RESEARCH

Detecting information from Twitter on landslide hazards in Italy using deep learning models

Rachele Franceschini[1,3]*, Ascanio Rosi[2], Filippo Catani[2] and Nicola Casagli[1,3]

Abstract

Background Mass media are a new and important source of information for any natural disaster, mass emergency, pandemic, economic or political event, or extreme weather event affecting one or more communities in a country. Several techniques have been developed for data mining in social media for many natural events, but few of them have been applied to the automatic extraction of landslide events. In this study, Twitter has been investigated to detect data about landslide events in Italian-language. The main aim is to obtain an automatic text classification on the basis of information about natural hazards. The text classification for landslide events in Italian-language has still not been applied to detect this type of natural hazard.

Results Over 13,000 data were extracted within Twitter considering five keywords referring to landslide events. The dataset was classified manually, providing a solid base for applying deep learning. The combination of BERT + CNN has been chosen for text classification and two different pre-processing approaches and bert-model have been applied. BERT-multicase + CNN without preprocessing archived the highest values of accuracy, equal to 96% and AUC of 0.96.

Conclusions Two advantages resulted from this studio: the Italian-language classified dataset for landslide events fills that present gap of analysing natural events using Twitter. BERT + CNN was trained to detect this information and proved to be an excellent classifier for the Italian language for landslide events.

Keywords Social media, Text classification, Landslide events, Deep learning

Introduction

The use of social media in detecting natural hazards is showing promising results (Holderness and Turpin 2015; Wang et al. 2018). The retrieval of data from technical reports and/or newspapers (Barman et al. 2021), using specific data mining algorithms also for different languages (Pennington et al. 2022), extends exploitable data for event detection and assessment. Studies indicate that social media can be regarded as sensors to detect the inception of natural disasters much faster than observatories (Goswami et al. 2018; O'Halloran et al. 2021). Such characteristic provides a unique opportunity to capture situations with a relatively high temporal and spatial resolution. Crowdsourcing platforms have become an indispensable part of people's everyday lives and a powerful tool for communication during emergencies. People post situation-sensitive information on social media related to what they are experiencing, witnessing, and/or hearing from other sources (Hughes and Palen 2009). Crowdsourcing platforms such as Twitter (since 2023 named

*Correspondence:
Rachele Franceschini
rfranceschini@ogs.it
[1] National Institute of Oceanography and Applied Geophysics - OGS, Borgo Grotta Gigante N. 42/C, Sgonico (Trieste), Italy
[2] Department of Geosciences, University of Padova, Via G. Gradenigo, 6, 35131 Padua, Italy
[3] Department of Earth Sciences, University of Firenze, Via Giorgio La Pira, 4, 50121 Florence, Italy

Springer Open

"X") are being used as a medium of communication every day and the amount of various information that can be found is overwhelming (Madichetty and Sridevi 2020). Twitter had more than 353.1 million active users in 2020 (Gilmary et al. 2023), with an average rate of 0.85–3% of tweets being geo-tagged, and approximately 7,000,000 geo-tagged tweets are posted per day (Huang et al. 2018). Tweets featured short messages (maximum of 280 characters) published in real-time, the ability to attach pictures and share GPS geolocation, and the provision of an application programming interface (API) made it possible to perform monitoring tasks (Fayjaloun et al. 2020). Twitter is widely used to detect different types of events using different parts of data provided by users, such as text, images, and video. Large crises often generate an explosion of social media activity. Researchers have observed a strong and immediate spread of tweets when a significant event happens (Comunello et al. 2016; Kryvasheyeu et al. 2016). Nowadays, many studios use Twitter as a source of data in difficult circumstances, such as elections (Di Giovanni et al. 2018; Chen et al. 2022; Eady et al. 2023), humanitarian crises such as pandemics (Hussain et al. 2021; Naseem et al. 2021) or wars (Pierri et al. 2023; Moreno-Mercado and Calatrava-García 2023), natural disasters such as hurricanes (Wang and Zhuang 2017; Karimiziarani and Moradkhani 2023; Florath and Keller 2022; Florath et al. 2024), floods (Gul et al. 2018; Kankanamge et al. 2020), wildfires (Wang et al. 2016; Zander et al. 2023), earthquakes (Sakaki et al. 2010; Earle et al. 2011; Splendiani and Capriello 2022), landslides (Ofli et al. 2022, 2023; Pennington et al. 2022) and infrastructure damages (Kozlowski et al. 2020). The earliest known cases of people using the microblogging service Twitter in an emergency occurred during severe wildfires near San Diego, California (United States) in 2007 (Imran et al. 2015). The largest documented peaks of tweets per minute observed during disasters are 20,000 tweets per minute during Hurricane Sandy in 2012 (United States of America) (Castillo 2016), approximately 17,000 tweets per minute during the Notre Dame fire in 2019 in France (Kozlowski et al. 2020) and more than 150,000 tweets in the first 48 h after the Mw 6.2 Amatrice earthquake (Italy, August 2016) (Francalanci et al. 2017).

1.1. Leveraging social media during crisis

Social networking sites have multiple roles: enhancing situational awareness, accelerating information spreading, and monitoring (Luna and Pennock 2018). Researchers have applied social media data to study disaster management in recent years due to these benefits (Li et al. 2021). For example, Kryvasheyeu et al. (2016) examined Twitter activities of 50 metropolitan areas in the United States before, during, and after

the 2012 Hurricane Sandy. Their analysis manifested that the spatiotemporal distribution of disaster-related messages could support the real-time monitoring and assessment of disaster damage. Another example is provided by Pennington et al. (2022), which use a semantic analysis in different languages to detect landslide events over the world within Twitter. In particular, they harvest and classify photos and images (Ofli et al. 2022, 2023) attachment on tweet text.

The other dominant use of social media information in disaster crises lies in communication and cooperation (Li et al. 2021). Social networks can help explore interactions of online users and understand the critical roles for emergency information propagation (Kim and Hastak 2018). During a natural disaster, help is needed for people who are endangered, and it is important to support the spread of information (Dragović et al. 2019).

After a crisis, information regarding who needs help, the locations of vulnerable people, and the regions with substantial damage (Dragović et al. 2019) is shared. For example, volunteers from Tufts University after the earthquake in Haiti created a map that has helped survivors and volunteers send rescue information via messages on Twitter. Within 15 days, more than 2500 messages were received (Fraustino et al. 2012; Gao et al. 2011). Having easy access to tweets coming from people would afford new possibilities for emergency response, such as a contribution to the real-time assessment of impacts, criticalities, and needs.

Textual analysis of vocabulary, words, grammar, and other features such as text sentiment provides information that can be leveraged in the post-event environment (Li et al. 2021). Such textual analysis is now widely used in sociology, psychology, marketing, and elsewhere to conclude what appears to be relatively descriptive and qualitative information (Das et al. 2019; Mahoney et al. 2019; Majumdar and Bose 2019; Osorio-Arjona and García-Palomares 2019; Plunz et al. 2019; Reboredo and Ugolini 2018). Much of the recent research on social media and disasters capturing textual analysis converges on themes of early warning (Wu and Cui 2018), damage assessment (Fan et al. 2018), and behaviour analysis (Li et al. 2021). For example, Li et al. (2020) tracked the online public's mental and behavioural responses following a large-scale blackout in New York City in 2019. The research revealed that sentiment analysis and online textual information could help imply community resilience in a place. Li et al. (2021) leveraged text classification pipelines to develop a rapid damage assessment model for earthquake events based on Twitter data, while Fan et al. (2018) applied text similarity calculations and developed a graph-based method for detecting credible

situation information related to infrastructure disruptions in hurricane disasters.

The social media-based method advances the use of big data and citizen-as-sensor approaches, offering an additional technique and near-real-time information source for decision-makers to understand emergency evacuation patterns. Previous studies have shown the potential of social media data and textual analysis to support disaster response and recovery. However, limited research exists on extracting and understanding information dissemination across social media networks.

