Twitter-Based Flood Damage Mapping Using Text Classification and Geoparsing Techniques

Zhasman Amirgholizadeh*, Mohammad Hassan Vahidnia, Ali Akbar Matkan

Center for Remote Sensing and GIS Research, Faculty of Earth Sciences, Shahid Beheshti University, Tehran, Iran
z.amirgholizade@mail.sbu.ac.ir, mh_vahidnia@sbu.ac.ir, a-matkan@sbu.ac.ir

KEY WORDS: Text Classification, Geoparsing, Natural language processing, Natural Disaster, Machine Learning, Social Media

ABSTRACT:

Social media is a prominent source of real-time information for disaster understanding. This study provides a method that uses natural language processing (NLP), machine learning, and geoparsing for flood damage mapping and classification from Twitter data. It addresses a key research gap in prior work, which has primarily focused on classifying tweets as either damage-related or non-damage. In contrast, the framework proposed in this study categorizes tweets into multiple risk classes, enabling more detailed assessment and enhancing spatial resolution by analyzing risk at the city level rather than providing only a broad overview. For this purpose, we retrieved a dataset of 3,000 tweets from India and Pakistan from the CrisisNLP database. After cleaning the text, 1,000 tweets were manually annotated into three damage classes. Three machine learning classifiers—Support Vector Machine (SVM), Random Forest (RF), and Logistic Regression—were trained after applying Term Frequency-Inverse Document Frequency (TF-IDF) vectorization. SVM performed best in terms of accuracy, precision, and recall. The trained models were used to label the remaining 2,000 tweets. For spatial analysis, a rule-based geoparsing strategy was used with a curated list of states and cities, and geographic coordinates were retrieved using the Geopy library. Tweets were then grouped by location, and predicted flood damage trends were mapped in ArcGIS. Validation involved visual comparison of satellite image before and after the flood, confirming damage detection in selected cities. Results indicate that combining social media analysis with geospatial techniques can effectively assess flood damage in areas lacking organized or geotagged data.

1. INTRODUCTION

Natural disasters—including earthquakes, floods, storms, and droughts—are responsible for approximately 40,000 to 50,000 deaths globally each year. Although this represents a relatively small proportion of total global deaths, natural disasters have devastating impacts on local communities, often triggering mass migration and substantial economic losses. Each year, millions are displaced or rendered homeless, and these economic losses are especially difficult to recover from in impoverished nations (Ritchie et al. 2022). Such disasters disproportionately affect vulnerable populations, resulting in heightened psychological, emotional, and physical stress. Research shows that these events exacerbate health outcomes among marginalized communities, particularly in areas where disaster response is inadequate (Benevolenza & DeRigne, 2019).

The effects of natural disasters are not only immediate but also long-term, creating serious challenges for economic and social development. While both developed and developing countries are affected, the economic consequences are typically more severe in low-income nations. In many developing countries, disasters can cause a sudden increase in poverty levels, pushing vulnerable populations into deeper economic hardship. Disasters destroy livelihoods, infrastructure, and access to essential services, making recovery even more difficult (Kreimer, 2001). Effective disaster management plays a crucial role in mitigating these negative impacts. Evidence suggests that well-designed policies—particularly those focused on early warning systems, community preparedness, infrastructure resilience, and coordinated response—can significantly reduce human and economic losses. These policies must be inclusive and account for the unique vulnerabilities of different population groups, including women, children, the elderly, and individuals with disabilities. The intersection of systems theory, social capital theory, and vulnerability theory underscores the importance of a collective, community-driven approach to disaster response (Samuel, 2024).

Since the mid-1990s, social media has transformed how individuals communicate during emergencies and disasters. Platforms such as Twitter, Facebook, and YouTube are increasingly used to share real-time information, issue warnings, check on the safety of others, and support disaster response efforts. While most emergency management agencies currently use social media passively—for instance, posting updates—there is growing recognition of its potential as an active emergency management tool. This includes issuing alerts, receiving assistance requests, and conducting damage assessments. As evidence of successful social media use in disaster response accumulates, interest is growing in formally integrating these tools into emergency planning (Lindsay, 2011).
Social media has opened the public sphere to more decentralized and participatory communication during disasters. For example, during the 2007 Southern California wildfires, unofficial information shared via social media—so-called "backchannel" practices—became authoritative sources of real-time updates,

* Corresponding author

even as their legitimacy was questioned by official agencies. These developments reflect the increasing importance of community-generated information in crisis situations and signal a shift toward more open and responsive disaster management systems (Sutton et al. 2008).

