data, remove irrelevant or inconsistent elements, and ensure that the dataset was suitable for effective analysis.

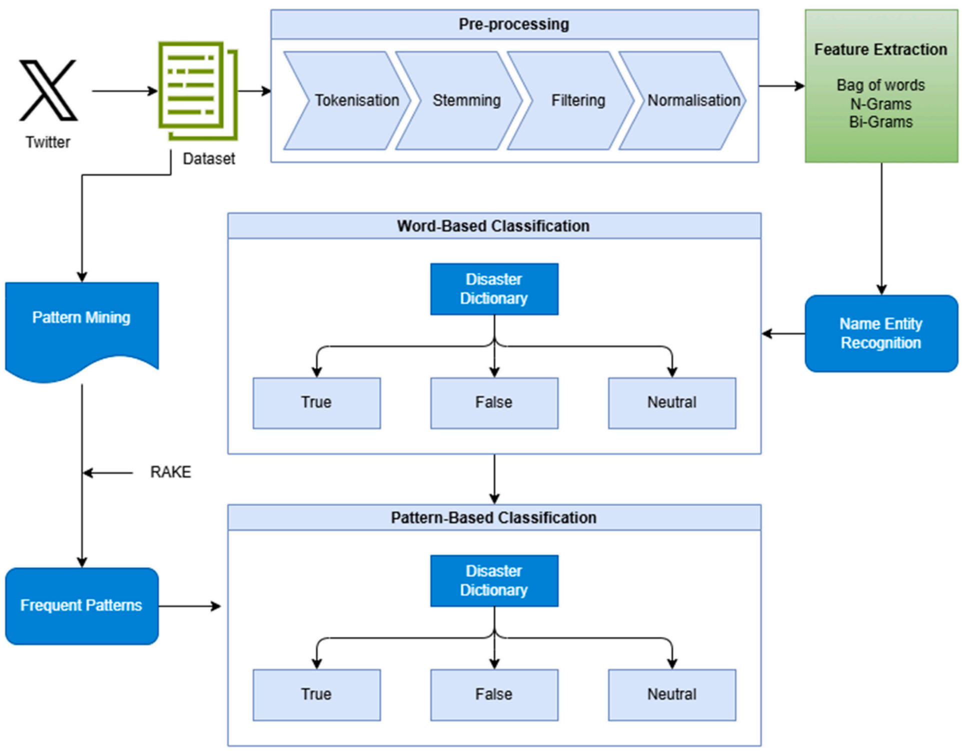

Fig. 1. The overall methodology of the proposed DisasterSense framework, integrating dictionary-driven named entity recognition and pattern-based extraction to enhance disaster text classification.

3.2. *Preprocessing*

Data preprocessing in natural language processing (NLP) is a critical phase that prepares raw text for reliable analysis and model development. The process begins with data cleaning, during which noise and irrelevant elements, such as empty strings, ellipses, stop words, URLs, emojis, and superfluous punctuation, were removed to reduce dimensionality and retain only meaningful content for disaster-related text classification. After cleaning, the text was subjected to tokenisation, which segments the text into smaller units or tokens that may represent words, sub-words, or characters. In this study, tokenisation was performed using the Natural Language Toolkit (Loper & Bird, 2002), enabling subsequent operations such as filtering, stemming, lemmatisation, and transformation into numerical representations (Eamwiwat et al., 2019). To refine the tokenised data further, filtering was applied to eliminate residual noise and irrelevant patterns while retaining information essential to the classification task, thereby improving the model's ability to detect meaningful patterns and semantic relationships (Singh et al., 2024). Following filtering, stemming was used to reduce words to their root forms by removing prefixes and suffixes so that different inflections of the same word (e.g., *walking, walks, walk*) are treated as a single lexical unit. This step enhances the effectiveness of tasks such as information retrieval and text classification (Arianto & Budi, 2020a). The final stage, normalisation, standardises text into a consistent format to ensure that different variants of the same word or expression are processed uniformly, which supports improved performance across various NLP tasks, including classification, retrieval, and sentiment analysis (Kashina et al., 2020). Once these preprocessing steps were completed, the refined text data was prepared for feature extraction and subsequent model training.

3.3. *Feature extraction*

After cleaning and performing preprocessing on social media data. The next step was to apply some useful feature extraction techniques to the data. This allows the extraction of some important features from the text data, which are helpful while implementing machine learning and deep learning models. Thus, we applied Bag of Words (BoW), commonly used for text classification. It consists of individual words and counting the frequency of each word. Each word becomes a feature, and the number or frequency of that word in the text becomes its value (Janiesch et al., 2021). The words that contain high frequency and are commonly used in the dataset were extracted. The frequent words were used to create the disaster dictionary. The arrangement of features as they occur in texts is known as an n-gram. These components can be any type of element that appears sequentially in texts, including words, characters, POS tags, and numbers (Sidorov et al., 2014). The most common n-grams were also used in our disaster dictionary. Contiguous sequences of N-words in a given text are useful for identifying significant patterns and background information. For the successful handling of emergencies, it is necessary to collect voluminous amounts of information, and the extraction of such data can be advantageous for disaster management. Another technique, bi-grams, in which a sequence of two contiguous pieces from a string of tokens, which can be words, was implemented. Bi-grams are often utilised in computational linguistics and natural language processing (NLP) for text analysis, language modelling, and other purposes (Dasgupta et al., 2025). When attempting to decipher a text's structure and patterns, bi-grams help identify frequently occurring word or character pairings (Anisimov et al., 2017). The pairs of words extracted by bi-grams were updated in the dictionary. Popular word embedding models such as Word2Vec, GloVe, and FastText can convert words to numerical vectors of a fixed dimension (Wang et al., 2020). In place of conventional vector space models used in distributional semantics, word embedding models provide vector representations of words for system training. Despite introducing some computational complexity, these models are highly regarded for their effectiveness. One intriguing aspect of word embedding models is their ability to function well with unsupervised data, which means they do not necessitate a big, annotated data set for training (Chen et al., 2021). We used dependency parsing as it focuses on the grammatical structure of the sentence and determines the link between the words. Sentence dependencies are represented as a tree structure. A tree with nodes representing words and edges representing the interdependence between them is what dependency parsing produces. The primary verb in a phrase is usually the root of the tree (Sun et al., 2023). A generative probabilistic technique, Latent Dirichlet Allocation (LDA), was also applied in our work. The fundamental concept is that the texts are shown as arbitrary combinations of latent subjects, each of which is defined by a distribution of words (Jelodar et al., 2019). Character-level embedding was incorporated to enhance system performance, consistent with prior work demonstrating its benefits for neural machine translation, language modelling, and document classification (Luong & Manning, 2016). Such representations allow the model to handle morphological variations and out-of-vocabulary terms more effectively. Term Frequency-Inverse Document Frequency (TF-IDF) was also applied to weight words based on their importance across documents. Words that appear frequently in one document but infrequently in others receive higher weights, allowing important terms to be emphasised during classification (Zhou, 2022). All feature extraction techniques were combined with various machine learning and deep learning models for classification. The performance of all model feature combinations is reported in Section 5.4.