Related work

The progress in text mining and machine learning techniques provides valuable tools for extracting information from social media data. Li et al. (2021) employed text classification pipelines to create a swift damage assessment model for earthquake events using Twitter data. Zahra et al. (2020) integrated textual features (bag-of-words) with machine learning classifiers to identify eyewitness messages during flood, earthquake, and hurricane disasters. Similarly, Devaraj et al. (2020) employed word embeddings with neural networks to categorize urgent help requests during hurricanes. The use of a deep learning model has started to become more common for natural language processing (NLP) tasks. Numerous deep learning models with Bidirectional Encoder Representations from transformers (BERT) architecture have been proposed in these few years for text classification of tweet text, such as sentiment analysis (Alaparthi and Mishra 2021; Geetha and Renuka 2021) or event detection (Jain et al 2019;

Madichetty and Muthukumarasamy 2020; Liu et al 2021; Huang et al 2022; Dharma and Winarko 2022; Zhou et al. 2022; Lai et al. 2022) using different languages (see Table 1). Various methodologies were applied for classifying crises, either using the BERT classifier alone or incorporating the classifier with other deep learning techniques. Jain et al. (2019), Liu et al. (2021), and Zhou et al. (2022) used the BERT method alone to check whether the text describes a crisis or not. Jain et al. (2019) demonstrated the effectiveness of the classifier for flood, hurricane, and earthquake events. Both studies used a binary classification (0 and 1) estimating a maximum accuracy of 95% with Liu et al. (2021). Zhou et al. (2022), use the BERT method for recovery activities after a natural event, classifying text according to aid information, complete interridge, and victims. Zhou et al. (2022), have chosen original English tweets containing the 5-digit zip code of coastal Texas as potential rescue request tweets. This study for each label compares a different model, also implementing other deep learning techniques, such as convolutional neural networks (CNN), or Long short-term memory (LSTM) in the BERT model. Huang et al. (2022) and Dharma and Winarko (2022) implemented other deep learning models to the BERT model to improve text classification. In particular, the implementation of CNN with BERT embedding is predominant, and Dharma and Winarko (2022) used this combination to delineate whether text is informative with respect to a crisis or not. This study, by Dharma and Winarko (2022), uses extracted Twitter data for different natural events: earthquakes, floods, pyroclastic flows,

Table 1 Different approaches analysing events

Authors	Target	Event	Language	Model	Results	
Jain et al. (2019)	Relevant to crisis or no	Earthquake	English	BERT	P	83%
		Flood			R	77%
		Hurricane			F1	78%
					A	77%
Liu et al. (2021)	Relevant to crisis or no	Crisislex	English	DistilBERT	F1	95%
					A	95%
Zhou et al. (2022)	Help or no help	Hurricane	English	BERT + CNN	P	89%
					R	93%
					F1	90%
Dharma and Winarko (2022)	Relevant to crisis or no	Earthquake, Flood, Pyroclastic flow, Drought, Typhoon, Tsunami, Landslide, Eruption	Indonesian	CNN + BERT	P	97.63%
					R	96.64%
					F1	97.13%
					A	97%
Sánchez et al. (2022)	Relevant to crisis or no	Flood	Italian	XLM-RoBERTa	F1	84%
		Earthquake		MT-BERT	F1	82%

Results the metrics have been represented using the abbreviation, R: Recall; P: Precision; F1, and A: Accuracy

eruptions, tsunamis, drought, landslides, typhoons, and others. These data constitute a single dataset in the Indonesian language. The data were manually classified with a Boolean value in which the value 1 is distributed to the data with disaster and a non-disaster was classified with 0. CNN with pre-trained BERT embedding was able to get the best result (Dharma and Winarko 2022). The accuracy was 97.16%.

At the same time, Sánchez et al. (2022) introduced a new multilingual and multi-domain crisis dataset, containing 53 crisis events and more than 160,000 messages. They proposed an empirical transfer learning, using crisis data from high-resource languages (such as English) to classify data from other languages (such as Italian, Spanish, and French). The authors used different models for the binary classification of tweets that are related and unrelated to the crisis. Considering the Italian-language, the authors considered data from flood events (on Sardinia and Genova events) and earthquake events (on L'Aquila events). Both datasets come from SoSItalyT4 by Cresci et al. (2015). The best performance, for a flood event, was archived in the Cross-lingual & Multi-domain scenario, using the XLM-RoBERTa model (Conneau and Lample 2019). In this scenario, the training featured multiple domains in one language (e.g. floods, earthquakes,

and hurricanes in English) while the test set was characterized by a new event in another language (e.g. flood in Italian). The highest F1 value was 84%.

A high performance was achieved using Multilingual & Multi-domain scenarios with Machine Translation (MT)+BERT for earthquake events. In this case, the model was trained with English and Italian tweets about floods, earthquakes, and hurricanes to then classify earthquake-related messages in Italian. The best value of F1 was 82%.

Research context

Italy, with an area of almost 300,000 km^2, is divided into 20 regions (Fig. 1a), 107 provinces, and 7926 municipalities, with 158 Hydrological Warning Zones (HZWs) based on morphology and basin boundaries. The country is predominantly hilly and mountainous (Guzzetti 2000), with the Alps and the Apennines (Fig. 1b) as the main mountain ranges. Italy is particularly susceptible to hydrogeological instability and is the European country with the widest territorial distribution and the highest recurrence of large landslides (Salvati et al. 2010; Avvisati et al. 2019). The IFFI database (Italian Inventory of Landslides; Trigila et al. 2007) includes over 600,000 landslides affecting 7.9% of the national territory. Every year,
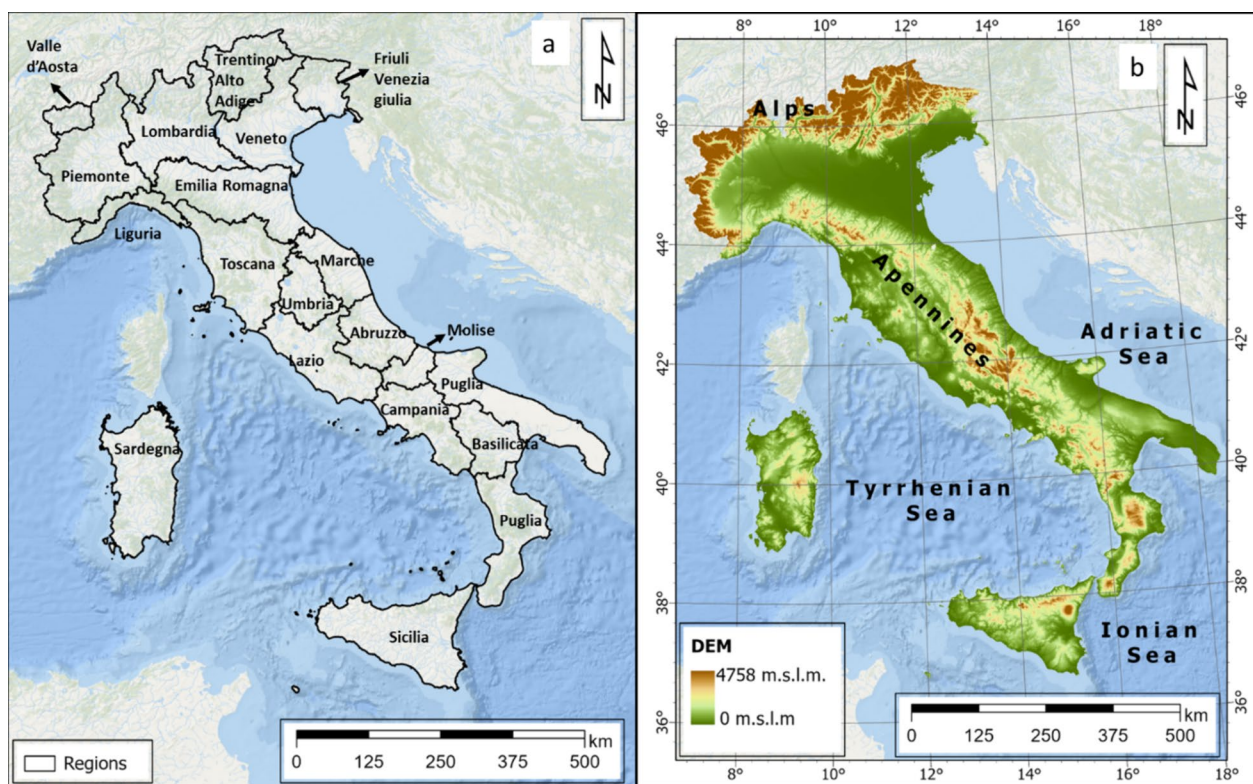