Recent research highlights the growing importance of processing social media data during large-scale crises. Imran et al. (2015) argue that platforms such as Twitter have the potential to provide real-time situational information during emergencies. Accordingly, various computational methods have been developed to filter, classify, and extract valuable information from the massive volume of messages available online. However, detecting damage-related content on Twitter remains a challenging task, as such messages constitute only a small subset of situational data. To improve detection, the SESIW methodology was proposed, combining statistical features (e.g., hashtags, mentions, digits) with informative terms extracted using Term Frequency-Inverse Document Frequency (TF-IDF). This approach outperformed conventional models in classifying damage-related tweets (Madichetty & Sridevi, 2019).

The concept of "citizens as sensors" has gained traction, and Volunteered Geographic Information (VGI) sourced from social media is increasingly central to disaster monitoring. With advances in deep learning and natural language processing (NLP), it is now possible to extract valuable information—such as flood depth and damage severity—from social media data and integrate it with remote sensing data to enhance disaster analysis (Feng et al. 2022).

In addition, machine learning models offer powerful tools for interpreting language patterns and identifying meaningful content. As deep learning and neural network techniques have evolved, NLP methods have improved significantly, enabling automatic feature extraction and enhanced performance in tasks such as sentiment analysis, entity recognition, and text classification. Given the rising frequency of natural disasters and the growing complexity of crisis management, leveraging social media data—particularly from Twitter—as a critical source of big data is highly valuable.

Most current research examining social media data in disaster contexts has focused primarily on detecting disaster-related content and mapping its geographical distribution. However, many of these studies do not account for the severity of the damage, which limits their utility for effective response planning and prioritization. Unlike previous work that categorizes tweets simply as damage or non-damage and generates damage maps based on the frequency of relevant tweets over broad spatial units such as grids or counties (Ahadzadeh & Malek, 2021), our study introduces a more nuanced, multi-class classification of damage severity (e.g., moderate, severe, extreme) directly from raw tweet text. This study aims to apply NLP and machine learning techniques to extract damage-related information from raw textual data to support rapid decision-making during disasters.

Moreover, rather than conducting analysis at large administrative levels, we assess damage at a finer spatial resolution—namely, the city level—allowing for more precise and actionable flood damage mapping. This increased granularity in both classification and spatial resolution represents a key advancement and strength of our approach. For instance, studies such as Zou, He, Wang, and Liang (2025) focus on the spatial distribution of disaster impacts but overlook intensity. Addressing this gap is critical for enhancing our understanding of disasters using unstructured, real-time text data from platforms like Twitter. This study seeks to fill that gap by extracting not only geographic location information but also varying levels of damage severity from raw tweet texts, thereby enabling detailed city-level flood damage assessments.

2. METHODOLOGY

2.1 Data Collection and Preprocessing

The data used in this study was originally presented by Imran et al. (2016), who provide detailed documentation of the data collection and annotation process. The dataset comprises 3,000 tweets related to flood disasters in India and Pakistan that occurred in **September 2014**, sourced from the CrisisNLP website. To evaluate the robustness of the proposed model, data from both countries were combined into a single dataset. Of the total, 1,000 tweets were manually labeled for supervised training, while the remaining 2,000 tweets were reserved for prediction and spatial analysis.

Tweet texts were pre-processed using the Python NLTK library prior to classification. The preprocessing steps included: (1) Removal of URLs and user tags (e.g., "@username"); (2) Conversion of all text to lowercase; (3) Removal of punctuation; (4) Tokenization using word_tokenize from NLTK; (5) Removal of stopwords using NLTK's English stopword list; (6) Stemming using the PorterStemmer algorithm.

