International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11s
DOI: https://doi.org/10.17762/ijritcc.v11i11s.8174
***Article Received:** 28 June 2023 **Revised:** 26 August 2023 **Accepted:** 07 September 2023*

An Ensemble Learning Approach for Fast Disaster Response using Social Media Analytics

M. Ayyadurai[1], K. Sujatha[2], R. Pavithra Guru[3], D. Sasirekha[4], A. Umamageswari[5], S. Deepa[6]
[1]Assistant Professor, Department of Computer Science and Engineering,
SRM Institute of Science and Technology, Ramapuram Campus, Chennai
ayyadurm@srmist.edu.in
[2]Assistant Professor, Department of Computer Science and Engineering,
SRM Institute of Science and Technology, Ramapuram Campus, Chennai
sujathak@srmist.edu.in
[3]Assistant Professor, Department of Computer Science and Engineering,
SRM Institute of Science and Technology, Ramapuram Campus, Chennai
pavithrr5@srmist.edu.in
[4]Assistant Professor, Department of Computer Science and Engineering,
SRM Institute of Science and Technology, Ramapuram Campus, Chennai
sasirekd1@srmist.edu.in
[5]Assistant Professor, Department of Computer Science and Engineering,
SRM Institute of Science and Technology, Ramapuram Campus, Chennai
r.umaramesh@gmail.com
[6]Associate Professor, Department of Computer Science and Engineering,
SRM Institute of Science and Technology, Ramapuram Campus, Chennai
deepas1@srmist.edu.in

Abstract— Natural disaster happens, as a result of natural hazards that cause financial, environmental or human losses. Natural disasters strike unexpectedly, affecting the lives of tens of thousands of people. During the flood, social media sites were also heavily used to disseminate information about flooded areas, rescue agencies, food and relief centres. This work proposes an ensemble learning strategy for combining and analysing social media data in order to close the gap and progress in catastrophic situation. To enable scalability and broad accessibility of the dynamic streaming of multimodal data namely text, image, audio and video, this work is designed around social media data. A fusion technique was employed at the decision level, based on a database of 15 characteristics for more than 300 disasters around the world (Trained with MNIST dataset 60000 training images and 10000 testing images). This work allows the collected multimodal social media data to share a common semantic space, making individual variable prediction easier. Each merged numerical vector(tensors) of text and audio is sent into the K-CNN algorithm, which is an unsupervised learning algorithm (K-CNN), and the image and video data is given to a deep learning based Progressive Neural Artificial Search (PNAS). The trained data acts as a predictor for future incidents, allowing for the estimation of total deaths, total individuals impacted, and total damage, as well as specific suggestions for food, shelter and housing inspections. To make such a prediction, the trained model is presented a satellite image from before the accident as well as the geographic and demographic conditions, which is expected to result in a prediction accuracy of more than 85%.

Keywords- Ensemble Algorithm, Disaster management, K-CNN, PNAS

I. INTRODUCTION & RELATED WORK

India's experience with disaster management over the past twenty years has prompted a reconsideration of the issues with disaster mitigation and the development of a better structure to enable the efficient and effective performance of disaster relief operations. However, when actual examples are looked at, it becomes clear that only concentrating on disaster response and recovery phases (specifically disaster crisis management) and being unable to give mitigation and preparedness phases adequate care and attention.

First and foremost, this study defines critical terminology for an effective crisis management system and gathers real-time social media samples to inform the rescue department's reaction. Then it will provide suggestions for disaster management through advanced technologies like Machine learning and deep learning algorithms to provide solutions for the problems identified in the current system. In the final section, which is the conclusion, a successful disaster management model that follows an Indian example is tried to be presented.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11s
DOI: https://doi.org/10.17762/ijritcc.v11i11s.8174
***Article Received:** 28 June 2023 **Revised:** 26 August 2023 **Accepted:** 07 September 2023*

Figure 1: Social media Images of Various Disaster

Unmanned aerial vehicles (UAVs) are helping to assess the region impacted by the tragedy and to build a network of communication between the rescue workers and the available cellular infrastructure [1].Using remote sensing technologies, such as Synthetic Aperture Radar (SAR) and VIEWS, to calculate the impact of the central Java earthquake and tsunami [2].