3.4. Name entity recognition

Named Entity Recognition (NER) plays a critical role in disaster management by identifying key entities such as locations, people, organisations, and resources from social media content and other textual data. The integration of NER with sentiment analysis provides valuable insight into the psychological condition of affected individuals and communities, enabling disaster response teams to tailor support and relief efforts accordingly. NER also facilitates resource allocation by helping disaster management teams efficiently identify requests for assistance, supplies, or services shared on social media, thereby enabling the prioritisation of urgent needs.

By analysing social media posts using NER, disaster management teams gain a clearer understanding of affected areas, the scale of the disaster, and ongoing response efforts. Furthermore, NER is essential for distinguishing credible information from unreliable or misleading sources, ensuring that decision-making is guided by accurate and trustworthy data (Demuth et al., 2012). The application of NER in crisis mapping supports the automatic population of critical information such as incident locations, required resources, and affected populations on digital maps, thus improving coordination and situational awareness. Overall, NER enhances disaster management strategies by enabling extensive data extraction, continuous monitoring, sentiment interpretation, resource optimisation, credibility assessment, and crisis mapping, ultimately contributing to more efficient and effective disaster response efforts (Sufi & Khalil, 2022).

3.5. Dictionary-driven NER

Named Entity Recognition (NER) identifies and classifies named entities in a text using predefined inventories such as names, locations, and organisations. It compares words in the text with dictionary entries to discover matches. Dictionary-based NER is a fast and simple method for identifying named entities, although its performance is constrained by the completeness and accuracy of the underlying dictionaries (Quimbaya et al., 2016). In this research, a novel method for categorising man-made and natural disasters using social media data was applied. By using NER on disaster-related social media content and feature extraction techniques, including Bag of Words, n-grams, and bi-grams, the dictionary employed in this study was constructed as a collection of disaster-related keywords (Yang et al., 2019). First, the gathered data was cleaned through preprocessing. Following preprocessing, important phrases were extracted using feature extraction techniques, and these terms were used to create a word-based disaster dictionary, as shown in Fig. 1. The extracted terms supported the classification of disasters into two categories: man-made and natural.

The dictionary-driven NER approach classified each text by matching tokens against predefined keyword lists. Words corresponding to natural disaster terms were assigned to the natural disaster category, while those matching man-made disaster terms were assigned to the man-made category. When no match was found, the word was classified as neutral or irrelevant, as demonstrated in Algorithm 1. We used the Python library spaCy (Honnibal et al., 2020) to develop the disaster dictionary for classifying natural and man-made disasters. This research was structured into two phases: word-based classification and pattern-based classification.

3.6. Disaster classification (Word-based comparison)

The first step in this phase involved processing the collected social media data in textual form. The data underwent preprocessing to remove noise and irrelevant content, ensuring a cleaner input for subsequent analysis. Once the word-based disaster dictionary had been constructed, the proposed approach (Algorithm 1) accepted both the text data and the dictionary as input. It then compared the words in the dictionary with the text data to classify each entry as true for natural disasters, false for man-made disasters, or neutral if no relevant match was found. The classification was based on a predefined list of keywords representing both natural and man-made disasters. The natural disaster keywords included terms such as earthquakes, volcanic eruptions, tsunamis, landslides, hurricanes, cyclones, typhoons, tornadoes, blizzards, floods, droughts, extreme heat waves, flash floods, avalanches, and pandemics. In contrast, the man-made disaster category included keywords such as oil spills, nuclear accidents, industrial explosions, chemical spills, deforestation, mining accidents, building collapses, transport accidents, terrorist attacks, fires, waste mismanagement, and dam failures. Any text that did not match either category was classified as neutral, ensuring that only relevant disaster-related content contributed to the classification process.

3.7. Finding the relation between words and patterns

Following the first stage of text-based classification, the second phase applied a Pattern Mining approach (Fournier-Viger et al., 2017) to the categorised disaster-related and non-disaster text data. After completing the word-based classification, we extracted common terms that appeared across both output files and subsequently identified recurring patterns from these shared terms, as illustrated in Fig. 1. This process enabled the detection of recurring phrases and trends present within the classified data. Association rule mining was used to determine the relationships between frequently co-occurring words and patterns (Diaz-Garcia et al., 2023). When Pattern Mining was applied to the classified dataset, the primary objective was to identify patterns that could support the system in distinguishing between social media content related to natural versus man-made disasters.

The next stage involved creating a pattern-based disaster dictionary that incorporated these frequently occurring patterns associated with both types of disasters. This dictionary was constructed using the Rapid Automatic Keyword Extraction (RAKE) model, which facilitated the extraction of meaningful pattern structures for enhanced classification.

3.8. By applying the rake model for patterns extraction

RAKE is an unsupervised technique for extracting keywords and phrases from text. Because it does not require training data, it is considered adaptable and efficient for various applications, including text mining, information retrieval, and summarisation (Thushara et al., 2017). To capture word sequences in context, RAKE begins by dividing the text into sentences. It then identifies the key content words in each

Algorithm 1
Disaster Dictionary Construction and Text Classification.

```
Input: text_data
dictionary ← GenerateBagOfWords(text_data) ∪ GenerateNGrams(text_data, n ≥ 2)
natural_keywords ← ∅
man_made_keywords ← ∅
For each text t in text_data do
    For each term x in ApplyNER(t) ∪ ApplyRAKE(t) do
        If IsRelatedToNaturalDisaster(x) then
            natural_keywords ← natural_keywords ∪ {x}
        Else if IsRelatedToManMadeDisaster(x) then
            man_made_keywords ← man_made_keywords ∪ {x}
        End if
    End for
End for
dictionary ← dictionary ∪ natural_keywords ∪ man_made_keywords
results ← empty list
For each text t in text_data do
    n ← CountOverlap(t, natural_keywords)
    m ← CountOverlap(t, man_made_keywords)
    If n > m then
        Append true to results
    Else if m > n then
        Append false to results
    Else
        Append neutral to results
    End if
End for
Output: dictionary, natural_keywords, man_made_keywords, results
```

sentence by removing common stopwords (such as *is, the, and*), leaving only the most informative lexical items. These remaining content words are grouped into candidate phrases based on their proximity within the sentence, using stopwords as natural boundaries. This process produces a list of potential phrases that may be relevant to the underlying text.