The processed text was stored in a separate column labeled clean_text and preserved for downstream tasks. This cleaned and standardized textual corpus significantly improved the performance of the subsequent models.

2.2 Text Classification

For supervised learning, 1,000 tweets were manually classified based on a scheme developed by the researcher, taking into account the number of reported deaths and the extent of infrastructure damage. The tweets were categorized into four classes: one for irrelevant content and three representing varying levels of flood damage severity. Examples of Tweets and manual damage assessment is provided in **Table 1**. The details of the classes interpretations are as follows:

Class 0 (Irrelevant): Tweets that are unrelated to flood damage or contain vague/general information not useful for evaluation.

Class 1 (Damage1): Tweets indicating moderate damage, typically reporting fewer than 50 deaths and minimal or no infrastructure loss.

Class 2 (Damage2): Tweets indicating heavy damage, usually reporting 50–200 deaths along with mentions of infrastructure damage (e.g., road blockages, destroyed homes).

Class 3 (Damage3): Tweets reporting severe damage, characterized by over 200 fatalities and substantial infrastructure destruction.

Damage Score	Tweet Text (Original)
0	I love you but boy. Do you need 1,000 youtube channels?!!! I cant keep up!

1	Top India Stories From WSJ: Nepal Flooda, Landslides Kill At Least54
2	180 Dead as Floods Wash Away Homes in Nepal, India-ABC News
3	#MaikyLinaresHere Scores Killed in Flooding in Nepal and India: More than 11,000 homes have been damaged and r….

Table 1. Examples of Tweets and Manual Damage Assessment

To prepare the textual data for classification, the Term Frequency–Inverse Document Frequency (TF-IDF) method was used to convert the preprocessed tweets into numerical feature vectors (Qaiser & Ali, 2018). TF-IDF helps to highlight terms that are important in a specific tweet but rare across the overall dataset. The TF-IDF score of a term t in a document d is calculated as follows:

$$TF - IDF = TF(t,d) \times log\left(\frac{N}{DF(t)}\right) \quad (1)$$

where *t* represents term, *d* denotes document, *TF*(*t, d*) stands for term frequency in document *d*, *DF*(*t*) represents document frequency of term *t*, and *N* denotes the total number of documents.

Using TF-IDF vectors, this study employed classical machine learning models due to the relatively small or moderate size of the dataset, including Logistic Regression (LR), Random Forest (RF), and Support Vector Machine (SVM). These models are well-established in text classification tasks because of their robustness and interpretability in handling datasets of limited size (Vahidnia, 2024). They were subsequently used to predict class labels for the remaining 2,000 tweets, which were then utilized for further spatial and damage analysis.

2.3 Geoparsing and Mapping

Following the classification process, the 2,000 predicted tweets were further processed to extract spatial information. A rule-based geoparsing approach was employed, utilizing a hand-curated gazetteer of Indian states and cities. Regular expression pattern matching with word boundaries was applied to the sanitized tweet texts to detect exact city name mentions. If no city name was identified, a secondary search was conducted for state names as a fallback option. This process generated two new columns in the dataset: matched_city and matched_state.

For each city identified in the tweet texts, geographic coordinates (latitude and longitude) were obtained using the Geopy library and the Nominatim geocoding service. If a city was not found in one country (India or Pakistan), the query was retried in the other country, as the dataset contained locations from both nations. A delay of 1.5 seconds was introduced between requests to reduce the risk of service blocking.

The outcome was a geocoded list of urban areas, each associated with a specific damage class (Damage1, Damage2, or Damage3) as predicted by the classifiers. These classifications were then tabulated using pandas.crosstab to analyze the distribution of damage classifications across cities and states. For example, if a city had 156 tweets labeled as Damage2, it was categorized as experiencing "high damage." The aggregated spatial data was exported to Excel for visualization.

Damage maps were created using ArcGIS to illustrate the geographic distribution and intensity of flood impacts. Additionally, satellite images from before and after the flood events were manually reviewed for each city to qualitatively evaluate the accuracy of the damage classification and mapping results. **Figure 1** illustrates the workflow of the proposed methodology for flood damage mapping and detection from Twitter data.