The use of naturalistic walking massive data from video recordings to study pedestrian navigation in enclosed open spaces By taking into consideration realistic pedestrian behaviour, simulation approaches based on such augmented models can assist practitioners, such as public building designers, in maximising designs for open places [3]. Drones were utilized to find missing individuals and monitor the terrain during the Uttarakhand floods of 2013, providing authorities with vital updated information. IIT Madras students recently built an AI-enabled drone that can assist authorities in providing important information on persons stuck in disaster-stricken areas.

Google's experimental effort to assess the flood level condition in Patna, which was done in partnership with the Central Water Commission [4][13]. Google claims that a map-based notice was sent to anyone living within a thousand square kilometers of Patna using its AI-based model. The map showed which areas were more likely to be flooded and which ones were less likely to be flooded [5]. The accuracy of Google's model, according to the company, was over 90%.

The Kerala State IT Mission had a brainstorming session and decided to design a crisis management system using the crowd sourcing method and deploy it as quickly as possible using the DevOps methodology. At the URL www.keralarescue.in, a crisis management platform was designed, hosted, and made available to the public. Within 12 hours of the first day of inundation, it was up and operating. This was made possible by the contributions of techies from all over the world who offered to write, review, and recode to improve the platform's capabilities in response to the state's needs. We extracted and categorised tweets created during a natural disaster based on factors like whether the tweet included caution and advice or if it provided information about the casualty and damage along with the information source [6][14] in order to improve our assessment and knowledge of a disaster situation. Researchers used CNN to examine a sample of tweets sent out after Hurricane Sandy in 2012. This analysis was improved upon by the addition of other categorization categories, such as infrastructure and resource [7].

If analyzed quickly, this online data can be beneficial to humanitarian groups [8]. The Qatar Computing Research Institute (QCRI) has created a number of Artificial Intelligence (AI) technologies to improve disaster response and management. These technologies include technologies that use machine learning to automatically classify text communications (like tweets) into humanitarian categories including need reports, injured or deceased folks, and picture processing. The systems collect and analyze data from social media in real time to assist humanitarian organizations in gaining situational, researchers worked on RNN [9]. The LSTM architecture is a feedback-connected artificial recurrent neural network (RNN). As a result, it can handle extended sequences of data values, such as videos. The data movement during algorithm execution is controlled by three gates in LSTMs: input gate, forget gate, and output gate. Other RNNs have a vanishing gradient problem, which LSTMs are built to address and solve [17].

The Data collection is a crucial part of any disaster management because the affected region people will be in apprehensive and they may not get proper channel to reach the current situation immediately. The reason behind delayed response from the Government or different rescue department is still lack of communication to help most affected people immediately. The main uniqueness of our proposed methodology is bringing notice of highly affected region and informing such complicated area to the Government and Non-government organization [10][15]. Our proposed methodology will provide a facility for the affected people to send the data from affected people in the form of different categories such as text, image, audio and video.

The collected information from the affected people will be forwarded to our system to predict the emergency help required region and send those details to different rescue team. Many existing solutions are restricted with single modal and multimodal with single algorithms for prediction. The proposed methodology will analyze the text and audio information received from them is converted as tensors and analyzed using one model. Videos and images are analyzed using different suitable model to provide better prediction. Multi modal based solution using suitable model according to the information type is a novel one [11][16]. The entire application will be in single framework from the collection of data with the help of social media or chatbot analyzing the data using suitable model and the sending the relevant data to the rescue team [12]. The methodology will provide efficient solution since all modules are connected under single framework even though different sub process involved in it.follow.