Fig. 1 **a** Regions of Italy. **b** Digital elevation model (DEM)

Franceschini *et al. Geoenvironmental Disasters*   (2024) 11:22

Page 5 of 14

thousands of landslides occur, some causing casualties, evacuations, and damage to infrastructure. For example, in 2017, there were 172 events (Trigila and Iadanza 2018). Legambiente (2021) documented 1181 extreme weather events from 2010, causing damage in 637 municipalities and resulting in 264 human casualties. The National Research Council (CNR) recorded the evacuation of over 27,000 people between 2016 and 2020, rising to 320,000 since 1971. The regions most affected by extreme events since 2010 are Sicily and Lombardy with 144 and 124 events, respectively (Franceschini et al. 2022a, b).

Material and methods

Several techniques have been developed for data mining in social media for many natural events, but little research has been produced for automatic extraction of landslide events using text as the primary source. This work focused on landslide event detection and text classification using data extracted from Twitter. Data mining technique on Twitter was applied using appropriate keywords extracted from newspaper headlines (Franceschini et al. 2022a) in the Italian-language (Fig. 2). The starting dataset, composed of 13,349 tweets text, was classified manually considering landslide information provided by users. The dataset constitutes a solid base for applying deep learning using transformer architecture. There is a large literature on transformer models and BERT is

currently one of the most effective language models in terms of performance when different NLP tasks like text classification are concerned (Velankar et al. 2023). The literature review, presented in Sect. 1.1, has shown how BERT captures the language context in an efficient way (Jawahar et al. 2019).

The combination of BERT + CNN by Mozafari et al. (2020) was chosen for classifying tweet data with binary classification (0 and 1).

Data collection and analysis

Franceschini et al. (2022a) analysed the word frequency within online newspaper headlines about landslide events. The most frequent words have been chosen as keywords for collecting and analysing tweets. On this regard, the most relevant keywords used have been: "frana", "smottamento", "scivolamento", "crollo", "dissesto" (translation: "landslide", and its synonyms as: "slip", "fall", 'instability"). Overall, 13,349 Tweets were collected through Twitter API in several periods from 2011 to 2019. Periods have been selected based on the temporal distribution of newspaper articles analysed by Franceschini et al. (2022a, b). Each slot has tweets describing one or more landslide events and no landslide events. This allows the best model performance to be defined. From Twitter, it was possible to obtain different metadata
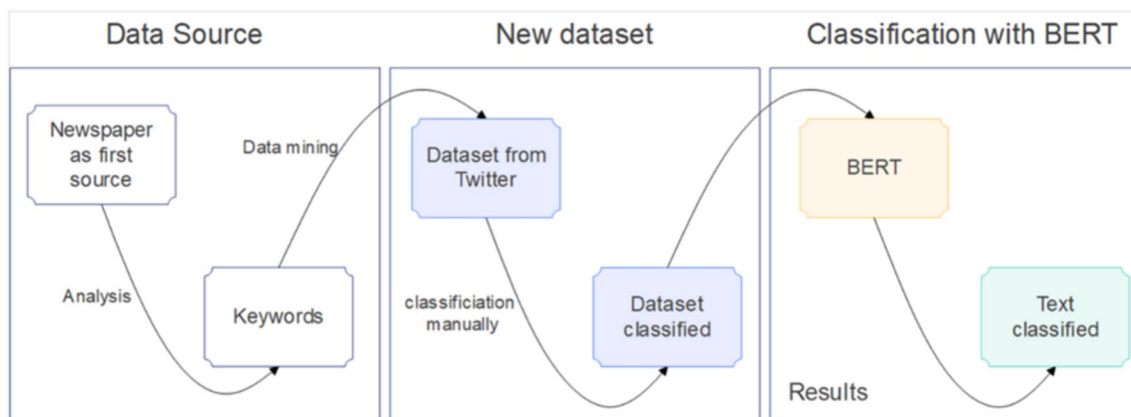

Fig. 2 Workflow for the implementation of this research

Table 2 Metadata and features list from Twitter. In bold the tweets text

Metadata	Features
User object	Fid, author_id, name, username, author_created_at, author_description, author_entities, author_location, public_metrics, author_url, author_verified
Tweet Object	Id_text, **text**, created_at, lang, entities, public metrics, in_reply_to_user_id, source
Place objects	Geo

such as User object, Tweet Object, and Place objects (Table 2).

The tweet database underwent manual classification based on the relevance of the landslide event. The classification features by 2 labels: landslide Yes (label 1) or No (label 0) (Table 3). The manual classification has been created to construct the ground truth for the training set and validation for deep learning techniques. According to the classification criteria, 4806 tweets were labelled as "landslide" and 8544 tweets were classified as "no landslide".

Data augmentation

The dataset was randomly divided into 80% for training and 10% for testing. The training dataset features 10,812 data spread in 6883 with label 0 ("no landslide") and 3829 data with label 1 ("landslide"). This dataset was further randomly divided by 10%, resulting in the validation dataset featured by 1202 data with 765 data in label 0 and 437 within label 1. The test set is characterized by 1335 data, spread in 850 with label 0 and 485 with label 1 (Table 4).

Model: bidirectional encoder representations from transformers (BERT)

Bidirectional Encoder Representations from Transformers (BERT), a multi-layer bidirectional transformer encoder, is pre-trained on the English Wikipedia and the Book Corpus, containing 2500 M and 800 M tokens, respectively. As our task involves classifying text into "landslide or "no landslide", a critical step is to analyse contextual information extracted from the pre-trained layers of BERT. Following this, we fine-tune the model using annotated datasets specific to our 'landslide' and 'no landslide' classification task, updating the weights with labelled data.

BERT processes input sequences with a maximum length of 512 tokens, producing a 768-dimensional vector representation of the sequence. Two special tokens, [CLS] (classification embedding) and [SEP] (segment separator), are inserted into each input sequence by BERT-multicase. The [CLS] embedding, positioned at the

Table 4 Summarize of data distribution for each dataset

Dataset	Count	Label 1—Landslide	Label 0—no landslide
Training	10,812	3829	6883
Validation	1202	437	765
Test	1335	485	850

sequence's outset, contains the special representation for classification. We extract the representation of the entire sequence from the first [CLS] token in the final hidden layer for the "landslide"/"no landslide" classification task. While BERT uses [SEP] for segment separation, we do not utilize it in our classification task. Different layers of a neural network capture various levels of syntactic and semantic information. The lower layers may capture more general information, while the higher layers tend to contain task-specific details (Devlin et al. 2018).

Pre-processing and Word embedding methods

The complete database was pre-processed using two different procedures (Table 5). In the first case, different parameters have been removed and changed:

Removed: HTML special entities, Italian stop words, tickers, numbers, hyperlinks, hashtags, punctuation, special characters, words with 2 or fewer letters, multiple whitespaces including new line characters, single space remaining at the front of the tweet and emoticons.

Modified: @username to AT_USER, and put all lowercase.

In the second case, no pre-processing was applied.

