AI Driven Natural Hazard Prediction System

Vishvesh Mukesh
Department of AIT-CSE
Chandigarh University
Punjab, India
vishveshmukesh04@gmail.com

Gautam Menon
Department of AIT-CSE
Chandigarh University
Punjab, India
officially.gm11@gmail.com

Hrithik Joshi
Department of AIT-CSE
Chandigarh University
Punjab, India
hrithikjoshi03@gmail.com

Sparsh Saraf
Department of AIT-CSE
Chandigarh University
Punjab, India
sparshsfofficial@gmail.com

Jyotsna Chauhan
Department of AIT-CSE
Chandigarh University
Punjab, India
Jyotsnachauhan1113@gmail.com

Dr. Raghav Mehra
Department of AIT-CSE
Chandigarh University, Gharuan,
Mohali1-140413,
Punjab, India
Raghav.mehrain@gmail.com

***Abstract*— Natural hazards are dangerous and can cause a lot of harm to human life and property. Given this, it is essential to have methods that are advanced enough to predict and get information in time for a quicker response. The paper that is the subject of this report, gives an explanation of a machine learning system, trained on historical disaster data through an AI algorithm, which was used to determine the disaster type as an example of a natural hazard. The data set consists of characteristics like dead people, economic losses, and the geographical location of the region where the disaster occurred, with disaster type as the last field of information.**

The models such as Decision Tree, K-Nearest Neighbours (KNN), and Support Vector Machine (SVM) were implemented and their performance was evaluated. The decision tree model turned out to be the correct model with 84.31% precision, followed by SVM with 82.35% and KNN with 79.41%. Confusion matrices and classification reports were the tools used for this analysis, and graphical error analysis was the next tool employed in the prediction patterns analysis. The decision tree model was also proven to be more transparent thus aiding its real-world applications in disaster management systems.

The research puts forward that AI could be the way of the future for hazard prediction and also proposes subsequent improvements in the area like updating records in real-time, using synthetic data to balance classes, and the ensemble-based model to bolster the immune system of the system.

Keywords— Artificial Intelligence (AI), Machine Learning (ML), Deep Learning (DL), Natural Hazard Prediction, Disaster Forecasting, Early Warning Systems, Remote Sensing, Geospatial Data Analysis, Internet of Things (IoT), Big Data in Disaster Management, Predictive Analytics, Risk Assessment, Climate Change and AI, Neural Networks in Hazard Prediction, Earthquake Prediction Models, Flood Forecasting with AI, Wildfire Detection Systems, Hurricane and Cyclone Prediction, Resilience and Disaster Mitigation, AI-driven Decision Support Systems.

I. Introduction

Natural hazards, including earthquakes, floods, wildfires, and hurricanes, pose significant threats to human life, infrastructure, and economic stability. Traditional methods of disaster prediction rely on historical data, physical modeling, and sensor networks, which often lack the speed and accuracy required for timely warnings. However, with the rapid advancements in Artificial Intelligence (AI) and Machine Learning (ML), scientists and disaster management agencies are now leveraging data-driven approaches to enhance predictive capabilities. AI-driven hazard prediction systems analyze vast amounts of geospatial, climatic, and seismic data to identify patterns, forecast disasters, and provide early warnings with unprecedented precision. These systems integrate satellite imagery, IoT sensors, social media feeds, and historical disaster records to train deep learning models, enabling real-time risk assessment and mitigation strategies.

The growing frequency and intensity of natural disasters due to climate change further emphasize the need for AI-powered solutions. For instance, flood prediction models now use convolutional neural networks (CNNs) to process satellite data and predict water levels, while earthquake early warning (EEW) systems employ recurrent neural networks (RNNs) to detect seismic precursors. Similarly, wildfire spread prediction relies on reinforcement learning to simulate fire behaviour under varying weather conditions. Governments and organizations worldwide—such as NASA, Google's Flood Forecasting Initiative, and the USGS ShakeAlert system—have already deployed AI-based tools to minimize disaster impact.