II. PROPOSED METHOD

To achieve the objective of providing a quick response to any sort of disaster that's happening in the country, we will have to go through sequence of methods to extend help to the needy and the people at risk. The steps comprise of ensemble learning processes namely Training, Testing, and Validation. The first, called training is concerned with the data utilized as training data. The data processing method is crucial in this case. The testing step focuses architecture of ensemble learning approach that learns the training data. The third step, validation, was all about verifying the results. Figure 2 shows the entire architecture to be followed in the proposed project work. The efficiency of the methodology revolves around the proposed dataset collection process.

The data sources are of three types:

- Disaster Characteristics
- Context aware risk information
- Post disaster condition.

2.1. Disaster Characteristics:

Seven factors are taken into account to define each natural catastrophe: the location, the time, the kind of disaster, the scale of the disaster, the total number of fatalities, the total number of impacted persons, and the total damages. (https://www.emdat.be/).

2.2. Context Aware Risk Information:

Information on risk management, elevation, different climates, natural resources, the employment-to-population ratio, the value of construction, the amount of total renewable power, the population, and the areas of the three closest administrative divisions (https://data.un.org/).

2.3. Post disaster condition:

These factors, such as the state of the roads, rivers, or land usage, influence judgments on where to live and how to live, and they are generated from satellite images using Google Maps.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11s
DOI: https://doi.org/10.17762/ijritcc.v11i11s.8174
***Article Received:** 28 June 2023 **Revised:** 26 August 2023 **Accepted:** 07 September 2023*

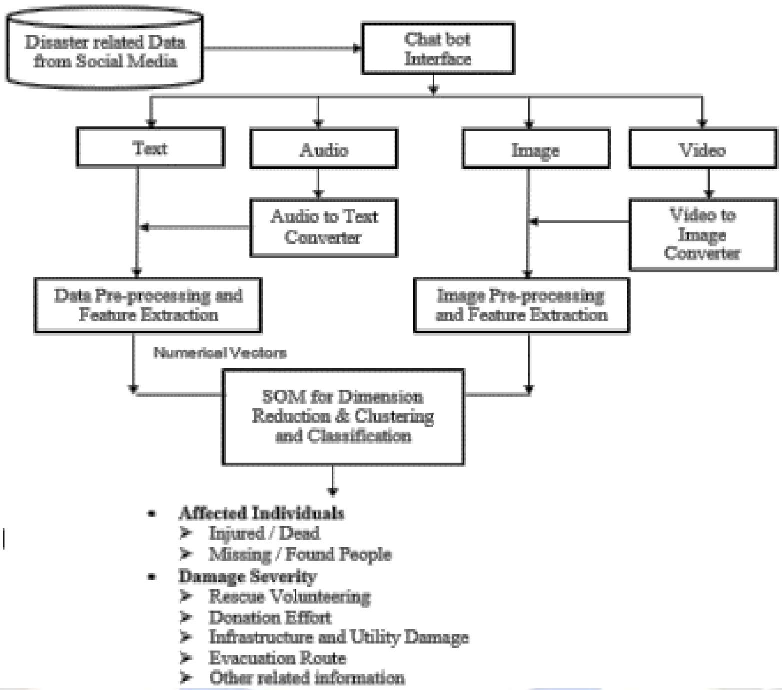

Figure 2: Architecture Diagram of Proposed Method

2.4. Step-by Step Process:

Step 1: Training:

To remove outliers and extract numerical values, each data modality is pre-processed. The one-hot encoding method was used to convert alphanumeric and alphabetical data into numerical values. This method creates a dictionary of all the words in the data and assigns each one a number that encodes it as an integer. An unsupervised extracting features approach was utilised to translate each image in order to extract numerical information from satellite photographs. This procedure is unsupervised since the satellite photos are presented as a collection of unlabeled instances. We initially constructed an AI feature extraction VGG-16 network that can train using satellite images using the densely integrated convolution layer model in order to extract features. It extract features from the entire dataset with the necessary dimensions after training, resulting in a list of numerical vectors to represent natural disasters.