For each pre-processing two models were chosen for tokenization: BERT-multicase by "tensorflow/bert_multi_cased_preprocess/1" (Devlin et al. 2018) and DistilBERT-multicased by "jeongukjae/distilbert_multi_cased_preprocess/2".
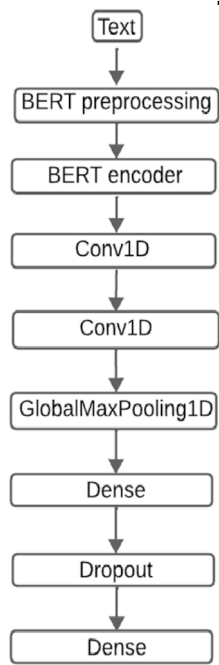
After tokenization, the tokens are converted into embedding vectors through the embedding process. BERT includes inside itself an embedding mechanism of words. Each word is represented as a vector with

Table 3 Examples of manual classification for tweet text considering information about landslides

Examples of Tweet text	Landslide yes or no
#SanremoNews Frana sulla Statale 20 ad Airole: questa mattina riunione tra i Sindaci in Prefettura *#SanremoNews Landslide on the 20th road in Airole: meeting between mayors and prefecture this morning*	Yes
Quale migliore idea di un bel #pontesullostretto: reggio calabria, crollato il tetto della sala calipari, auditorium di palazzo campanella, sede del consiglio regionale *what better idea than a nice #pontesullostretto: reggio calabria, collapsed roof of calipari hall, auditorium of campanella palace, seat of the regional council*	No

English translations are reported in italic

Table 5 BERT + CNN architecture

Layer (type)	Output Shape		Param distilbert_multi cased	Param bert_multi_cased
Text	None, 128		0	0
preprocessing	Distilbert : 'input word ids' 'input mask' Bert : 'input mask' 'input type ids' 'input words ids'	None, 128	0	0
BERT_encoder	None, 128, 768		134734080	177853441
Conv1d	None, 127, 32		49184	49184
Conv1d	None, 126, 64		4160	4160
GlobalMaxPooling1D	None, 64		0	0
Dense	None, 512		33280	33280
Dropout	None, 512		0	0
Classifier (dense)	None, 3		1539	1539

Total params: 134822243
Trainable params: 134822243
Non-trainable params: 0

Total params: 177941604
Trainable params: 177941603
Non-trainable params: 1

(Flow diagram on left: Text → BERT preprocessing → BERT encoder → Conv1D → Conv1D → GlobalMaxPooling1D → Dense → Dropout → Dense)

constant dimension and BERT takes into consideration both the preceding and subsequent contexts. These embedding vectors contain contextual information and represent how each token contributes to the context of the entire text sequence.

Fine-tuning strategies

Two different transformer encoders have been applied: bert_multi_cased_L-12_H-768_A-12/3 and distilbert_multi_cased_L-6_H-768_A-12/1. This allows to adapt the model to specific "landslide" and "no landslide" classification tasks, taking advantage of the strengths offered by both models in capturing contextual information effectively.

BERT-multicase is employed to take advantage of its ability to capture contextual representations of language in different languages. The choice to use BERT-multicase is based on its ability to handle case-sensitive text effectively, making it suitable for a wide range of language contexts. This pre-trained model provides a model with the ability to capture complex semantic and syntactic relations, which can later be exploited for binary text classification during fine-tuning. In this way, our model benefits from the richness of linguistic representations provided by BERT-multicase to best fit the specific classification task we are addressing.

On the other hand, DistilBERT-multicase is a lighter and computationally efficient alternative to BERT-multicase. DistilBERT-multicase, while retaining the essential qualities of BERT-multicase, allows for greater efficiency in computational resources, making fine-tuning more accessible on platforms with computational power limitations. This choice allows to maintain of high contextual language representation performance while optimizing available resources. Further details about these transformer encoder architectures are provided in (Devlin et al. 2018; Mozafari et al. 2020).

CNN layers have been added at the end of each transformer encoder this allows for employing all output from architecture. Concatenating the output vectors of each transformer encoder to create a matrix. Following this, a convolutional operation (CNN) is executed with

a window of size (3, hidden size of BERT, which is 768 in BERT-multicase and DistilBERT-multicase model), and the maximum value is derived for each transformer encoder through max-pooling on the convolution output (Table 5). By concatenating these values, a vector is formed, serving as input to a fully connected network. The classification operation is then performed by applying softmax to the input (Mozafari et al. 2020).

The parameters list in Table 6 have been maintained constant during training and for each interaction. The maximum number of epochs for the training is set to 80. An early stopping of 15 has been set to avoid unnecessary iterations with a lack of improvement during training.

Assessment metrics

The metrics used to evaluate the performance of the model were accuracy, confusion matrix, precision, recall, F1 score, and the calculation of the AUC with the rock curve.

- Accuracy is the ratio of the number of correct predictions to the total number of input samples.

$$Accuracy = \frac{number\ of\ correct\ predictions}{total\ number\ of\ predictions\ made} \quad (1)$$

- The confusion matrix describes the complete performance of the model. True Positive (TP) and True Negative (TN) are data predicted correctly, in contrast to False Negative (FN) and False Positive (FP).

	Predicted : no or 0	**Predicted: yes or 1**
Real: no or 0	TN	FP
Real: yes or 1	FN	TP

- Precision is the number of correct positive results divided by the number of positive results predicted by the classifier.

$$Precision = \frac{TruePositives}{TruePrositive + FalsePositives} \quad (2)$$

- Recall is the number of correct positive results divided by the number of all relevant samples.

$$Recall = \frac{TruePositives}{TruePrositive + FalseNegatives} \quad (3)$$

- The F1 score is the harmonic mean between precision and recall. The range is [0, 1]. This parameter describes the precision of the classification and its robustness.

$$F1\ score = 2*\frac{1}{\frac{1}{precision} + \frac{1}{recall}} \quad (4)$$

- The receiver operating characteristic (ROC) curve and the area under the curve (AUC) of a classifier are equal to the probability that the classifier will rank a randomly chosen positive example higher than a randomly chosen negative example. ROC is a probability curve, and AUC represents the degree of separability. The higher the AUC is, the better the model is at predicting. The ROC curve is plotted with a True Positive Rate (TPR) or sensitivity against a False Positive Rate (FPR).

$$TPR = \frac{True\ Positive}{False\ Negative + True\ Positive} \quad (5)$$

$$FPR = \frac{False\ Positive}{True\ Negative + False\ Positive} \quad (6)$$

Results

The manually classified dataset provides a solid base for applying deep learning to classify tweets. BERT+CNN was trained on the aforementioned classified dataset in the Italian-language with the goal of distinguishing whether a tweet describes a landslide event. Overall, four tests have been carried out, changing preprocessing and transformer encoders (BERT-multicase and DistilBERT-multicase). The model "BERT+CNN with bert_multi_cased_L-12_H-768_A-12/3." has collected the highest values of accuracy, even 96.10% after 19 epochs (Table 7).

For a further analysis of the results, a confusion matrix and ROC curve with AUC were utilised (Fig. 3a, b). In total, 1283 tweets out of 1335 data have been correctly classified: "no landslide" represented in white and "landslide" in grey in Fig. 3a. Only 16 FN and 36 FP were harvested. Figure 3b describes the ROC curve with an area under the curve of 0.96.