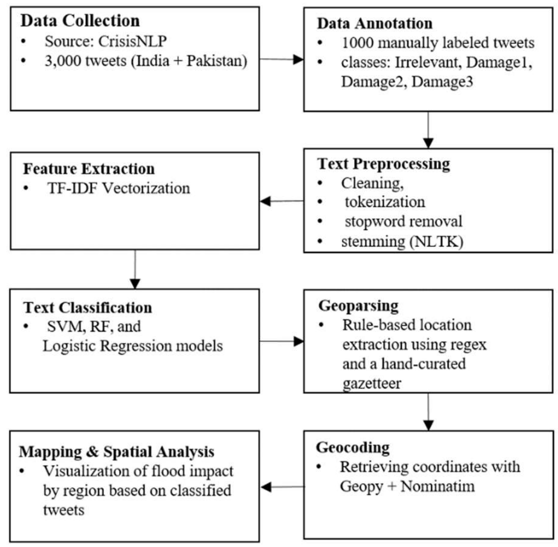

Figure 1. Workflow of the Proposed Methodology for Twitter-Based Flood Damage Detection and Mapping

3. IMPLEMENTATION AND RESULTS

3.1 Model Performance Evaluation

The performance of three machine learning models in classifying damage-labeled tweets was evaluated. Accuracy, precision, recall, and F1-score metrics are reported for each model.

The Logistic Regression model achieved a satisfactory overall accuracy of 0.79 in classifying the tweets. Precision across the different classes ranged from 0.71 to 1.00, with the highest precision (1.00) observed for Class 0 (no damage). However, the recall for Class 0 was lower at 0.66, suggesting that many true examples of this class were misclassified. The model's average F1-score was approximately 0.79, indicating a moderate balance between precision and recall.

The Random Forest model slightly outperformed Logistic Regression, yielding an accuracy of 0.80. It showed improved recall for Class 0 at 0.73 compared to 0.66 in Logistic Regression. The precision for Class 0 was 0.97, indicating strong performance in correctly identifying no-damage instances. The model achieved a mean F1-score of 0.80, reflecting a notable improvement in the balance of performance metrics.

The Support Vector Machine (SVM) model achieved the highest overall accuracy at 0.815 among the three models. Its precision, recall, and F1-scores were equal to or better than those of the other models across all damage classes. Notably, for Class 3 (severe damage), the SVM achieved a precision of 0.94 and a recall of 0.83, demonstrating strong capability in identifying high-damage tweets. The average F1-score for the SVM model was 0.82, suggesting an optimal trade-off between precision and recall.

For each machine learning algorithm to perform optimally, different settings were used for the hyperparameters. For example, in the case of the Random Forest algorithm, different hyperparameters were considered, such as the number of trees to be used and the depth limit, where the best combination was to use 200 trees and 15 depth limit, giving an accuracy and F1 score of 0.800. The best combination of hyperparameters was considered in other algorithms, where the SVM gave the best results with the linear kernel parameter and C = 1, while in the case of the Logistic Regression algorithm, it was saga with C = 2.

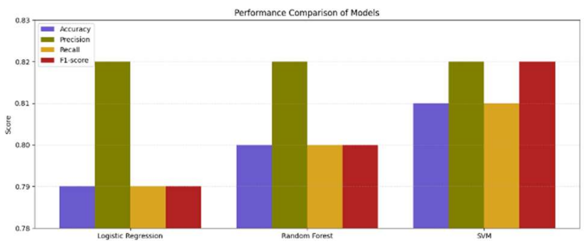

Figure 2. Performance Comparison of the Three Models (Accuracy, Precision, Recall, and F1-Score across All Damage Classes)

Model	Hyperparameters Tested	Best Configuration
Support Vector Machine (SVM)	kernel = {linear, rbf}, C = {1, 10}	kernel =linear, C = 1
Random Forest (RF)	n_estimators = {100, 200}, max_depth = {5, 10, 15}	n_estimators = 200, max_depth = 15
Logistic Regression (LR)	Solver={liblinear, saga}, C = {0.5, 1, 2}	solver = saga, C = 2