AI's predictive power for natural disasters is crucial, but challenges such as data privacy and algorithmic bias remain. Collaboration among geologists, data analysts, and policymakers is needed to integrate AI into disaster response systems. This research paper examines AI's potential in disaster preparation, response, and saving lives, revealing both possibilities and obstacles in creating resilient, proactive, and reactive societies.

II. Background to AI driven Hazard prediction systems

For many years natural calamities have been a big potential threat to humans and buildings, and that has always been the case, while traditional methods of forecast have been statistical models and physical simulation that often are not accurate and do not offer real-time predictions. Hurricane or earthquake intensity is increasing as a result of global warming. The situation is so bad that there is an urgent need for a solution which has not been seen so much, and the natural outcome of this is the use of artificial intelligence (AI). AI-based tools resort to the techniques of machine learning and deep learning to sort through a conclave of information, which includes data derived from satellites, seismic sensors, weather stations, as well as social media, for an all-inclusive and much quicker disaster forecasting. AI can struggle with real-time high-dimensional data, provoke lightning-speed data processing, and help the system catch routines that are invisible to the eyes of people or traditional algorithms; these are the reasons link to the technique becoming unlike the old ones.

Disaster research had a massive transformation due to AI that was capable of using CNNs and LSTMs. Out of the blue, AI has been the solution to predicting the societal system changes, and predicting the trend and even a lot of similar tasks. A CNN-based flood mapping model can now predict using land cover type. Their predictive ability comes from CNNs, while LSTMs access seismic data and compare the data to previous data points.

A. AI Technologies in Hazard Prediction

Several AI techniques contribute to hazard prediction, each tailored to different disaster types and data sources. Some of the most commonly used AI-driven methods include:

- ***Machine Learning (ML) and Deep Learning (DL):*** Machine learning algorithms, which include both supervised and unsupervised learning models, analyze large datasets to identify patterns and connections. Supervised machine learning (ML) approaches such as decision trees, support vector machines (SVMs), and random forests are trained on labeled historical data to categorize risk levels and forecast hazard events. Unsupervised machine learning, such as clustering algorithms, aids in the detection of anomalies in seismic activity, atmospheric pressure, and oceanic currents that could suggest an approaching calamity.

  Deep learning is a subset of machine learning that analyzes complicated datasets using artificial neural networks (ANNs). CNNs handle spatial data from satellite images to detect flood-prone areas and wildfire outbreaks, whereas RNNs and LSTM networks examine sequential data to predict storms and earthquakes.

- ***Big Data Analytics and Cloud Computing:*** The integration of AI, big data analytics, and cloud computing improves hazard prediction skills by allowing for real-time data processing and scalability. Cloud-based platforms collect data from a variety of sources, such as weather stations, social media feeds, and emergency reports, while AI-powered algorithms extract valuable information. This method improves catastrophe preparedness and the effectiveness of emergency response efforts.

B. AI-Powered Systems for Natural Hazard Prediction

- **Earthquake Prediction Systems**: Because seismic activity is complicated and unpredictable, earthquake prediction remains one of the most difficult aspects of hazard forecasting. AI models use seismic patterns, subsurface stress accumulation, and microseismic signals to identify early warning indicators. For example:
  1. **Japan's AI-based Earthquake Early Warning System** utilizes neural networks to analyze seismic waves and issue alerts within seconds of initial tremors.
  2. **Google's Android Earthquake Alerts System** uses smartphone accelerometers as mini seismometers, collecting data globally to enhance earthquake prediction models.
- **Flood Prediction Systems:** Floods cause extensive damage to infrastructure and human settlements, necessitating accurate forecasts and early warning systems. Artificial intelligence-based flood prediction systems estimate flood hazards using hydrological models, meteorological data, and satellite imagery. Notable systems include:
  1. **IBM's AI Flood Forecasting System**, which integrates ML models with weather data and topographical analysis to predict flood levels days in advance.
  2. **Google Flood Forecasting Initiative**, which employs deep learning models trained on historical and real-time hydrological data to provide accurate flood warnings in vulnerable regions.
- **Wildfire Detection and Prediction Systems:** Climate change has led to an increase in the frequency of wildfires, necessitating proactive monitoring and response systems. AI-powered wildfire prediction systems use satellite photos, weather data, and vegetation patterns to pinpoint fire-prone locations. Examples include:
  1. **NASA's FIRMS (Fire Information for Resource Management System)**, which uses AI-enhanced remote sensing to detect and track wildfires in near real-time.
  2. **California's AI Wildfire Prediction Model**, which processes geospatial data to forecast wildfire outbreaks and optimize firefighting strategies.