RAKE assigns scores to keywords and phrases by considering both word frequency and word degree. A word's *degree* refers to the number of other content words it co-occurs with, while its *frequency* denotes how often it appears in the text. RAKE calculates an individual word's score using the degree-to-frequency ratio. The final score of a phrase is computed by summing the scores of its constituent words. Higher-scoring phrases are interpreted as more significant because they are more strongly associated with the core content (Arianto & Budi, 2020b). Equation 1 presents the formula for calculating an individual word score, while Equation 2 illustrates how the overall phrase score is obtained by summing the scores of its component words.

$$Score(w) = Degree(w) \,/\, Frequency(w) \tag{1}$$

$$Score(phrase) = \Sigma\, Score(w), \quad w \in phrase \tag{2}$$

Where:

In this study, w denotes a candidate word extracted from the input text. The term Degree(w) represents the total number of co-occurrences of the word w with other words appearing within the same phrase, capturing its level of connectivity in the corpus. The frequency of a word, Frequency(w), is defined as the number of times the word w occurs in the corpus. Based on these properties, an importance score (w) is assigned to each candidate word according to Equation (1).

A phrase refers to a multi-word candidate expression identified during the keyword extraction process, where each word $w \in$ phrase contributes individually to the phrase's significance. The overall importance of a phrase, denoted as Score(phrase), is computed by aggregating the scores of its constituent words, as defined in Equation (2).

Equation (1) computes the significance of each word by normalising its co-occurrence degree with its textual frequency, ensuring that words with broader contextual associations receive higher weights. Equation (2) then sums the individual word scores to produce a final score for each candidate phrase. Phrases containing multiple high-scoring words achieve higher overall importance, allowing RAKE to effectively identify meaningful domain-specific keywords for the disaster dictionary.

3.9. Pseudocode

The pseudocode presents a two-phase approach for disaster-related text classification. In the first phase, a disaster dictionary is constructed directly from the input text data. The process begins by generating a bag-of-words representation and n-grams (where $n \geq 2$). These components are combined to form the initial dictionary. Two empty keyword sets, natural_keywords and man_made_keywords, are then initialized.

For each text in the dataset, Named Entity Recognition (NER) and RAKE are applied jointly to extract relevant terms. Each extracted term is evaluated to determine whether it corresponds to a natural disaster or a man-made disaster. Terms identified as natural disaster-related are added to the natural_keywords set, while those related to man-made disasters are added to the man_made_keywords set. After processing all texts, both keyword sets are merged with the initial dictionary to produce the final disaster dictionary.

In the second phase, classification is performed. An empty results list is initialized. For each text, the number of overlapping terms with natural_keywords and man_made_keywords is calculated. If the count of natural keyword matches exceeds that of man-made keywords, the text is labelled true. If the count of man-made keyword matches is higher, it is labelled false. If both counts are equal or no clear dominance is observed, the text is labelled neutral. The algorithm outputs the final dictionary, the two keyword sets, and the classification results.

4. Results

4.1. Data

The data set utilised for training in our study is a collection of disaster-related tweets and is online and available at Kaggle (Disasters on social media, 2026). It includes 10875 tweets on various sorts of natural and other types of disasters. The Tweets posted by different users from all over the world highlight disaster-related events and incidents, including natural and man-made disasters. Tweets are classified based on keywords into two categories: natural disasters and man-made disasters. Following that, another data set of 674 disaster-related events that occurred in Australia was used for testing (Australia disaster events by category impact data and geographic location, 2026). It contains information related to different disaster-related events and incidents that occurred in the Australian region.

The dataset utilised for testing was manually labelled. Three domain experts independently reviewed and classified each disaster-related instance into one of three categories: natural disaster, man-made disaster, or neutral. After the initial labelling phase, all instances with conflicting annotations were collectively discussed and resolved through consensus to ensure consistency and reduce subjectivity. Natural disasters were defined as events directly linked to natural hazards or environmental phenomena such as floods, earthquakes, storms, or landslides. Man-made disasters included events caused by human actions, negligence, technological failures, industrial accidents, or conflict-related incidents. Instances that did not correspond to either natural or man-made disasters, or that provided general information without indicating a specific hazardous event, were labelled as neutral. In cases where an event involved both natural and man-made elements, the annotators evaluated the dominant causal context and assigned the label based on the primary driver of the scenario. This systematic, expert-driven annotation process ensures a well-defined labelling schema and enhances the reliability and interpretability of the classification results. All of the models were evaluated using measuring parameters to quantify their efficiency.

4.2. Evaluation

Our models were evaluated using the following parameters: accuracy, precision, recall and F1 score. An accuracy metric is the most straightforward way to assess a model's performance. It is calculated as the proportion of accurately predicted observations to total observations (Mishra, 2018). When the target classes are evenly distributed, accuracy provides a meaningful performance estimate. Accuracy ranges from 0 to 1, where 1 represents completely correct predictions and 0 represents entirely incorrect predictions. The equations below present the different parameters used to measure model performance:

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \quad (3)$$

$$Precision = (TP) / (TP + FP) \quad (4)$$

$$Recall = (TP) / (TP + FP) \quad (5)$$

$$F1\ Score = (2 \times Precision \times Recall) / (Precision + Recall) \quad (6)$$

True Positives (TP) are occurrences in which the model properly predicted the positive class. False Positives (FP) are instances in which the model mistakenly predicted a positive class when it was negative. False Negatives (FN) are instances in which the model mistakenly predicted a negative class when it was positive (Dalianis, 2018).

4.3. Model evaluation and comparative analysis

In this section, we compare our proposed approach, DisasterSense, a pattern-based classification method using the RAKE model, with DB-NER (a word-based classification technique) and other traditional machine learning and deep learning models trained with various feature extraction methods. Fig. 2 illustrates a bar graph showcasing the best performance metrics of each model when combined with their most effective feature extraction techniques, evaluated using accuracy, precision, recall, and F1 score.

Among traditional models, Logistic Regression (LR) combined with n-grams achieved the highest accuracy of 0.680 compared to other feature extraction techniques applied to LR. The Support Vector Machine (SVM) using TF-IDF performed comparably, achieving an accuracy of 0.678 and the highest precision of 0.859 across all models. The K-Nearest Neighbours (KNN) model performed consistently across n-gram, BoW, and bi-gram features, with an accuracy of 0.653.

Among deep learning models, the Recurrent Neural Network (RNN) combined with bi-grams achieved an accuracy of 0.657, while LSTM with TF-IDF and CNN with BoW produced similar performances, with accuracies of 0.689 and 0.690, respectively. BERT, known for its effectiveness in text classification

(Garrido-Merchan et al., 2023), achieved an accuracy of 0.666 when used with n-grams among the BERT-family methods included for comparison.