Step 2: Testing

To study the connections between the current tragedies, this work suggested an unsupervised Self-organizing map (SOM) computer vision clustering approach. A comparison to maps, which are often used tools for decision support in disaster management, is also provided by SOM. On maps, proximity denotes similarity. The input layer of the SOM is composed of the 19-dimensional vectors acquired in the preceding phases. Each natural catastrophe is positioned according to how similar its weight value is to that of its neighbours in the output layer of the SOM, which is depicted as a similarity landscape. The average of a original n-dimensional observational data that is associated with each SOM or Best Matching Unit (BMU) node after iteration.

Step 3: Validation

This essay focused on the natural calamities that will occur in 2020. (Floods, earthquakes, floods, and landslides). To verify a forecast that wasn't based on the data used to make it? It selected a few natural catastrophes in the year 2020 to serve as Validation Data. Each natural catastrophe in the VD will be encoded, normalised, and fused, as was stated in the preceding sections. The values to be anticipated are the number of fatalities, the number of persons impacted, and the economic losses. Because the objective of this stage is to give decision-makers a severity rating of the catastrophe so they can determine the necessary resources, this study uses the EM-DAT categorization. As a result, the forecast will give a range of effect, such as low, medium, and high, rather than an exact number. Table 1 name the categories based on the number of deaths and value of economic damages from low to high (Low, medium, High).

ISSN: 2321-8169 Volume: 11 Issue: 11s
DOI: https://doi.org/10.17762/ijritcc.v11i11s.8174
***Article Received:** 28 June 2023 **Revised:** 26 August 2023 **Accepted:** 07 September 2023*

TABLE 1. CATEGORIES OF PREDICTIONS

Type	Number of fatalities	Economic Damage (US$)
High	>3,000	> 5 Billion US$
Medium	>500 to <3,000	>3 to 5 Billion US$
Low	>1 to <500	<3 Billion US$

III. EVALUATION METRICS AND RESULT ANALYSIS

Decision-makers can influence the results of later disaster response operations by making early decisions based on predictions of the impact of a natural catastrophe on a community in terms of the total number of deaths, total number of person impacted, and total damages. These facts (anticipated impact) often help with situational awareness. For decision-making, certain response clusters, like the shelter and housing sector, need for certain data trained using the MNIST dataset, which includes social media data for real-time analysis and 60 000 training pictures and 10,000 test images.

TABLE 2: CATEGORIES OF MAJOR NATURAL DISASTERS IN INDIA FOR DEATH RATE

Location	Type	Deaths	Category
West Bengal (2020)	Hurricane	103	Low
Kerala (2018)	Flood	504	Medium
Chennai (2015)	Flood	289	Low
Andhra (2014)	Strom	68	Low
Jammu and Kashmir (2014)	Flood	665	Medium
Uttaranchal (2013)	Flood	5748	High
Gujarat (2006)	Flood	350	Low
Gujarat (2001)	Earth quake	19737	High

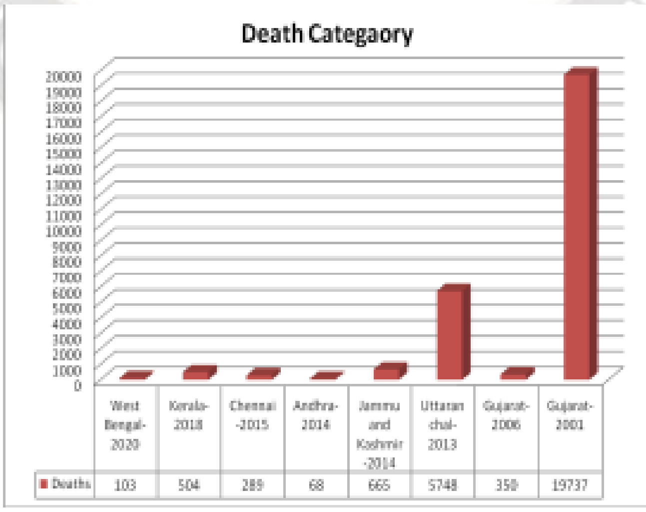

Figure 3: Death Category Risk Analysis

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11s
DOI: https://doi.org/10.17762/ijritcc.v11i11s.8174
***Article Received:** 28 June 2023 **Revised:** 26 August 2023 **Accepted:** 07 September 2023*

Table 2 and Figure 3 shows the category analysis of death rate happened in India for last 20 years during major natural disasters. This data provides clear visibility to the experts those who involved in the rescue process during disasters. The proposed method provides the predictions according the categories.