AI-powered natural hazard prediction systems have transformed disaster forecasting by increasing accuracy, lowering false alarms, and allowing for real-time decision-making. These systems use a variety of data sources, including satellite images, sensor networks, and meteorological data, to identify risk indicators and provide early warnings. AI breakthroughs such as deep learning, geospatial analytics, and big data integration have dramatically improved predictive skills across a wide range of danger kinds.

Despite their success, AI-based hazard prediction systems confront a number of hurdles, including data quality issues, computing constraints, and ethical concerns about data privacy and access. Furthermore, assuring the interpretability and dependability of AI models is an important topic of research. Moving forward, interdisciplinary collaboration, increased investment in AI-driven disaster resilience, and the integration of AI with community-based disaster preparedness initiatives will be critical in mitigating the catastrophic effects of natural disasters.

This article delves deeper into these AI-driven developments, examining their techniques, uses, and future prospects. Understanding the potential of AI in hazard prediction allows us to build more effective solutions to protect people and infrastructure from natural disasters.

III. Literature Survey

In poor countries like India, the use of new technologies such as Artificial Intelligence (AI), Satellite-based systems (GIS), Sensor Networks (WSNs), Drones (UAVs), and Data analytics is completely changing the disaster management sector by reducing infrastructural vulnerabilities via the latest responsive methodologies [1][17]. WSNs bring down the monitoring costs by 70%, UAVs decrease the response times by 25%, and the CNN-based models boost the prediction accuracy of an earthquake by 23% [17][18]. Using data from satellites, sensors, and social media, AI can help in the improvement of resource acts, including enhanced coordination of humanitarian efforts [2]. Big Data Technology not only leads to enhanced predictive capabilities but also raises the accuracy levels in the identification of floods by 87% and landlines by 79% [8].

Explainable AI (XAI), Federated Learning, and Large Language Models (LLMs) have the power to increase the effectiveness of the prediction process and reduce the communication time for the operations thus XAI is expected to improve the efficiency by 50% [18] and LLMs, being used for the issue of crisis communication, show better results than conventional methods [4]. High-resolution satellite data (<1 m) and Deep Convolutional Neural Networks (DCNNs) reach land use detection higher than 98% accurate whilst models like HIM-STRAT reach snow cover simulation at a level of 86.5% accuracy [6][19]. In the Middle Ganga Plain area prone to flood, the application of ANFIS-GA has yielded out with 0.924 prediction rate and the validation accuracy has been 0.883 [20].

It should be mentioned that a number of complex problems such as data isolation, poor infrastructure, and ethical issues have arisen [1]. Using a technique such as AHP, a multi-criteria strategy can be developed which highlights that nearly 30% of India's land and nearly 40% of its population are at a high risk of a disaster although the local level of accuracy should be improved [5]. In order to gain support for the cause of disaster resilience worldwide, promotion of the metier, and the use of AI which is widely applicable are key elements [1][3][9]..

Big Data analytics have greatly improved the detection of natural hazards, which has led to the flood prediction accuracy to reach 87%, landslide detection to 79%, and efficiency of early warning to be 23% compared to traditional methods [8]. Although, the major predicament for reliable prediction lies in data source diversity incorporation. This is because the risk assessment of the world is moving from single-hazard systems to multi-hazard ones that are a result of climate change, which, for instance, is the floods increase of the 30% by 2050. The present situation recommends that a standardized, easily extended risk assessment system is required. It is stated that at the moment there are 1.5 billion people living in high-risk zones, with annual disaster-related losses costing as much as $250 billion, which is a clear signal that the need for standardized, scalable risk assessment tools is there [9].