The DB-NER method, as shown in Table 1, demonstrated strong performance based on word-level classification, achieving an accuracy of 0.717, second only to the proposed approach. *DisasterSense*, which

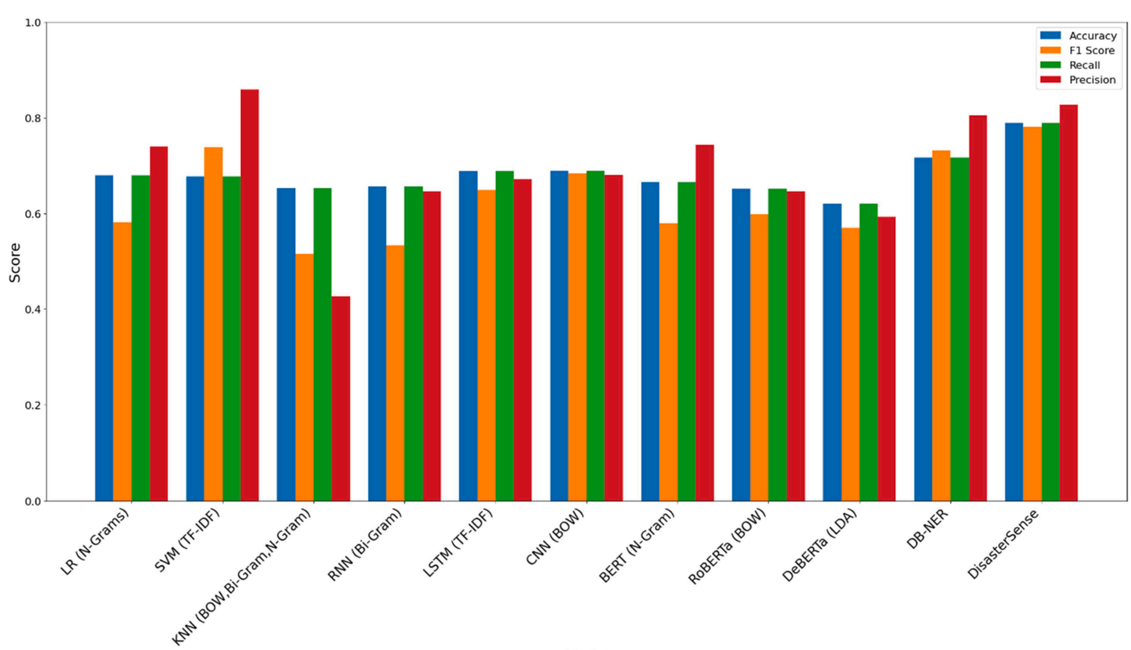

Fig. 2. Comparison of the best-performing models using various feature extraction techniques against our proposed approach, DisasterSense.

Table 1
Best results of all models compared with our proposed method.

Model + Technique	Accuracy	F1 Score	Recall	Precision
LR + N-Gram	0.680	0.582	0.680	0.740
SVM + TF-IDF	0.678	0.739	0.678	0.859
KNN + Bi-Gram, N-Gram, BoW	0.653	0.516	0.653	0.427
RNN + Bi-Gram	0.657	0.533	0.657	0.646
LSTM + TF-IDF	0.689	0.649	0.689	0.672
CNN + BoW	0.690	0.684	0.690	0.681
BERT + N-Gram	0.666	0.581	0.666	0.744
RoBERTa + BOW	0.652	0.599	0.652	0.646
DeBERTa + LDA	0.621	0.571	0.621	0.592
DB-NER + Bi-Gram, N-Gram, BoW	0.717	0.732	0.717	0.805
DisasterSense	0.789	0.781	0.789	0.828

leveraged a pattern-based classification approach with dictionary-driven named entity recognition trained using the RAKE model, outperformed all other methods. It achieved the highest accuracy of 0.789 and the second-highest precision of 0.828, demonstrating superior effectiveness in classifying disaster-related data into natural and man-made categories. Overall, the proposed pattern-based method showed the best performance among all model combinations with different feature extraction techniques discussed in Section 5.4.

These results demonstrated the comparative performance of various feature extraction techniques when combined with different models. As highlighted in Table 1, DB-NER used a word-based dictionary for classification and was trained using n-gram, bi-gram, and BoW techniques, along with named entity recognition. By employing a word-based dictionary, DB-NER performed as the second-best model among all methods included in the experimental comparison. The model achieved an accuracy of 0.717, with corresponding F1-score, recall, and precision values of 0.732, 0.717, and 0.805, respectively.

DB-NER used a word-based dictionary for the classification of natural and man-made disaster data, which was later updated with patterns extracted from the data using the RAKE model. Following this update, the *DisasterSense* pattern-based approach further improved classification performance.

4.4. Ablation study

Machine learning models such as Logistic Regression (LR), Support Vector Machine (SVM), and K Nearest Neighbour (KNN) need large amounts of data for training. We had considered logistic regression since logistic regression is an easy-to-understand, efficient technique for binary and multinomial classification tasks, and it is frequently applied to classification issues. It is appropriate for large-scale text datasets since it often uses less memory and processing resources than more intricate models (Shah et al., 2020). Support Vector Machine (SVM) is appropriate for text classification if there are several characteristics (words or n-grams), as they perform well with high-dimensional data (Devi et al., 2019). KNN is an appealing option for disaster classification tasks, especially where high interpretability and ease of implementation are goals, despite potential scalability and imbalance issues. This is due to its simplicity and flexibility in handling various feature representations (Sreenivasulu & Sridevi, 2020). By identifying language patterns associated with disasters, such as "earthquake just struck" or "flood alert," which frequently denote catastrophic events, RNNs can assist in disaster categorisation (Eligüzel et al., 2022). CNNs can recognise terms like "emergency alert" or "severe weather warning," which are essential for recognising communications about disasters (Madichetty & M, 2021). However, CNNs' limited window widths make it difficult for them to collect context beyond localised patterns, making them ineffective for extended relationships across text. The vanishing gradient problem is resolved by Long Short-Term Memory Networks (LSTMs), with a gating mechanism to preserve information over long sequences. Due to its ability to capture long-term dependencies in text, LSTMs are well-suited for catastrophe categorisation jobs where it is essential to comprehend how context changes over time (Wang & Li, 2022). The sophisticated method enables BERT to recognise tiny clues in literature about disasters, making it especially effective for comprehending complex language. In words like "roads are impassable due to landslide" or "urgent relief needed," for instance, BERT may identify underlying meanings even when they convey sensitive attitudes relating to disasters (Zhou et al., 2022). These models, along with the combination of different feature extraction techniques, were applied and compared with our proposed method for disaster data classification. The heat map shows all models trained by various feature extraction techniques and compared with other well-known models used in our experiment, as shown in Fig. 3. These results, by comparing accuracy, precision, F1 score and recall, were revealed in a heat matrix. As shown on the top of the matrix, the accuracy results for each model utilising various feature extraction methods are compared, in which the accuracy results of deep learning models were compared with machine learning models. The hyper-parameter utilised for the experiment with batch size 64 and 10 epochs to evaluate the results of different models. The CNN deep learning model trained with the Bag of Words (BoW) produced the best accuracy of 0.690 among machine and deep learning methods. At 0.689, the LSTM model, which was trained using TF-IDF, achieved the second-highest accuracy. This is due to the efficient feature modelling of both models and the self-capability of automatically extracting key features from a dataset. Also, these models perform well on large datasets. On the other side, utilising BoW, the logistic regression (LR) attained a high accuracy of 0.680 among machine learning models, but the support vector machine (SVM), trained with TF-IDF, had achieved an accuracy of 0.678. Among the transformer models, BERT achieved the highest accuracy with the n-gram representation 0.666, followed by RoBERTa with (BoW) 0.652 and DeBERTa with (LDA) 0.621. All other models, along with the combination of different feature extraction techniques, are shown in Fig. 3, highlighting dark colours with the highest accuracy and light colours with the lowest. The SVM model, which was also trained on TF-IDF, came in second with a precision of 0.859, while logistic regression (LR), which was trained on TF-IDF, had the best precision of 0.869. In contrast, the BERT model with dependency parsing came with 0.764, and the recurrent neural network (RNN) trained with N-grams attained a precision of 0.774.