Table 2. Hyperparameter Tuning of Machine Learning Models

Overall, the SVM model proved to be the most suitable for the classification of damage-labeled tweets, based on the evaluation metrics. While both Random Forest and Logistic Regression also performed well, the SVM showed superior accuracy and F1-score, making it the most appropriate model for this task.

the performance comparison of the three models is presented in Figure 2, which graphically represents the average values of the accuracy, precision, recall, and F1-score of each model for all damage classes. These results are further supported in **Table 2** by providing the best combination of hyperparameters for each model based on the resulting values of accuracy and F1-score.

3.2 State-level Damage Mapping

To determine the extent and severity of flood damage, the output of three machine learning models—Support Vector Machine (SVM), Logistic Regression, and Random Forest—were merged at the provincial scale. Each state was given one of the three extents of severity of damage—Moderate, Severe, or Extreme—based on the proportion of tweets containing damage. states with sparse tweet data were marked as "No data," not indicating the absence of damage, but missing data coverage. **Figures 3** to **5** present the spatial pattern of flood damage intensity predicted by each model with **Figure 3** presenting SVM model output, **Figure 4** presenting Logistic Regression, and **Figure 5** presenting Random Forest predictions.

As a whole, the models indicate overall agreement in predicting which states will be impacted by flood events but considerable variation in forecasting severity. SVM predicts more degrees of damage in some states such as Extreme damage predictions for Uttarakhand, Jammu & Kashmir, and Odisha. Logistic Regression suggests an overall wider spatial extent of Severe or Extreme harm throughout the nation, whereas Random Forest offers more centrally balanced estimates overall apart from Severe damage to be expected in Uttarakhand. These deviations reflect the varying sensitivities and decision boundaries pertaining to each model method when inferring spatial patterns of flood-related indicators from social media data.

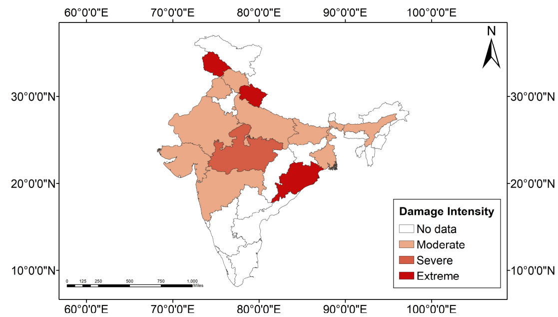

Figure 3. Classified damage map acquired from SVM model

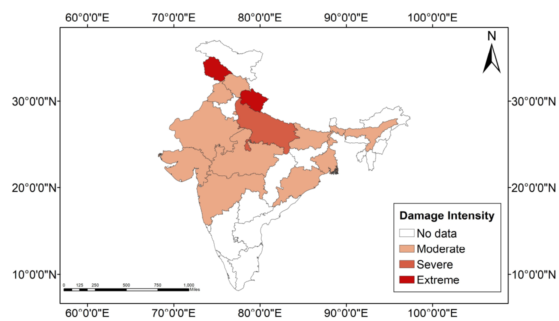

Figure 4. Classified damage map acquired from Logistic Regression

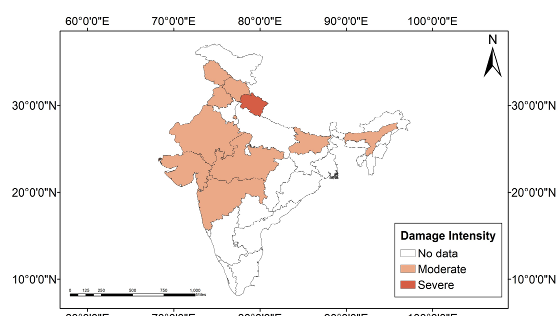

Figure 5. Classified damage map acquired from Random Forest

For each state, the final class was determined as the most frequently predicted class among all tweets assigned to that state. For example, if most tweets from a state were predicted as class

1 (moderate), then the assigned class for that state was "Moderate."