It can be observed that AI is bringing about the great change in the catastrophe risk assessment, the response to of infrastructure, and the resilience of the environment. With xBD data and the high-resolution of satellite images (<1 m), artificial intelligence is efficiently and with great accuracy, using deep learning to assess the damage caused by disasters on buildings; nevertheless, generalization still waits for a solution [10]. Indeed, in the case of thunderstorm forecasting, AI is able to gather information from radar, lightning, as well as the instruments of weather detecting in order to provide real-time hazard possibilities based on the 1 km grid that are further confirmed by explainable AI for their reliability. Thus, AI will help improve the weather predictions.

For the things mentioned above, it is clear that the Machine Learning (ML) algorithm for the soil liquefaction data analysis that was accessed in the smart cities with over 96% accuracy was based on the use of such models as gradient boosting and neural networks that greatly made it possible to create a high-resolution and accurate risk map for the area that was planned. In a broader sense, with the help of AI, not only the environment-sound recovery and health matters are prone to their perfect solutions, even predicting the disasters through AI, diagnosing the mental health issues, and giving the recovery aid are the things now can be done easily. There is a statistic that in 2021, AI managed to respond to 432 natural disasters that the impact caused, in the end, 101.8 million people and a loss of $252.1 billion [13].

AI has also helped in infrastructure vulnerability assessment, particularly in determining the susceptibility of data centers to natural disasters and power outages. A statewide survey of 2,660 data centers in the United States identified tornadoes, hurricanes, and earthquakes as primary risks, while spatial analytic approaches revealed low-risk zones, giving decision-makers with practical information to improve resilience [14]. These findings highlight AI's growing importance in disaster management by boosting prediction accuracy, streamlining emergency response, and strengthening infrastructure resilience. Future research should concentrate on improving AI models for different environments, combining interdisciplinary techniques, and resolving ethical concerns in order to optimize AI's impact on disaster preparedness.

Artificial intelligence (AI) is becoming increasingly important in natural disaster prediction, management, and environmental monitoring. In landslide-prone areas such as the Garhwal Himalaya, Support Vector Machines (SVMs) outperform classic statistical models. The L2-SVM-MFN model, with an AUC value of 0.829, efficiently maps landslide susceptibility, facilitating risk assessment in data-scarce areas [15]. A comprehensive analysis of AI applications in disaster management identifies six types of AI-powered models and discusses their impact on early warning systems. AI-enhanced EWSs increase predicting accuracy by 30-40% while reducing catastrophe reaction times by 20%. However, obstacles with data availability and ethical considerations continue, with 60% of research identifying data quality issues and 25% highlighting ethical concerns [16].

Emerging technologies such as Wireless Sensor Networks (WSNs), Unmanned Aerial Vehicles (UAVs), and the Internet of Things (IoT) are transforming multi-hazard disaster response—WSNs cut monitoring costs by 70%, while UAVs improve situational awareness and reduce response times by 25%. Over 150 countries use satellite-based monitoring for efficient disaster mitigation [17]. AI-driven systems, including CNN-based models, have enhanced earthquake prediction accuracy by 23%, though 40% of failures are due to outdated or missing datasets. The adoption of Explainable AI (XAI) and Federated Learning could improve real-time prediction efficiency by 50% [18].

AI and data fusion techniques like Deep Convolutional Neural Networks (DCNNs) have achieved over 98% accuracy in remote sensing for land use and environmental monitoring, with calls to enhance XAI for better transparency in decision-making [19]. In flood-prone regions like the Middle Ganga Plain (MGP), ANFIS-GA outperforms other models (AUC success rate: 0.922, prediction rate: 0.924, validation accuracy: 0.883) for flood susceptibility mapping using 12 conditioning parameters [20].

To strengthen preparedness, especially in countries like India—among the top five globally in disaster-related mortality—a national disaster management database using open-source Big Data platforms (e.g., Hadoop) is recommended for real-time analytics from ML, satellites, geological observatories, and social media [21].