Comparison and validation dataset

A validation was conducted comparing newspaper articles collected using the SeCaGN (Semantic Engine to Classification and Geotagging News) model (Battistini et al. 2013) and Twitter. The period was selected based on significant publication dates of articles during the Vaia

Table 6 Parameter settings for the model

Max length	Batch size	Epoch	Optimizer	Learning rate	Early stopping
128	32	80	Adam	3e-5	15

Values have been retrieved from the state-of-the-art and retained constants (such as Max length, Batch size, and Learning rate)

Table 7 Metrics for each model

Dataset	Model	Pre-trained parameters	A (%)	P (%)	R (%)	F1 (%)
Without pre-processing	BERT+CNN	distilbert_multi_cased_L-6_H-768_A-12/1	95.50	93.27	94.43	93.85
		bert_multi_cased_L-12_H-768_A-12/3	**96.10**	**92.87**	**96.70**	**94.74**
Pre-processing applied		distilbert_multi_cased_L-6_H-768_A-12/1	93.55	88.73	94.22	91.40
		bert_multi_cased_L-12_H-768_A-12/3	93.78	90.52	92.57	91.55

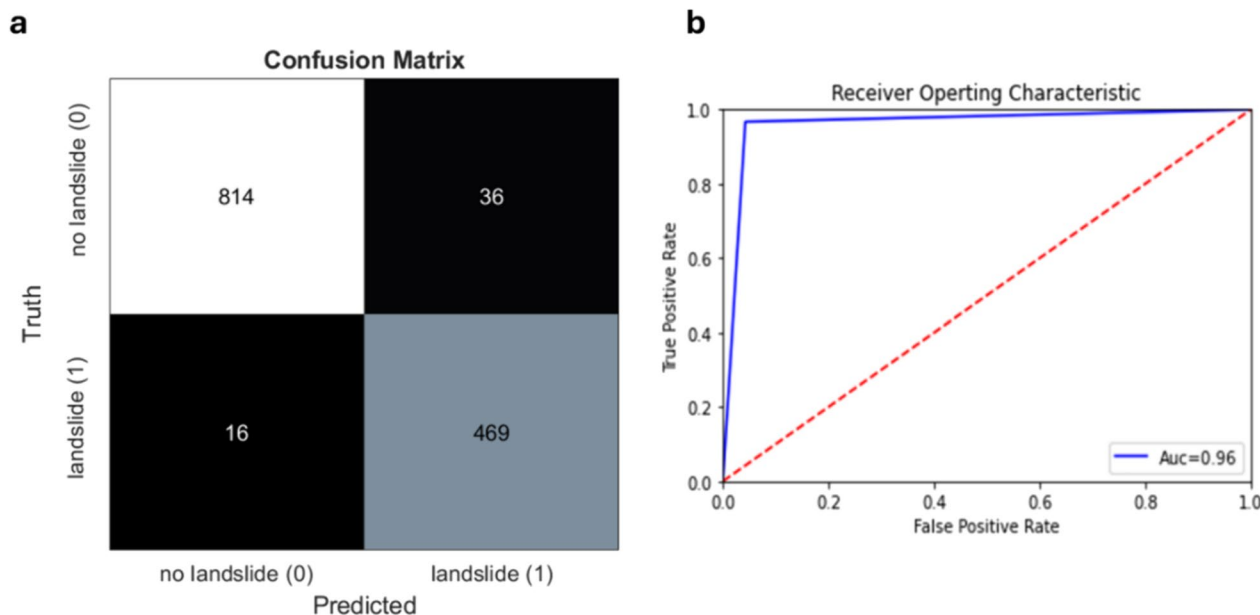Higher metrics are highlighted in bold. Accuracy (A); Precision (P); Recall (R) and F1

Fig. 3 After training, the model was tested on the classified dataset. The confusion matrix in **a** and the ROC curve and AUC in **b**

event, which affected Northeast Italy from October 27th to 30th, 2018 (Rainato et al. 2021; Meena et al. 2022). This storm caused landslides, floods, interruptions in electric supply, road traffic disruptions, and the most severe forest loss ever documented in Italy, totalling about 8.5 million m^3 of growing stock felled over 41,000 hectares, due to extremely high wind gusts, storm surges, and heavy precipitation, as reported by Biolchi et al. (2019) and Cavaleri et al. (2019).

The timeframe for social media analysis (Google News and Twitter) spans from October 26th to November 28th, 2018. This period was chosen based on the occurrence of landslides during and after the event, as well as significant publication dates of articles. The newspaper articles identified by the SeCaGN (Battisti et al. 2013) were classified into three classes, as suggested by Franceschini et al. (2022a), based on landslide information, localization, and time. The dataset comprises 2829 data points in class 1 (date and location of landslide can be retrieved from the news), 1524 in class 2 (articles referred to landslide, but

with low spatial and temporal accuracy), and 1 in class 0 (news not related to landslides), totalling 4354 data points.

A new dataset from Twitter, consisting of 16,026 data points, was extracted and subjected to classification using the BERT+CNN model with the highest accuracy score. Overall, 12,882 tweets were identified as not related to landslides (label 0), while 3144 tweets pertained to natural hazards (label 1).

Both datasets were primarily compared based on temporal distribution (Fig. 4) rather than spatial distribution. The classification of tweets did not consider the precise localization of landslides; thus, the sum of news in classes 1 and 2 was considered. Peaks of articles and tweets occurred on October 28th, with 731 and 602, respectively (Fig. 4). Additional distribution matches were observed on November 2nd, 4th, 6th, and 11th, 2018.

The correlation allowed to validate the ability of code extraction and classification of the data of tweets, in line with what was reported by newspapers.
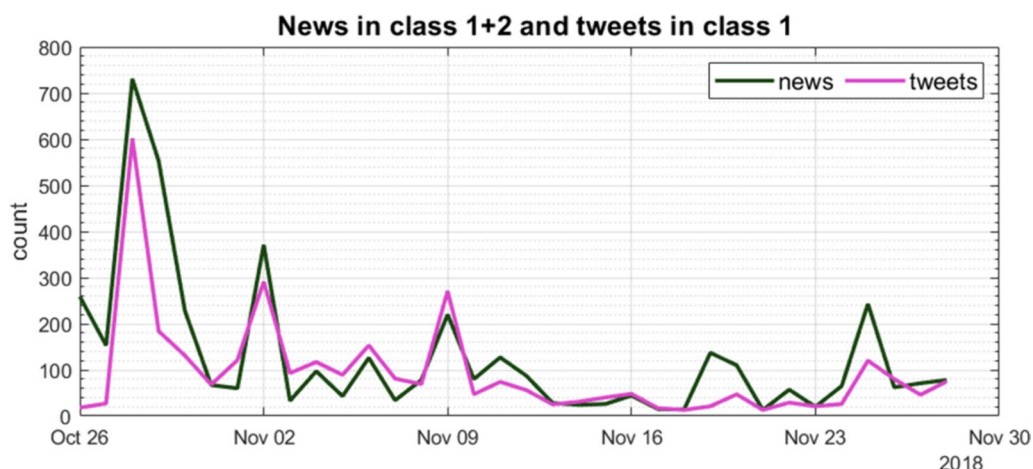
Fig. 4 Daily distribution analysis between news in class 1 and 2 from Google News and tweets in class 1

Discussion

Using Twitter as a data source, this work demonstrates how data mining techniques can be employed to collect and generate datasets tailored to specific events, such as landslide occurrences. The keywords and period time of extraction were chosen based on the analysis of newspaper articles presented by Franceschini et al. (2022a, b). However, even with the overwhelming amount of data that can be found on Twitter, often it is not enough to use keywords alone to obtain useful tweets (Nguyen et al. 2017). As demonstrated by Zhou et al. (2022), several issues may arise due to the nature of big social media data. Tweet analysis tends to favour those who use social media more often. Uneven usage of social media may lead to biased consequences. Moreover, social media posts suffer from locational bias, temporal bias, and reliability issues. In some cases, the presence/absence of tweets could be affected by other factors such as disruption in communication services, socio-demographic factors (the events affecting socially vulnerable populations get less attention) and absence of points of attraction. These factors can alter the real distribution of event hazard, leading from one hand to underestimate the presence of data in rural areas, forest or without a journalistic relevance and to the other hand to overestimate the hazards in most relevant areas from a journalistic point of view. Such issues should be considered while further analysing the spatiotemporal patterns of the identified tweets for detecting vulnerable communities or assessing disaster damages (Zhou et al. 2022).