Some entries in the table are marked as "No Damage," but this does not necessarily mean that tweets from that state were completely absent. Rather, it indicates cases in which all tweets were classified as class 0 (non-related to flood damage) by the respective model. Additionally, there were some inconsistencies among the predictions of the models.

The inconsistency observed in the Random Forest model may be attributed to differences in the distribution and characteristics of the sample sets across various damage classes. The model tended to be biased toward detecting non-damage tweets with high precision but relatively lower recall, while the moderate damage class was often overestimated—resulting in higher recall but lower precision. These variations are most likely a consequence of data imbalance and the similarity of textual features among adjacent classes, which makes clear separation more challenging for the model.

Such inconsistencies are reflected in the final damage level predictions across states, where the Random Forest model occasionally assigned different classes than the SVM and Logistic Regression models, particularly for regions with fewer or less distinctive samples, such as Odisha.

State	SVM	Logistic Regression	Random Forest
Assam	Moderate	Moderate	Moderate
Bihar	Moderate	Moderate	Moderate
Delhi	Moderate	Moderate	Moderate
Gujarat	Moderate	Moderate	Moderate
Jammu & Keshmir	Extreme	Extreme	Moderate
Madhya Pradesh	Severe	Moderate	Moderate
Odisha	Extreme	Moderate	NoDamage
Punjab	Moderate	Moderate	Moderate
Uttar pradesh	Moderate	Severe	NoDamage
Uttarakhand	Extreme	Extreme	Severe
West Bengal	Moderate	Moderate	NoDamage

Table 3. Predicted flood damage level in states

Additionally, **Table 3** provides a tabular summary of the predicted damage levels across states for each model, complementing the visual interpretation presented in the provincial damage maps.

3.3 City-level Damage Mapping

The output of the Support Vector Machine (SVM) model was selected for city-level analysis, as it demonstrated the best performance among the three classifiers.

Nine cities across different regions of India were identified and marked on the map. Srinagar in Jammu & Kashmir was the only city categorized as experiencing extreme damage. Moradabad and Dehradun, located in Uttar Pradesh and Uttarakhand respectively, were classified as severely affected. The remaining cities—Guwahati, Muzaffarpur, Delhi, Ujjain, Bargarh, and Kolkata—were all categorized as experiencing moderate damage.

This spatial distribution indicates that flood damage was not concentrated in a single area but rather spread across the northern and eastern parts of India. The presence of higher damage levels in states traditionally known to be at high flood risk—such as Jammu & Kashmir, Uttar Pradesh, and Uttarakhand—supports the reliability and contextual accuracy of the model's predictions. Importantly, the city-level findings reveal localized impacts that may be obscured in broader regional analyses. Such granularity is crucial for effective and targeted disaster response and planning.

To provide contextual clarity, two maps were included. **Figure 6** illustrates the predicted damage intensity at the state level, with dots representing the cities analyzed. **Figure 7** presents a more detailed view, showing the same nine cities with damage intensities visualized within their administrative boundaries. Transitioning from point-based to area-based representations allows for a clearer understanding of intra-city flood impacts.

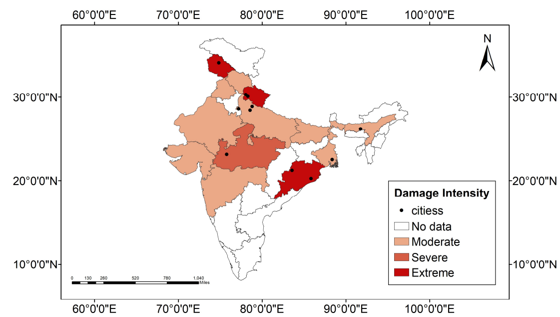

Figure 6. Predicted flood damage intensity at the state level with city markers

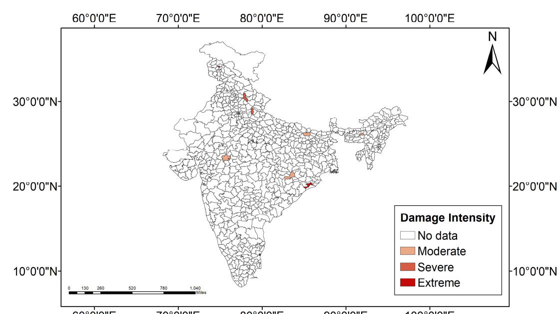

Figure 7. Predicted flood damage intensity at the city level using administrative boundaries