The CNN model with bi-grams was also close, obtaining 0.742 precision among the others. With the best score of 0.742 across all feature extraction methods, logistic regression (LR) beat all other models for the F1 score. When paired with TF-IDF, the SVM model achieved the second-highest F1 score of 0.739, and when mixed with BoW, the convolutional neural network (CNN) achieved the third-highest score of 0.685. Comparing the results of recall, the CNN model with BoW had achieved the best recall, 0.690, closely followed by the LSTM model, 0.689. Other models also showed good performance, with SVM reaching 0.678 when trained on TF-IDF and logistic regression (LR) reaching 0.680 when paired with BoW. One of our significant discoveries was the performance gap between machine learning and deep learning models. Regarding accuracy, deep learning models such as CNN trained on BoW and LSTM trained on TF-IDF beat machine learning models such as LR trained on BoW and SVM trained on TF-IDF, providing the best accuracy among machine learning models. However, in terms of precision, machine learning models such as LR and SVM outperformed the deep learning models, including RNN and the Transformer-based models RoBERTa, DeBERTa, and BERT. Transformer-based models are inherently capable of learning rich contextual representations and extracting useful features directly from the raw text. As a result, most external feature extraction techniques did not lead to performance improvements, particularly in the case of RoBERTa and DeBERTa, where the models' built-in feature extraction mechanisms already captured the necessary linguistic and semantic information. The F1 score of machine learning models, LR and SVM, was likewise higher than that of deep learning models like CNN and LSTM. Recall scores followed a similar pattern to

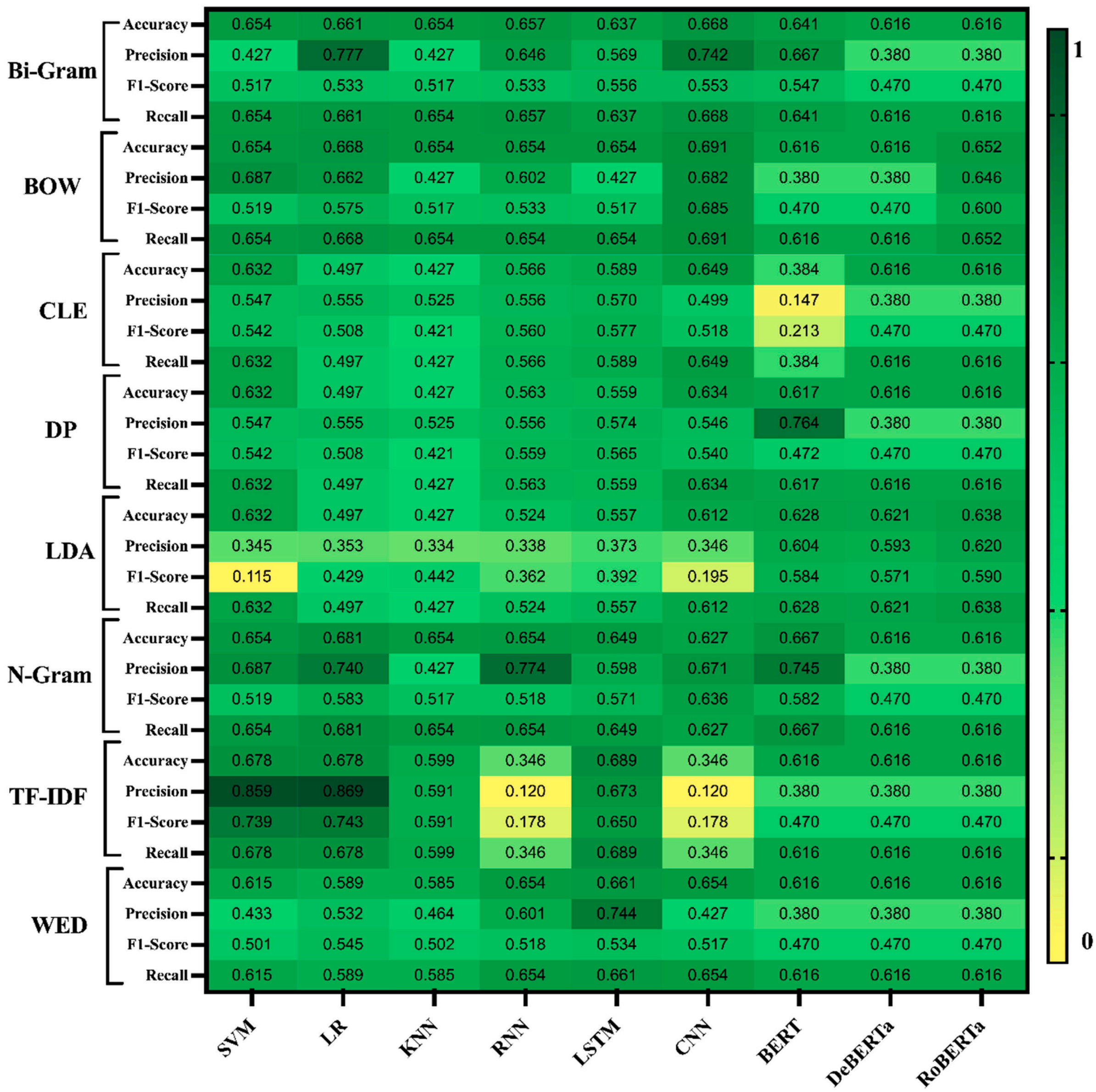

Fig. 3. Results comparison of Models, along with the combination of feature extraction techniques trained.

accuracy, with deep learning models beating machine learning approaches. The results in the matrix show the variation in terms of measuring parameters, and this is due to the capability of each model for the classification of disaster data by utilising the feature extraction techniques.

5. Conclusion and future works

This paper developed a unified dictionary-driven framework that successfully distinguished natural and man-made disasters without requiring large, annotated datasets. It introduced a DB-NER–based dictionary expansion method that automatically extracted disaster-related entities to enhance contextual understanding and implemented a RAKE-driven pattern-mining component that captured frequent lexical structures to strengthen semantic representation.