The dataset from Twitter was subjected to a binary classification based on landslide information. Overall, 4806 tweets labelled as "landslide" (class 1), and 8544 tweets have been classified as "no landslide" (class 0). The high number of tweets classified in class 0 demonstrates

the difficult handling and ambiguity that characterize the data from Twitter. Therefore, a strong filtering system must be applied to handle these data.

The manually classified dataset provided a solid base for applying deep learning using the natural language technique. The combination between BERT + CNN by Mozafarti et al. (2020) has been chosen for classification. From the language modelling perspective, BERT + CNN uses all the information included in different layers of pre-trained BERT during the fine-tuning phase and adding level CNN on fine tuning it is possible to obtain the best performance. This information contains both syntactical and contextual features coming from lower layers to higher layers of BERT. The model has been iterated considering a binary classification (0 and 1) based on landslide information in the Italian language. This analysis led to a considerable advancement of the BERT classifier, which until now was very often used for a variety of analyses in the English language and different fields, as confirmed by Gasparetto et al. (2022). Two advantages resulted from this work:

  (i) the Italian-language classified dataset for landslide events fills that present gap of analysing natural events using Twitter, which has not yet been exploited to a great extent;
  (ii) BERT + CNN was trained to detect this information and proved to be an excellent classifier for the Italian language for landslide events.

Two tests were carried out to obtain text classification, changing the setup of the pre-processing. This procedure was necessary to outline the best setting of the data cleaning parameters. Based on the results obtained for each model, the best trade-off is represented without pre-processing. In fact, it showed important values of

Franceschini *et al. Geoenvironmental Disasters* (2024) 11:22

Page 11 of 14

accuracy equal to 96% and AUC of 0.96. The application of any pre-processing completely undermines the text, risking changing the context and meaning. This results in a considerable loss of data. On the one hand, the non-application of pre-processing before the model resulted in a better classification of the data, on the other hand, the time of training significantly increased. However, compared with the state of the art, the proposed model improves on the performance of Liu et al. (2021) and nearly equals it Dharma and Winarko (2022).

A correlation between newspaper articles and Twitter was applied to outline a validation of model. The lack of coordinate by Twitter allowed analysing only of the temporal distribution of tweets. The daily distribution of each variable, about extreme Vaia event, showed a good temporal correlation. If the tweets presented a decrease almost immediately at the event, the news presented echoes in the next days. The reason can be linked to several aspects:

i) newspaper news needs more steps for publication than a tweet;
ii) the event(s) present an impact distribution over many days, hence articles also are published in the following days;
iii) the consequences of the event or the damage caused are felt in the following days, so the articles are published repeatedly over time.

From a practical point of view, this study provides useful perspectives for decision-makers to consider when using social media as an additional information resource for rapid damage assessment. BERT + CNN makes possible the detection of landslide events within tweets and brings state-of-the-art integration in NLP technology of text classification. At the same time, several problems may arise due to the nature of big social media data analysis and some limitations of this research. These problems should not be ignored when translating the research results into practice.

However, nowadays it needs to be given that Twitter limited the possibility of extracting data in real-time, due to API changes. Further, a subscription has to be necessary to obtain data.

Limitations in this study should be acknowledged for future research. Owing to privacy concerns, the geographical information of tweets is only accessible if users explicitly choose to share it. Consequently, it becomes challenging to ensure that a tweet or retweet accurately depicts the actual location in the event of a landslide. The main limitation of this approach is that Geo-tagged tweets are limited to only 1% of the whole incoming tweets (Singh et al. 2024). Few researchers have used the

content-based techniques. An approach more efficient than GPS and use-profile locations has been proposed by Vieweg et al. (2010), and Giridhar et al. (2015). Authors claim that there are more chance than users mention the event location in the text.

Ongoing and further work will pay attention to the improvement of the text classification models and then define the location of the tweet that describes a landslide event.

Conclusion

Several techniques have been developed for data mining in social media for many natural events, but they have rarely been applied to landslide events. This makes it possible to fill the gap in the literature with respect to this natural hazard. With appropriate keywords in Italian-language, a dataset was harvested using Twitter as source. The database features 13,349 tweets and was classified manually, providing a solid base for applying deep learning. The Italian-language classified dataset for landslide events fills that present gap in analysing natural events. The combination BERT + CNN has been used to classify the dataset into two classes (0 and 1) based on landslide information. This analysis leads to a considerable advancement of the BERT + CNN classifier, which until now was very often used to analyse data in the English language for different fields. The model demonstrated an efficient classification task. The best performance recorded without pre-processing showed significant accuracy, equal to 96% and an AUC of 0.96, locating itself between implementing models with CNN. This study confirms that relevant information on landslide hazards can be obtained by data mining from Twitter, even during emergencies. Such data, properly filtered and classified, may be of notable help in increasing our present capability of calibrating and validating early warning models, with particular reference to data-scarce areas and back-analysis of undocumented past events.

Franceschini *et al. Geoenvironmental Disasters*   (2024) 11:22

Page 12 of 14

ROC   Receiver operating characteristic
SeCaGN   Semantic engine to classification and geotagging news
SEP   Segment separator
TN   True negative
TP   True positive
TPR   True positive rate

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author upon reasonable request.

Declarations

Competing interests

The authors declare that they have no competing interests.

References

Alaparthi S, Mishra M (2021) BERT: a sentiment analysis odyssey. J Mark Anal 9(2):118–126

Avvisati G, Sessa EB, Bellucci E, Colucci O, Marfè B, Marotta E, Nave R, Peluso R, Ricci T (2019) Tomasone M (2019) Perception of risk for natural hazards in Campania Region (Southern Italy). Int J Dis Risk Red 40:101164

Barman R, Ehrmann M, Clematide S, Oliveira SA, Kaplan F (2021) Combining visual and textual features for semantic segmentation of historical newspapers. J Data Min Digit Humanit

Battistini A, Segoni S, Manzo G, Catani F, Casagli N (2013) Web data mining for automatic inventory of geohazards at national scale. Appl Geogr 147–158.

Biolchi S, Denamiel C, Devoto S, Korbar T, Macovaz V, Scicchitano G et al (2019) Impact of the October 2018 storm Vaia on coastal boulders in the northern Adriatic Sea. Water 11(11):2229

Castillo C (2016) Big crisis data: social media in disasters and time-critical situations. Cambridge University Press, London

Cavaleri L, Bajo M, Barbariol F, Bastianini M, Benetazzo A, Bertotti L et al (2019) The October 29, 2018 storm in Northern Italy—an exceptional event and its modeling. Prog Oceanogr 178:102178

Chen E, Deb A, Ferrara E. (2022) # Election2020: the first public Twitter dataset on the 2020 US Presidential election. J Comput Soc Sci 1–18

Comunello F, Parisi L, Lauciani V, Magnoni F, Casarotti E (2016) Tweeting after an earthquake: user localization and communication patterns during the 2012 Emilia seismic sequence. Ann Geophys 59(5):0537

Conneau A, Lample G (2019) Cross-lingual language model pretraining. Adv Neural Inf Process Syst 32

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 social web for disaster management (SWDM'15) workshop

Das S, Dutta A, Medina G, Minjares-Kyle L, Elgart Z (2019) Extracting patterns from Twitter to promote biking. IATSS Res 43(1):51–59. https://doi.org/10.1016/j.iatssr.2018.09.002

Devaraj A, Murthy D, Dontula A (2020) Machine-learning methods for identifying social media-based requests for urgent help during hurricanes. Int J Disaster Risk Reduct 51:101757

Devlin J, Chang M W, Lee K, Toutanova K (2018) Bert: pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805