TABLE 3: CATEGORIES OF MAJOR NATURAL DISASTERS IN INDIA FOR ECONOMICAL LOSS

Location	Type	Economical Loss (US$ in Billion)	Category
West Bengal (2020)	Hurricane	13.5	High
Kerala (2018)	Flood	3.52	Medium
Chennai (2015)	Flood	2.37	Low
Andhra (2014)	Strom	7.56	High
Jammu and Kashmir (2014)	Flood	6.45	High
Uttaranchal (2013)	Flood	1.62	Low
Gujarat (2006)	Flood	4.3	High
Gujarat (2001)	Earthquake	6.13	High

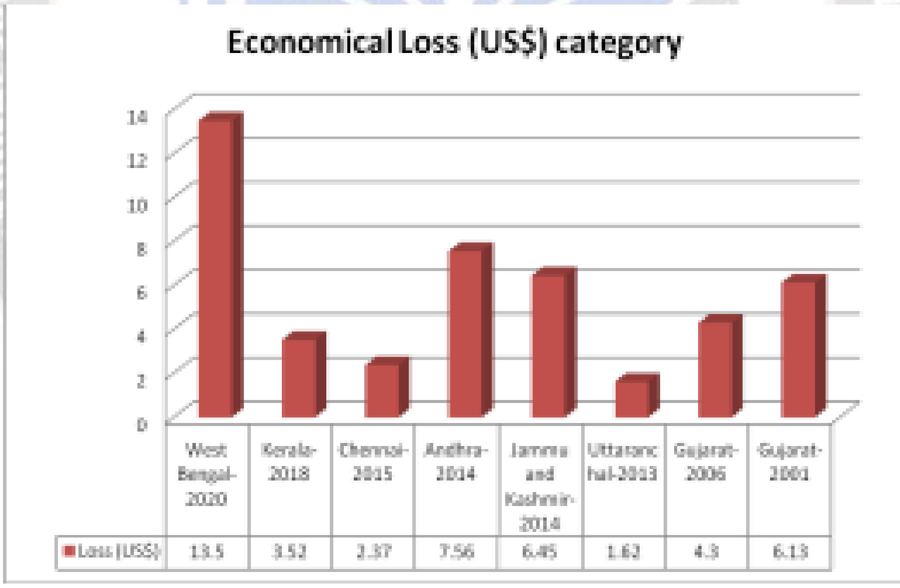

Figure 4: Economic loss Category Risk Analysis

Table 3 and Figure 3 shows the category analysis of economic loss happened in India for last 20 years during major natural disasters. This data provides clear visibility to the experts to manage the economy during disasters. The proposed method provides the predictions according the categories.

IV. CONCLUSION

National disaster management authority reinforced to fully integrate and implement hazard mitigation policies and plans by 2025, increasing the capacity of national actors and communities to implement effective risk reduction and response measures. This proposed work used an ensemble learning strategy for combining and analyzing social media data in order to close the gap and progress in catastrophic situation. This work allows the collected multimodal social media data to share a common semantic space, making individual variable prediction easier. K-CNN algorithm and Progressive Neural Artificial Search (PNAS). The proposed work predicted the accident as well as

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11s
DOI: https://doi.org/10.17762/ijritcc.v11i11s.8174
Article Received: 28 June 2023 Revised: 26 August 2023 Accepted: 07 September 2023

the geographic and demographic conditions, which is expected to result in a prediction accuracy of more than 85% .Authors and Affiliations

REFERENCES

[1] ECHO (2020). Shelter and settlements. Factsheet. European Civil Protection and Humanitarian Aid, Brussels

[2] Oxfam (2016). The Effectiveness and Efficiency of Interventions Supporting Shelter Self-Recovery Following Humanitarian Crises: An evidence synthesis protocol. Humanitarian Evidence Program, London. 10.21201/2016.605179

[3] Marmanis, D., Wegner, J. D., Galliani, S., Schindler, K., Datcu, M., & Stilla, U. (2016). Semantic segmentation of aerial images with an ensemble of CNSS. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2016, 3, 473-480.