The paper evaluated a range of traditional machine learning and deep learning, along with transformer-based approaches. It demonstrated that the proposed DisasterSense framework outperformed all baseline models, achieving the highest accuracy of 0.789. Scientifically, the work addressed key challenges in disaster informatics by reducing dependence on manual annotation, introducing an automated text-labelling strategy, and improving the efficiency and scalability of disaster information processing. Practically, accurate classification of natural and man-made disasters enables government agencies and emergency responders to assess impacts, prioritise resource allocation, and respond more effectively. By enabling rapid differentiation between the two disaster types, the proposed approach supports more informed decision-making and strengthens real-time disaster management capabilities.

For the Future, a multi-model strategy that further verifies and enhances disaster-related data could be produced by incorporating additional data sources, such as satellite imaging, news feeds, and emergency broadcasts. Enhancing pattern identification by combining transformer-based models with pattern mining may also help capture intricate contextual details that set related disaster types apart. To further confirm the model's resilience and usefulness, it should be compared to other historical catastrophic occurrences.

CRediT authorship contribution statement

Muhammad Salman Tiwana: Conceptualization, Methodology, Software, Data curation, Writing – original draft, Visualization, Project administration. **Ali Braytee:** Methodology, Software, Validation, Formal analysis, Writing – review & editing. **Faezeh Karimi:** Data curation, Investigation, Writing – review & editing. **Madhushi Bandara:** Formal analysis, Validation, Resources, Writing – review & editing. **Tony Jan:** Writing – review & editing, Investigation, Formal analysis. **Mukesh Prasad:** Supervision, Conceptualization, Validation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Adesokan, A., Madria, S., & Nguyen, L. (2023). Tweetace: A fine-grained classification of disaster tweets using transformer model. In *2023 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)* (pp. 1–9). https://doi.org/10.1109/AIPR60534.2023.10440656

Ahmad, W., Shahzad, A. R., Amin, M. A., Bangyal, W. H., Alahmadi, T. J., & Khan, S. H. (2025). Machine learning driven dashboard for chronic myeloid leukemia prediction using protein sequences. *PloS One, 20*(6), Article e0321761. https://doi.org/10.1371/journal.pone.0321761

Albladi, A., Islam, M., & Seals, C. (2025). Sentiment analysis of twitter data using NLP models: A comprehensive review. *IEEE Access*.

Alzamzami, F., & El Saddik, A. (2021). Monitoring cyber SentiHate social behavior during COVID-19 pandemic in North America. *IEEE Access, 9*, 91184–91208.

Anand, M., Sahay, K. B., Ahmed, M. A., Sultan, D., Chandan, R. R., & Singh, B. (2023). Deep learning and natural language processing in computation for offensive language detection in online social networks by feature selection and ensemble classification techniques. *Theoretical Computer Science, 943*, 203–218.

Anisimov, I., Polyakov, V., Makarova, E., & Solovyev, V. (2017). Spelling correction in english: Joint use of bi-grams and chunking. In *2017 Intelligent Systems Conference (IntelliSys)* (pp. 886–892). https://doi.org/10.1109/IntelliSys.2017.8324234

Arianto, D., & Budi, I. (2020a). Aspect-based sentiment analysis on indonesia's tourism destinations based on google maps user code-mixed reviews (study case: borobudur and prambanan temples). In *Proceedings of the 34th Pacific Asia Conference on Language, Information and Computation* (pp. 359–367). https://aclanthology.org/2020.paclic-1.41/.

Arianto, D., & Budi, I. (2020b). Aspect-based sentiment analysis on Indonesia's Tourism Destinations Based on Google Maps User Code-Mixed Reviews (Study Case: Borobudur and Prambanan Temples). In *Pacific Asia Conference on Language, Information and Computation*. https://api.semanticscholar.org/CorpusID:234487229.

Australia disaster events by category impact data and geographic location. (2026). https://www.data.gov.au/data/dataset/disaster-events-with-category-impact-and-location/resource/ad5c6594-571e-4874-994c-a9f964d789df.

Bangyal, W. H., Amina, S., Shakir, R., Ubakanma, G., & Iqbal, M. (2023). Using deep learning models for COVID-19 related sentiment analysis on Twitter data. In *2023 International Conference on Human-Centered Cognitive Systems (HCCS)* (pp. 1–6). https://doi.org/10.1109/HCCS59561.2023.10452567

Battistoli, B., King, T., & White, E. (2018). Voices in the storm: The lost discourse of climate change in Hurricanes Harvey and Irma. *International Journal of Crisis Communication, 1*(1), 72–78. https://doi.org/10.31907/2617-121X.2017.01.02.02

Bhuvana, N., & Aram, I. A. (2019). Facebook and Whatsapp as disaster management tools during the Chennai (India) floods of 2015. *International Journal of Disaster Risk Reduction, 39*, Article 101135. https://doi.org/10.1016/j.ijdrr.2019.101135

Bikku, T., Chandrika, P., Kanyadari, A., Prathima, V., & Sai, B. B. (2022). Analysis of disaster tweets using natural language processing. In *Intelligent Computing and Applications: Proceedings of ICDIC 2020* (pp. 491–501). Springer.

Chen, Q., Wang, W., Huang, K., & Coenen, F. (2021). Zero-shot text classification via knowledge graph embedding for social media data. *IEEE Internet of Things Journal, 9*(12), 9205–9213. https://doi.org/10.1109/JIOT.2021.3093065

Cresci, S., Tesconi, M., Cimino, A., & Dell'Orletta, F. (2015). A linguistically-driven approach to cross-event damage assessment of natural disasters from social media messages. In *Proceedings of the 24th international conference on world wide web* (pp. 1195–1200). https://doi.org/10.1145/2740908.2741722

Dalianis, H. (2018). Evaluation metrics and evaluation. *Clinical Text Mining: secondary use of electronic patient records*, 45–53. https://doi.org/10.1007/978-3-319-78503-5_6

Dasgupta, S., Maity, C., Mukherjee, S., Singh, R., Dutta, D., & Jana, D. (2025). HITgram: A Platform for Experimenting with n-Gram Language Models. In *International Conference on Applied Algorithms* (pp. 92–104). https://doi.org/10.48550/arXiv.2412.10717

Demuth, J. L., Morss, R. E., Morrow, B. H., & Lazo, J. K. (2012). Creation and communication of hurricane risk information. *Bulletin of the American Meteorological Society, 93*(8), 1133–1145. https://doi.org/10.1175/BAMS-D-11-00150.1

Devi, G. R., Kumar, M. A., & Soman, K. (2019). Extraction of named entities from social media text in tamil language using N-gram embedding for disaster management. *Nature-Inspired Computation in Data Mining and Machine Learning* (pp. 207–223). Springer.