Dharma LSA, Winarko E (2022) Classifying natural disaster tweet using a convolutional neural network and BERT embedding. In: 2022 2nd international conference on information technology and education (ICIT&E), pp 23–30. https://doi.org/10.1109/ICITE54466.2022.9759860

Di Giovanni M, Brambilla M, Ceri S, Daniel F, Ramponi G. (2018) Content-based classification of political inclinations of twitter users. In: 2018 IEEE international conference on big data (big data). IEEE, pp 4321–4327

Dragović N, Vasiljević Đ, Stankov U, Vujičić M (2019) Go social for your own safety! Review of social networks use on natural disasters—case studies from worldwide. Open Geosci 11(1):352–366. https://doi.org/10.1515/geo-2019-0028

Eady G, Paskhalis T, Zilinsky J, Bonneau R, Nagler J, Tucker JA (2023) Exposure to the Russian Internet Research Agency foreign influence campaign on Twitter in the 2016 US election and its relationship to attitudes and voting behavior. Nat Commun 14(1):62

Earle PS, Bowden D, Guy M (2011) Twitter earthquake detection: earthquake monitoring in a social world. Ann Geophys 54(6):708–715

Fan C, Mostafavi A, Gupta A, Zhang C (2018) A system analytics framework for detecting infrastructure-related topics in disasters using social sensing in Advanced Computing Strategies for Engineering. Springer, Cham, pp 74–91

Fayjaloun R, Gehl P, Auclair S, Boulahya F, Guérin-Marthe S, Roullé A (2020) Integrating strong-motion recordings and Twitter data for a rapid shakemap of macroseismic intensity. Int J Disast Risk Reduct 52:101927. https://doi.org/10.1016/j.ijdrr.2020.101927

Florath J, Keller S (2022) Supervised machine learning approaches on multispectral remote sensing data for a combined detection of fire and burned area. Remote Sens 14(3):657

Florath J, Chanussot J, Keller S (2024) Rapid natural hazard extent estimation from twitter data: investigation for hurricane impact areas. Nat Hazards 1–22

Francalanci C, Guglielmino P, Montalcini M, Scalia G, Pernici B (2017) IMEXT: a method and system to extract geolocated images from Tweets—analysis of a case study. In: 2017 11th international conference on research challenges in information science (RCIS). IEEE, pp 382–390

Franceschini R, Rosi A, Catani F, Casagli N (2022a) Exploring a landslide inventory created by automated web data mining: the case of Italy. Landslides 1–13

Franceschini R, Rosi A, del Soldato M, Catani F, Casagli N (2022b) Integrating multiple information sources for landslide hazard assessment: the case of Italy. Sci Rep 12:20724

Fraustino JD, Brooke L, Yan J (2012) Social media use during disasters: a review of the knowledge base and gaps. Final Report to Human Factors/Behavioural Sciences Division, Science and Technology Directorate, U.S. Department of Homeland Security, National consortium for the study of terrorism and responses to terrorism (START), College Park, Maryland

Gao H, Barbier G, Goolsby R (2011) Harnessing the crowdsourcing power of social media for disaster relief. IEEE Intell Syst 26(3):10–14. https://doi.org/10.1109/MIS.2011.52

Gasparetto A, Zangari A, Marcuzzo M, Albarelli A (2022) A survey on text classification: practical perspectives on the Italian language. PLoS ONE 17(7):e0270904

Geetha MP, Renuka DK (2021) Improving the performance of aspect based sentiment analysis using fine-tuned Bert Base Uncased model. Int J Intell Netw 2:64–69

Gilmary R, Venkatesan A, Vaiyapuri G (2023) Detection of automated behavior on Twitter through approximate entropy and sample entropy. Pers Ubiquit Comput 27(1):91–105

Giridhar P, Abdelzaher T, George J, Kaplan L (2015) On quality of event localization from social network feeds. In: 2015 IEEE international conference on pervasive computing and communication workshops (PerCom workshops). IEEE, pp 75–80

Goswami S, Chakraborty S, Ghosh S, Chakrabarti A, Chakraborty B (2018) A review on application of data mining techniques to combat natural

Franceschini *et al. Geoenvironmental Disasters*   (2024) 11:22

Page 13 of 14

disasters. Ain Shams Eng J 9:365–378. https://doi.org/10.1016/j.asej.2016.01.012

Gul S, Shah TA, Ahad M, Mubashir M, Ahmad S, Gul M, Sheikh S (2018) Twitter sentiments related to natural calamities: analysing tweets related to the Jammu and Kashmir floods of 2014. Electron Libr 36(1):38–54

Guzzetti F (2000) Landslide fatalities and evaluation of landslide risk in Italy. Eng Geol 58:89–107

Holderness T, Turpin E (2015) From social media to geosocial intelligence: crowdsourcing civic co-management for flood response in Jakarta, Indonesia. In: Social media for government services. Springer, pp 115–133

Huang Y, Li Y, Shan J (2018) Spatial-temporal event detection from geo-tagged tweets. ISPRS Int J Geo Inf 7(4):150. https://doi.org/10.3390/ijgi7040150

Huang L, Shi P, Zhu H, Chen T (2022) Early detection of emergency events from social media: a new text clustering approach. Nat Hazards 111(1):851–875

Hughes AL, Palen L (2009) Twitter adoption and use in mass convergence and emergency events. Int J Emerg Manage 6(3):248–260

Hussain A, Tahir A, Hussain Z, Sheikh Z, Gogate M, Dashtipour K, Sheikh A (2021) Artificial intelligence–enabled analysis of public attitudes on facebook and twitter toward covid-19 vaccines in the united kingdom and the united states: observational study. J Med Internet Res 23(4):e26627

Imran M, Castillo C, Diaz F, Vieweg S (2015) Processing social media messages in mass emergency: a survey. ACM Comput Surv (CSUR) 47(4):67

Jain P, Ross R, Schoen-Phelan B (2019) Estimating distributed representation performance in disaster-related social media classification. In: 2019 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM). IEEE, pp 723–727

Jawahar G, Sagot B, Seddah D (2019) What does BERT learn about the structure of language? In: Proceedings of the 57th annual meeting of the association for computational linguistics. Association for Computational Linguistics, Florence, Italy, pp 3651–3657. Available: https://aclanthology.org/P19-1356

Kankanamge N, Yigitcanlar T, Goonetilleke A, Kamruzzaman M (2020) Determining disaster severity through social media analysis: testing the methodology with South East Queensland Flood tweets. Int J Disaster Risk Reduct 42:101360

Karimiziarani M, Moradkhani H (2023) Social response and Disaster management: insights from twitter data Assimilation on Hurricane Ian. Int J Disaster Risk Reduct 95:103865

Kim J, Hastak M (2018) Social network analysis: characteristics of online social networks after a disaster. Int J Inform Manage 38(1):86–96. https://doi.org/10.1016/j.ijinfomgt.2017.08.003

Kozlowski D, Lannelongue E, Saudemont F, Benamara F, Mari A, Moriceau V, Boumadane A (2020) A three-level classification of French tweets in ecological crises. Inf Process Manage 57:102284. https://doi.org/10.1016/j.ipm.2020.102284

Kryvasheyeu Y, Chen H, Obradovich N, Moro E, Van Hentenryck P, Fowler J, Cebrian M (2016) Rapid assessment of disaster damage using social media activity. Sci Adv 2(3):e1500779

Lai K, Porter JR, Amodeo M, Miller D, Marston M, Armal S (2022) A natural language processing approach to understanding context in the extraction and geocoding of historical floods, storms, and adaptation measures. Inf Process Manage 59(1):102735

Legambiente (2021) Osservatorio Città Clima: https://cittaclima.it/2021/12/29/emergenza-clima-il-bilancio-del-2021-dellosservatorio-cittaclima/

Li L, Ma Z, Cao T (2020) Leveraging social media data to study the community resilience of New York City to 2019 power outage. Int J Disaster Risk Reduct 51:101776