[4] Warnier, M., Alkema, V., Comes, T., & Van de Walle, B. (2020). Humanitarian access, interrupted: dynamic near real-time network analytics and mapping for reaching communities in disaster-affected countries. OR Spectrum, 42(3), 815-834.

[5] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[6] Chen, T., Lu, D., Kan, M.-Y., and Cui, P. (2013). "Understanding and classifying image tweets". In: ACM International Conference on Multimedia, pp. 781–784.

[7] Prof. Barry Wiling. (2018). Identification of Mouth Cancer laceration Using Machine Learning Approach. International Journal of New Practices in Management and Engineering, 7(03), 01 - 07. https://doi.org/10.17762/ijnpme.v7i03.66.

[8] Socher, R., & Fei-Fei, L. (2010). Connecting modalities: Semi-supervised segmentation and annotation of images using unaligned text corpora. In 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition

[9] Chowdhury, S. A., Stepanov, E. A., Danieli, M., & Riccardi, G. (2019). Automatic classification of speech overlaps: Feature representation and algorithms. Computer Speech & Language, 55, 145-167.

[10] Poria, S., Cambria, E., Howard, N., Huang, G. B., & Hussain, A. (2016). Fusing audio, visual and textual clues for sentiment analysis from multimodal content. Neuro computing, 174, 50-59.

[11] Vande Berg, L. R., Wenner, L. A., & Gronbeck, B. E. (2004). Media literacy and television criticism: Enabling an informed and engaged citizenry. American Behavioral Scientist, 48(2), 219-228

[12] Makarand L, M. . (2021). Earlier Detection of Gastric Cancer Using Augmented Deep Learning Techniques in Big Data with Medical Iot (Miot). Research Journal of Computer Systems and Engineering, 2(2), 22:26. Retrieved from https://technicaljournals.org/RJCSE/index.php/journal/article/view/28

[13] A. Umamageswari, S. Deepa, K. Raja, "An enhanced approach for leaf disease identification andclassification using deep learning techniques", Measurement: Sensors, Volume 24, 2022, 100568, ISSN2665-9174

[14] Alvarez-Marin, D., & Ochoa, K. S. (2020). Indexical Cities: Articulating Personal Models of Urban Preference with Geotagged Data. arXiv preprint arXiv:2001.10615.

[15] Ofli, F., Alam, F., & Imran, M. (2020). Analysis of social media data using multimodal deep learning for disaster response. arXiv preprint arXiv:2004.11838.

[16] Pereira, M., Pádua, F., Pereira, A., Benevenuto, F., & Dalip, D. (2016, March). Fusing audio, textual, and visual features for sentiment analysis of news videos. In Proceedings of the International AAAI Conference on Web and Social Media

[17] Umamageswari., Johnson, S.D., Sara, D., Kothandaraman, R. (2022). An enhanced identification andclassification algorithm for plant leaf diseases based on deep learning. Traitement du Signal, Vol. 39, No.3, pp. 1013-1018. https://doi.org/10.18280/ts.390328

[18] Azlina Abdullah, Ismail Musirin, Muhammad Murtadha Othman, Siti Rafidah Abdul Rahim, A.V. Sentilkumar. (2023). Multi-DGPV Planning Using Artificial Intelligence. International Journal of Intelligent Systems and Applications in Engineering, 11(4s), 377–391. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/2677.

[19] Murtagh, F., & Hernández-Pajares, M. (1995). The Kohonen self-organizing map method: an assessment. Journal of Classification, 12(2)

[20] Umamageswari A, Bharathiraja N, Irene DS. A Novel Fuzzy C-Means based Chameleon Swarm Algorithm for Segmentation and Progressive Neural Architecture Search for Plant Disease Classification.