Diaz-Garcia, J. A., Ruiz, M. D., & Martin-Bautista, M. J. (2023). A survey on the use of association rules mining techniques in textual social media. *Artificial Intelligence Review, 56*(2), 1175–1200. https://doi.org/10.1007/s10462-022-10196-3

Disasters on social media. (2026). https://www.kaggle.com/datasets/jannesklaas/disasters-on-social-media/data.

Dominguez-Péry, C., Vuddaraju, L. N. R., Corbett-Etchevers, I., & Tassabehji, R. (2021). Reducing maritime accidents in ships by tackling human error: A bibliometric review and research agenda. *Journal of Shipping and Trade, 6*(1), 20.

Eamwiwat, C., Thanasutives, P., Saetia, C., & Chalothorn, T. (2019). Using label noise filtering and ensemble method for sentiment analysis on Thai social data. In *2019 14th International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP)*.

Eligüzel, N., Çetinkaya, C., & Dereli, T. (2022). Application of named entity recognition on tweets during earthquake disaster: a deep learning-based approach. *Soft Computing, 26*(1), 395–421.

Fan, C., Wu, F., & Mostafavi, A. (2020). A hybrid machine learning pipeline for automated mapping of events and locations from social media in disasters. *IEEE access, 8*, 10478–10490.

Fanni, S. C., Febi, M., Aghakhanyan, G., & Neri, E. (2023). Natural language processing. *Introduction to Artificial Intelligence* (pp. 87–99). Springer. https://doi.org/10.1007/978-3-031-25928-9_5

Fauzi, M. A. (2023). Social media in disaster management: Review of the literature and future trends through bibliometric analysis. *Natural Hazards, 118*(2), 953–975.

Fournier-Viger, P., Lin, J. C. W., Kiran, R. U., Koh, Y. S., & Thomas, R. (2017). A survey of sequential pattern mining. *Data Science and Pattern Recognition, 1*(1), 54–77. https://doi.org/10.1145/3314107

Gardazi, N. M., Daud, A., Malik, M. K., Bukhari, A., Alsahfi, T., & Alshemaimri, B. (2025). BERT applications in natural language processing: A review. *Artificial Intelligence Review, 58*(6), 1–49.

Garrido-Merchan, E. C., Gozalo-Brizuela, R., & Gonzalez-Carvajal, S. (2023). Comparing BERT against traditional machine learning models in text classification. *Journal of Computational and Cognitive Engineering, 2*(4), 352–356.

Girsang, A. S., & Noveta, B. K. (2024). Six classes named entity recognition for mapping location of Indonesia natural disasters from twitter data. *International Journal of Intelligent Computing and Cybernetics, 17*(2), 395–414.

Habdank, M., Rodehutskors, N., & Koch, R. (2017). Relevancy assessment of tweets using supervised learning techniques: Mining emergency related tweets for automated relevancy classification. In *2017 4th International conference on information and communication technologies for disaster management (ICT-DM)*.

Hagiwara, M. (2021). *Real-world natural language processing: practical applications with deep learning*. Simon and Schuster. http://ieeexplore.ieee.org/document/10280435.

Honnibal, M., Montani, I., Van Landeghem, S., & Boyd, A. (2020). Industrial-strength natural language processing in python. *spaCy*. https://doi.org/10.5281/zenodo.1212303

Islam, T., Islam, M., Rudra, R., Pranto, I. H., Mahmud, M. T., & Foysal, M. M. R. (2023). Disaster related tweets analysis with machine learning approaches. In *2023 2nd International Conference on Futuristic Technologies (INCOFT)*.

Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets, 31*(3), 685–695. https://doi.org/10.1007/s12525-021-00475-2

Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2019). Latent Dirichlet allocation (LDA) and topic modeling: Models, applications, a survey. *Multimedia tools and applications, 78*(11), 15169–15211.

Kang, Y., Cai, Z., Tan, C.-W., Huang, Q., & Liu, H. (2020). Natural language processing (NLP) in management research: A literature review. *Journal of Management Analytics, 7*(2), 139–172. https://doi.org/10.1080/23270012.2020.1756939

Kashina, M., Lenivtceva, I., & Kopanitsa, G. (2020). Preprocessing of unstructured medical data: the impact of each preprocessing stage on classification. *Procedia Computer Science, 178*, 284–290. https://doi.org/10.1016/j.procs.2020.11.030

Ketmaneechairat, H., & Maliyaem, M. (2020). Natural language processing for disaster management using conditional random fields. *Journal of Advances in Information Technology, 11*(2).

Khan, I., Ali, A., Waqas, T., Ullah, S., Ullah, S., Shah, A. A., & Imran, S. (2022). Investing in disaster relief and recovery: A reactive approach of disaster management in Pakistan. *International journal of disaster risk reduction, 75*, Article 102975.

Khosravi, K., Pham, B. T., Chapi, K., Shirzadi, A., Shahabi, H., Revhaug, I., Prakash, I., & Bui, D. T. (2018). A comparative assessment of decision trees algorithms for flash flood susceptibility modeling at Haraz watershed, northern Iran. *Science of the Total Environment, 627*, 744–755.

Kim, J. M., Bae, J., Adhikari, M. D., & Yum, S. G. (2023). A study of deep learning algorithm usage in predicting building loss ratio due to typhoons: The case of southern part of the Korean Peninsula. *Frontiers in Earth Science, 11*, Article 1136346.

Koshy, R., & Elango, S. (2023). Multimodal tweet classification in disaster response systems using transformer-based bidirectional attention model. *Neural Computing and Applications, 35*(2), 1607–1627. https://doi.org/10.1007/s00521-022-07790-5

Lam, N. S., Meyer, M., Reams, M., Yang, S., Lee, K., Zou, L., Mihunov, V., Wang, K., Kirby, R., & Cai, H. (2023). Improving social media use for disaster resilience: Challenges and strategies. *International Journal of Digital Earth, 16*(1), 3023–3044.

Lensing, J. C., Choe, J. Y., Johnson, B. B., & Wang, J. (2025). Text mining of practical disaster reports: Case study on Cascadia earthquake preparedness. *Plos one, 20*(1), Article e0313259.

Lingad, J., Karimi, S., & Yin, J. (2013). Location extraction from disaster-related microblogs. In *Proceedings of the 22nd international conference on world wide web*.

Loper, E., & Bird, S. (2002). Nltk: The natural language toolkit. *arXiv preprint cs/0205028*. https://doi.org/10.48550/arXiv.cs/0205028

Lovera, F., & Cardinale, Y. (2025). An analytics approach to extracting location and sentiment insights from classified social media data. *Decision Analytics Journal*, Article 100624.

Luong, M.T., & Manning, C.D. (2016). Achieving open vocabulary neural machine translation with hybrid word-character models. *arXiv preprint arXiv:1604.00788*.

Madichetty, S., & M, S. (2021). A stacked convolutional neural network for detecting the resource tweets during a disaster. *Multimedia Tools And Applications, 80*(3), 3927–3949.

Mishra, A. (2018). Metrics to evaluate your machine learning algorithm. *Towards Data Science*, 24.