TABLE I. COMPARING DIFFERENT AI MODELS IN DISASTER PREDICTION

AI Model	Strengths	Weaknesses	Best Applications
Decision Trees (DT) & Random Forest (RF)	Simple, interpretable, effective for structured data	Prone to overfitting, limited in complex datasets	Landslide prediction, wildfire spread modeling
Support Vector MACHINES (SVM)	Effective for non-linear data, good for small datasets	Computationally expensive for large datasets	Earthquake damage assessment, storm classification
K-Nearest NEIGHBORS (KNN)	Easy to implement, minimal training required	Inefficient for large datasets	Flood risk estimation, localized hazard assessment
Artificial NEURAL NETWORKS (ANNs)	High accuracy, adaptable to multiple data types	Requires large datasets, lacks interpretability	Tsunami detection, hurricane prediction
Hybrid AI MODELS	Combines STRENGTHS OF MULTIPLE MODELS	Complex IMPLEMENTATION	Flood PREDICTION, MULTI-HAZARD ASSESSMENT

IV. MODEL

In this study, three widely used supervised learning algorithms were implemented to classify natural hazards based on historical impact data: **Decision Tree Classifier**, **K-Nearest Neighbors (KNN)**, and **Support Vector Machine (SVM)**. Each model was selected based on its distinct methodological advantages and applicability to multi-class classification tasks inherent in disaster prediction systems.

A. Decision-Tree Classifier

The Decision Tree Classifier is an algorithm that classifies rules area of the selected feature using a tree-like structure that takes data characteristics and class labels as inputs. The certain method for being easy to understand and transparency was the decision that was a crucial attribute in the given situation—disaster management which expects the model to be able to interpret and validate the results choosing the specific features. The tree structure facilitates the presentation of the decision-making process which is comprehensible, and this allows professionals in the field to find out how the change of a specific feature (e.g., the number of fatalities, economic damages, geographical area) determines the classification of disaster types such as floods, earthquakes, or wildfires.

The model operates by evaluating the "best" feature to split the data at each node using a criterion such as **Information Gain** or **Gini Impurity**:

- **Entropy**, a measure of impurity, is defined as:

$$H(D)=-\ {}_{i=1}\sum^{c} p_i \log^2 p_i \quad (1)$$

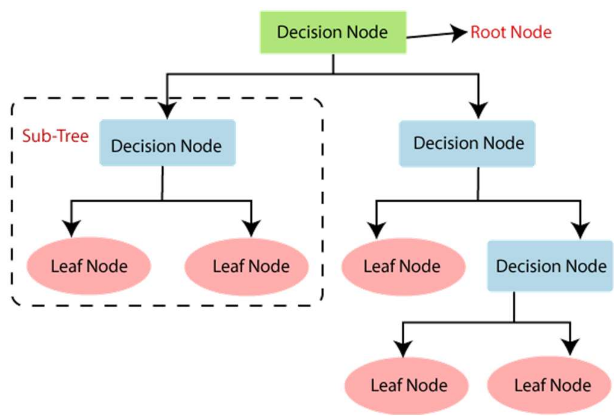

Fig. 1. Decision Tree Model Architecture

B. K-Nearest Neighbors (KNN)

K-Nearest Neighbors or KNN is a non-parametric, instance-based classifier that categorizes the new data points according to the group of neighboring data points which are defined by the majority class among them. KNN1 is highly efficient in cases where it is difficult to distinguish between different disaster types such as similar regional impacts or infrastructure damages. The technique relies on the fact that the events with similar disaster attributes have short distances among the features they occupy.

The commonly used **Euclidean Distance** formula is:

$$d(x,x')=\sqrt{}_{i=n}\sum^{n} (x_i-x_i')^2 \quad (2)$$

KNN is resource-intensive from the point that it does not need to be trained for prediction more so, the latter by calculating distance to all training points. However, its most straightforward functionality and efficient performance on structured data enable it to be used for multi-class disaster classification. The model that attained the best result from the classifiers applied was KNN with an accuracy of 79.41%.

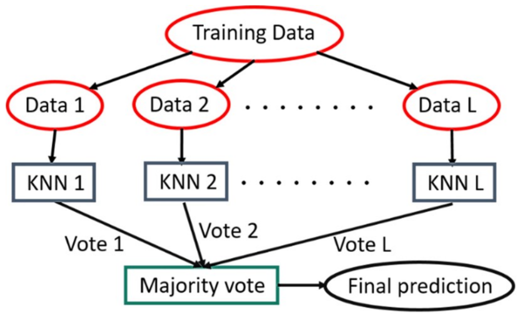

Fig. 2. Knn Model Architecture.