Li L, Bensi M, Cui Q, Baecher GB, Huang Y (2021) Social media crowdsourcing for rapid damage assessment following a sudden-onset natural hazard event. Int J Inf Manag 60:102378. https://doi.org/10.1016/j.ijinfomgt.2021.102378

Liu J, Singhal T, Blessing LT, Wood KL, Lim KH (2021) Crisisbert: a robust transformer for crisis classification and contextual crisis embedding. In: Proceedings of the 32nd ACM conference on hypertext and social media, pp 133–141

Luna S, Pennock MJ (2018) Social media applications and emergency management: a literature review and research agenda. Int J Disaster Risk Reduct 28:565–577

Madichetty S, Muthukumarasamy S (2020) Detection of situational information from Twitter during disaster using deep learning models. Sādhanā 45(1):1–13

Madichetty S, Sridevi M (2020) Improved classification of crisis-related data on Twitter using contextual representations. Proc Comput Sci 167:962–968

Mahoney J, Le Moignan E, Long K, Wilson M, Barnett J, Vines J, Lawson S (2019) Feeling alone among 317 million others: disclosures of loneliness on Twitter. Comput Hum Behav 98:20–30. https://doi.org/10.1016/j.chb.2019.03.024

Majumdar A, Bose I (2019) Do tweets create value? A multi-period analysis of Twitter use and content of tweets for manufacturing firms. Int J Prod Econ 216:1–11. https://doi.org/10.1016/j.ijpe.2019.04.008

Meena SR, Puliero S, Bhuyan K, Floris M, Catani F (2022) Assessing the importance of conditioning factor selection in landslide susceptibility for the province of Belluno (region of Veneto, northeastern Italy). Nat Hazard 22(4):1395–1417

Moreno-Mercado JM, Calatrava-García A (2023) Multilingual public diplomacy: strategic communication of Israeli Defence Forces (IDF) in Twitter during operation guardian of the walls. Media, War & Conflict 16(2):282–299

Mozafari M, Farahbakhsh R, Crespi N (2020) A BERT-based transfer learning approach for hate speech detection in online social media. In: Complex networks and their applications VIII: Volume 1 proceedings of the eighth international conference on complex networks and their applications COMPLEX NETWORKS 2019 8. Springer, pp 928–940

Naseem U, Razzak I, Khushi M, Eklund PW, Kim J (2021) COVIDSenti: a large-scale benchmark Twitter data set for COVID-19 sentiment analysis. IEEE Trans Comput Soc Syst 8(4):1003–1015

Nguyen DT, Al-Mannai KA, Joty S, Sajjad H, Imran M, Mitra P (2017) Robust classification of crisis-related data on social networks using convolutional neural networks. In: Proceedings of the 11th international conference on web and social media, ICWSM 2017. AAAI Press, pp 632–635

Ofli F, Imran M, Qazi U, Roch J, Pennington C, Banks V, Bossu R (2023) Landslide detection in real-time social media image streams. Neural Comput Appl 35(24):17809–17819

Ofli F, Qazi U, Imran M, Roch J, Pennington C, Banks V, Bossu R (2022) A real-time system for detecting landslide reports on social media using artificial intelligence. In: International conference on web engineering. Springer, Cham, pp 49–65

O'Halloran KL, Pal G, Jin M (2021) Multimodal approach to analysing big social and news media data. Discourse Context Med 40:100467

Osorio-Arjona J, García-Palomares JC (2019) Social media and urban mobility: using twitter to calculate home-work travel matrices. Cities 89:268–280. https://doi.org/10.1016/j.cities.2019.03.006

Pennington CV, Bossu R, Ofli F, Imran M, Qazi U, Roch J, Banks VJ (2022) A near-real-time global landslide incident reporting tool demonstrator using social media and artificial intelligence. Int J Disaster Risk Reduct 77:103089

Pierri F, Luceri L, Chen E, Ferrara E (2023) How does Twitter account moderation work? Dynamics of account creation and suspension on Twitter during major geopolitical events. EPJ Data Sci 12(1):43

Plunz RA, Zhou Y, Carrasco Vintimilla MI, Mckeown K, Yu T, Uguccioni L, Sutto MP (2019) Twitter sentiment in New York City parks as measure of well-being. Landsc Urban Plan 189:235–246. https://doi.org/10.1016/j.landurbplan.2019.04.024

Rainato R, Martini L, Pellegrini G, Picco L (2021) Hydrological, geomorphic and sedimentological responses of an alpine basin to a severe weather event (Vaia storm). CATENA 207:105600

Reboredo JC, Ugolini A (2018) The impact of Twitter sentiment on renewable energy stocks. Energy Econ 76:153–169. https://doi.org/10.1016/j.eneco.2018.10.014

Sakaki T, Okazaki M, Matsuo Y (2010) Earthquake shakes twitter users: real-time event detection by social sensors. In: Proceedings of the 19th international conference on world wide web, pp 851–860

Salvati P, Bianchi C, Rossi M, Guzzetti F (2010) Societal landslide and flood risk in Italy. Nat Hazards Earth Syst Sci 10:465–483

Sánchez C, Sarmiento H, Pérez J, Abeliuk A, Poblete B (2022) Cross-lingual and cross-domain crisis classification for low-resource scenarios. arXiv preprint arXiv:2209.02139

Singh J, Pandey D, Singh AK (2024) Event detection from real-time twitter streaming data using community detection algorithm. Multimed Tools Appl 83(8):23437–23464

Splendiani S, Capriello A (2022) Crisis communication, social media and natural disasters–the use of Twitter by local governments during the 2016 Italian earthquake. Corp Commun Int J 27(3):509–526

Trigila A, Iadanza C (2018) Landslides and floods in Italy: hazard and risk indicators-Summary Report 2018. Report number: 267bis/2018. Affiliation: Institute for Environmental Protection and Research (ISPRA). Project: National risk indicators. https://doi.org/10.13140/RG.2.2.14114.48328

Trigila A, Iadanza C, Guerrieri L (2007) The IFFI project (Italian landslide inventory): methodology and results. Guidelines for Mapping Areas at Risk of Landslides in Europe. In: Hervás J (eds). ISPRA, Rome, Italy, pp 15–18

Velankar A, Patil H, Joshi R (2023) Mono vs multilingual bert for hate speech detection and text classification: a case study in marathi. In: IAPR workshop on artificial neural networks in pattern recognition. Springer, Cham, pp 121–128

Vieweg S, Hughes AL, Starbird K, Palen L (2010) Microblogging during two natural hazards events: what twitter may contribute to situational awareness. In: Proceedings of the SIGCHI conference on human factors in computing systems. ACM, pp 1079–1088

Wang B, Zhuang J (2017) Crisis information distribution on Twitter: a content analysis of tweets during Hurricane Sandy. Nat Hazards 89:161–181

Wang Z, Ye X, Tsou MH (2016) Spatial, temporal, and content analysis of Twitter for wildfire hazards. Nat Hazards 83:523–540

Wang RQ, Mao H, Wang Y, Rae C, Shaw W (2018) Hyper-resolution monitoring of urban flooding with social media and crowdsourcing data. Comput Geosci 111:139–147

Wu D, Cui Y (2018) Disaster early warning and damage assessment analysis using social media data and geo-location information. Decis Support Syst 111:48–59

Zahra K, Imran M, Ostermann FO (2020) Automatic identification of eyewitness messages on twitter during disasters. Inf Process Manage 57(1):102107

Zander KK, Garnett ST, Ogie R, Alazab M, Nguyen D (2023) Trends in bushfire related tweets during the Australian 'Black Summer'of 2019/20. For Ecol Manage 545:121274

Zhou B, Zou L, Mostafavi A, Lin B, Yang M, Gharaibeh N, Mandal D (2022) VictimFinder: harvesting rescue requests in disaster response from social media with BERT. Comput Environ Urban Syst 95:101824

Publisher's Note