Interestingly, while state-level classifications provide a general sense of flood severity across larger regions, city-level predictions offer more nuanced insights. For instance, although Uttar Pradesh is categorized as having a moderate level of damage overall, the city of Moradabad within it is projected to have experienced severe damage. This highlights the importance of conducting disaster impact assessments at finer spatial resolutions to enhance the precision and effectiveness of response efforts.

Table 4 summarizes the nine cities identified, their corresponding states, and the damage severity levels predicted by the SVM model. These cities span various geographic regions of India, illustrating the widespread nature of the flood impacts.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-4-W8-2025-77-2026

3.1 Validation Using Satellite-Based Damage Assessment

To validate the accuracy of the tweet-based flood damage classifications, external verification was conducted using publicly available satellite-based crisis maps (**Figure 8**). For example, Srinagar, which was classified as highly impacted in the tweet-based model, was also confirmed to have experienced severe flooding according to the Google Crisis Response Map released in September 2014 (Google, 2014).

Moreover, according to the ReliefWeb situation map (India – Floods and Landslides, 25 September 2014) released by the United Nations Office for the Coordination of Humanitarian Affairs (OCHA) (**Figure 9**), the total death toll from the September 2014 flood in the state of Assam was reported to be approximately 36. In the results obtained from the model used in the proposed framework, Assam was classified under the moderate category of flood damage. This category corresponds to areas where the death toll was less than 50.

State	city	Damage intensity
Assam	Guwahati	Moderate
Bihar	Muzaffarpur	Moderate
Delhi	Delhi	Moderate
Jammu & Keshmir	Srinagar	Extreme
Madhya Pradesh	Ujjain	Moderate
Odisha	Bargarh	Moderate
Uttar Pradesh	Moradabad	Severe
Uttar Khand	Dehradun	Severe
West Bengal	Kolkata	Moderate

Table 4. Predicted flood damage in Indian cities

Figure 8. Satellite imagery depicting flood impact in Srinagar, obtained from *Google Crisis Response* (Google, 2014).

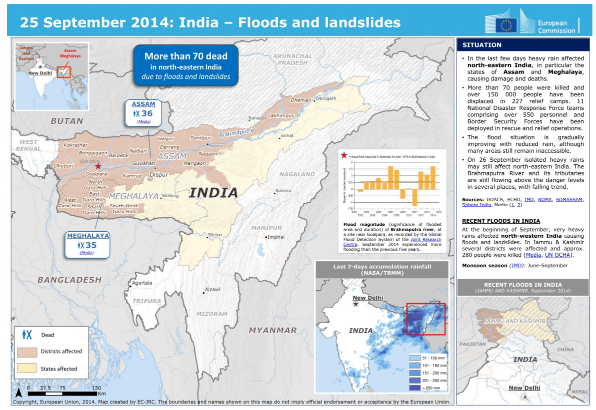

Figure 9. India Floods and Landslides – ReliefWeb Map (September 2014)

The efficiency of the GeoParsing component was qualitatively verified by manually reviewing a subset of tweets to ensure that the extracted place names accurately corresponded to real geographic locations at the city level. Furthermore, the effectiveness of the NLP preprocessing pipeline—comprising text cleaning, tokenization, and TF-IDF feature extraction—was indirectly validated through the performance of the machine learning models during the training and testing processes. The relatively high classification accuracy (above 0.79 for both Logistic Regression and SVM) indicates that the text preprocessing and feature extraction stages successfully generated reliable textual representations.

4. CONCLUSION

This study evaluated the predictive accuracy of multiple machine learning algorithms in assessing the intensity of flood damage based on Twitter data. Among the models tested, the Support Vector Machine (SVM) achieved the highest overall accuracy (0.815), along with notably high precision and recall values for detecting severe damage. While Logistic Regression exhibited lower overall accuracy, it still demonstrated practical utility, consistently identifying high-damage regions such as Odisha and Srinagar, which were corroborated by satellite-derived flood maps.