Moishin, M., Deo, R. C., Prasad, R., Raj, N., & Abdulla, S. (2021). Designing deep-based learning flood forecast model with ConvLSTM hybrid algorithm. *IEEE Access, 9*, 50982–50993.

Mutlu, N. G., Altuntas, S., & Dereli, T. (2023). The evaluation of occupational accident with sequential pattern mining. *Safety Science, 166*, Article 106212.

Novianti, P., Setyorini, D., & Rafflesia, U. (2017). K-Means cluster analysis in earthquake epicenter clustering. *International Journal of Advances in Intelligent Informatics, 3*(2), 81–89.

Pratama, R. A. M., & Pardede, H. F. (2023). Disaster tweet classifications using hybrid convolutional layers and gated recurrent unit. *Int. J. Com. Dig. Sys, 14*(1).

Quimbaya, A. P., Múnera, A. S., Rivera, R. A. G., Rodríguez, J. C. D., Velandia, O. M. M., Peña, A. A. G., & Labbé, C. (2016). Named entity recognition over electronic health records through a combined dictionary-based approach. *Procedia Computer Science, 100*, 55–61.

Ro, S. H., Li, Y., & Gong, J. (2024). A machine learning approach for post-disaster data curation. *Advanced Engineering Informatics, 60*, Article 102427.

Savale, V., Patil, S., Raut, A., Phadke, M., Sundkar, R., & Dongre, R. (2024). Disaster management ecosystem using crowdsourced data, deep learning and NLP. In *2024 International Conference on Sustainable Communication Networks and Application (ICSCNA)*.

Shafqat, Z., Iqbal, M., Bangyal, W. H., & Almakhles, D. (2022). Opinion mining of politics and inflation using roman Urdu dataset. In *2022 Human-Centered Cognitive Systems (HCCS)*.

Shah, K., Patel, H., Sanghvi, D., & Shah, M. (2020). A comparative analysis of logistic regression, random forest and KNN models for the text classification. *Augmented Human Research, 5*(1), 12.

Sharma, T., Singhal, A., Kundu, K., & Agarwal, N. (2022). Machine learning/deep learning for natural disasters. *Applications of Artificial Intelligence, Big Data and Internet of Things in Sustainable Development* (pp. 91–121). CRC Press.

Shidik, G. F., Saputra, F. O., Saraswati, G. W., Winarsih, N. A. S., Rohman, M. S., Pramunendar, R. A., Kusuma, E. J., Ratmana, D. O., Venus, V., & Andono, P. N. (2024). Indonesian disaster named entity recognition from multi source information using bidirectional LSTM (BiLSTM). *Journal of Open Innovation: Technology, Market, and Complexity, 10*(3), Article 100358.

Shinde, P. P., & Shah, S. (2018). A review of machine learning and deep learning applications. In *2018 Fourth international conference on computing communication control and automation (ICCUBEA)*.

Sidorov, G., Velasquez, F., Stamatatos, E., Gelbukh, A., & Chanona-Hernández, L. (2014). Syntactic n-grams as machine learning features for natural language processing. *Expert Systems with Applications, 41*(3), 853–860. https://doi.org/10.1016/j.eswa.2013.08.015

Singh, G., Singh, D., Sharma, R., & Bhardwaj, K. (2024). Potential approach for text-based emotion detection using NLP coupled with deep learning of sentiment analysis. In *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*.

Sreenivasulu, M., & Sridevi, M. (2020). Comparative study of statistical features to detect the target event during disaster. *Big Data Mining and Analytics, 3*(2), 121–130.

Subramani, S., Wang, H., Vu, H. Q., & Li, G. (2018). Domestic violence crisis identification from facebook posts based on deep learning. *IEEE Access, 6*, 54075–54085.

Sufi, F. K., & Khalil, I. (2022). Automated disaster monitoring from social media posts using AI-based location intelligence and sentiment analysis. *IEEE Transactions on Computational Social Systems*. https://doi.org/10.1109/TCSS.2022.3157142

Sun, K., Li, Z., & Zhao, H. (2023). Cross-lingual universal dependency parsing only from one monolingual treebank. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. https://doi.org/10.48550/arXiv.2012.13163

Suwaileh, R., Elsayed, T., Imran, M., & Sajjad, H. (2022). When a disaster happens, we are ready: Location mention recognition from crisis tweets. *International Journal Of Disaster Risk Reduction, 78*, Article 103107.

Thushara, M., Krishnapriya, M., & Nair, S. S. (2017). A model for auto-tagging of research papers based on keyphrase extraction methods. In *2017 International conference on advances in computing, communications and informatics (ICACCI)*.

Tounsi, A., & Temimi, M. (2023). A systematic review of natural language processing applications for hydrometeorological hazards assessment. *Natural Hazards, 116*(3), 2819–2870.

Wang, D., Su, J., & Yu, H. (2020). Feature extraction and analysis of natural language processing for deep learning English language. *IEEE Access, 8*, 46335–46345. https://doi.org/10.1109/ACCESS.2020.2974101

Wang, H., & Li, F. (2022). A text classification method based on LSTM and graph attention network. *Connection Science, 34*(1), 2466–2480.

Widyastuti, D. A. R. (2021). Using new media and social media in disaster communication. *Komunikator, 13*(2), 100–111.

Yang, C., Wang, Q., & Pan, S. (2024). Identification and threshold analysis of strong winds and heavy rain disaster factors based on frequent-pattern mining. *Urban Climate, 56*, Article 102015.

Yang, T., Xie, J., Li, G., Mou, N., Li, Z., Tian, C., & Zhao, J. (2019). Social media big data mining and spatio-temporal analysis on public emotions for disaster mitigation. *ISPRS International Journal of Geo-Information, 8*(1), 29. https://doi.org/10.3390/ijgi8010029

Yuan, X., Chang, W., Zhou, S., & Cheng, Y. (2018). Sequential pattern mining algorithm based on text data: taking the fault text records as an example. *Sustainability, 10*(11), 4330.

Zamir, M. T., Ullah, F., Tariq, R., Bangyal, W. H., Arif, M., & Gelbukh, A. (2024). Machine and deep learning algorithms for sentiment analysis during COVID-19: A vision to create fake news resistant society. *Plos one, 19*(12), Article e0315407.

Zhang, M., Huang, Q., & Liu, H. (2022). A multimodal data analysis approach to social media during natural disasters. *Sustainability, 14*(9), 5536.

Zhou, B., Zou, L., Mostafavi, A., Lin, B., Yang, M., Gharaibeh, N., Cai, H., Abedin, J., & Mandal, D. (2022). VictimFinder: Harvesting rescue requests in disaster response from social media with BERT. *Computers, Environment and Urban Systems, 95*, Article 101824.

Zhou, H. (2022). Research of text classification based on TF-IDF and CNN-LSTM. *Journal of Physics: Conference Series*.