A key strength of this research lies in its ability to extract both geographic location and damage severity directly from raw tweet texts, enabling fine-grained, city-level assessments of flood impact. This level of detail is particularly valuable for targeted disaster response and planning.

However, the study also faced several limitations. One major constraint was the limited volume of labeled tweets, which restricted the training capacity and potential accuracy of the machine learning models. Additionally, obtaining real-time or historical Twitter data often involves financial costs, which may hinder accessibility for some researchers and institutions.

Although satellite-based flood maps were used for validation purposes, their reliability may be compromised under cloud cover, especially during flood events. To address this, future studies could incorporate radar-based remote sensing data (e.g., Sentinel-1), which is less affected by weather conditions and could enhance validation accuracy. Furthermore, applying the proposed method to geotagged and well-organized Twitter

datasets could provide more precise spatial validation and improve assessment.
Moving forward, the expansion of labeled datasets, the integration of multi-source geospatial data, and the application of advanced deep learning techniques could significantly improve the robustness, scalability, and practical utility of urban-scale disaster impact assessments using social media data.

REFERENCES

Ahadzadeh, S., Malek, M. R., 2021: Earthquake damage assessment based on user generated data in social networks. *Sustainability*, *13*(9), 4814.

Benevolenza, M. A., DeRigne, L., 2019: The impact of climate change and natural disasters on vulnerable populations: A systematic review of literature. Journal of Human Behavior in the Social Environment, 29(2), 266-281.

Feng, Y., Huang, X., Sester, M., 2022: Extraction and analysis of natural disaster-related VGI from social media: review, opportunities and challenges. *International Journal of Geographical Information Science*, *36*(7), 1275-1316.

Google., 2014: *India Floods – Crisis Map (September 2014)*. Google Crisis Response. https://www.google.org/crisisresponse

Imran, M., Castillo, C., Diaz, F., Vieweg, S., 2015: Processing social media messages in mass emergency: A survey. *ACM Computing Surveys (CSUR)*, *47*(4), 1-38.

Imran, M., Mitra, P., Castillo, C., 2016: Twitter as a Lifeline: Humanannotated Twitter Corpora for NLP of Crisis-related Messages.

Kreimer, A., 2001: Social and economic impacts of natural disasters. International Geology Review, 43(5), 401-405.

Lindsay, B. R., 2011, September: *Social media and disasters: Current uses, future options, and policy considerations*.

Madichetty, S., Sridevi, M., 2019: Disaster damage assessment from the tweets using the combination of statistical features and informative words. *Social Network Analysis and Mining*, *9*(1), 42.

Qaiser, S., Ali, R., 2018: Text mining: use of TF-IDF to examine the relevance of words to documents. *International journal of computer applications*, *181*(1), 25-29.

Ritchie, H., Rosado, P., Roser, M., 2022: Natural disasters. Our World in Data.

ReliefWeb., 2014, September 25: *India – Floods and landslides (25 September 2014)* [Map]. United Nations Office for the Coordination of Humanitarian Affairs (OCHA). https://reliefweb.int/map/india/25-september-2014-india-floods-and-landslides

Samuel, E., 2024: The Effectiveness of Disaster Management Policies in Reducing the Impact of Natural Disasters in Canada. *Journal of Public Policy and Administration*, *9*(2), 26-38.

Sutton, J. N., Palen, L., Shklovski, I., 2008: Backchannels on the front lines: Emergency uses of social media in the 2007 Southern California Wildfires.

Vahidnia, M. H., 2024: Meta ensemble learning in geospatial sentiment analysis and community survey mapping: a water supply case study. *Earth Science Informatics*, *17*(4), 3233-3252.

Zou, L., He, Z., Wang, X., Liang, Y., 2025: Spatiotemporal Typhoon Damage Assessment: A Multi-Task Learning Method for Location Extraction and Damage Identification from Social Media Texts. *ISPRS International Journal of Geo-Information*, *14*(5), 189.