C. Support Vector Machine (SVM)

SVM is a potent tool that is designed to decide the optimal hyperplane that will considerably align with the classes in the feature space. With these features, we can use kernel functions to produce non-linear decision boundaries with SVM, so it is useful for high dimensional and intricate datasets. SVM can be used to find the best disaster type suitable for disasters with hardly recognizable differences from the location, disaster scale, and socio-economic impact. The model can be still a powerful tool for datasets with well-separated clusters regardless of the fact that it will not perform well on the imbalanced or heavily overlapping class situation. The strong theoretical basis and high-level robustness of SVM render it an ideal candidate for classification tasks when precision and boundary clarity are indispensable.

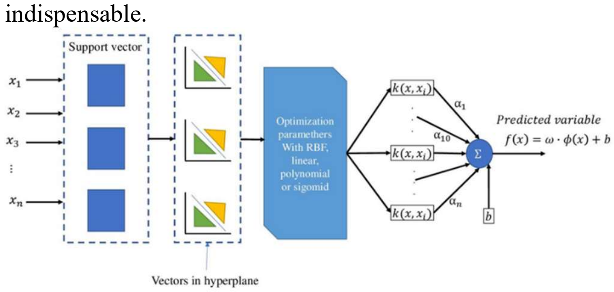

Fig. 3. SVM Model Architecture.

V. Methodology

An AI-powered natural hazard forecasting system that is useful was designed by testing many supervised machine learning models. The means adopted a process of data cleansing, model training, and the use of standard evaluation metrics for performance assessment. This part of the paper explains features of the dataset, method of modeling, and the comparison of three Decision Tree, K-Nearest Neighbors (KNN), and Support Vector Machine (SVM) classifiers.

A. Dataset Description and Preprocessing:

- One of the aspects of natural hazards contained in the dataset was the number of fatalities, economic loss, geographic location, and the year of occurrence.
- The predictive variable is the name of the natural calamity, e.g., flood, earthquake, wildfire, hurricane).
- The data was preprocessed by normalizing it, as well as by encoding the categorical variables and imputing the missing values to make the model ready.
- 80:20 split was made in such a way that 80 percent of the dataset was in the training set and the rest was in the testing set to check the generalization from the training phase.

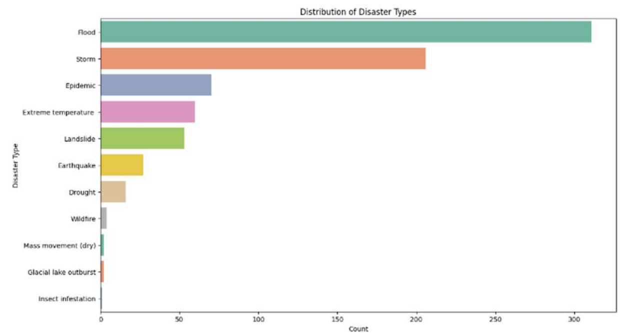

Fig. 4. Distribution of Disaster Types in Dataset

B. Model Evaluation Criteria:

The performance of each classification model was assessed using:

- Accuracy: Overall correctness of the model.
- Precision: The proportion of true positives among predicted positives.
- Recall: The proportion of true positives among actual positives.
- F1-Score: Harmonic mean of precision and recall, suitable for imbalanced data.

All metrics were calculated using weighted averages to accommodate class imbalance in the dataset.

C. Model Selection:

The machine learning model choice was mostly due to the necessity for keeping a balance in interpretability, computational efficiency, and classification performance across various disaster categories. Decision Tree, K-Nearest Neighbors (KNN), and Support Vector Machine (SVM) were preferred mostly because of their successful history in supervised classification tasks and because they have different advantages. Decision Tree has very good interpretability and can model relationships that are not necessary to be linear, which is extremely useful for decision support systems. KNN is the best option for simplicity and for the ability to find out local patterns in data, which is highly valuable when disaster events have the clustering behavior. SVM has the ability to operate stably even in the case of the high-dimensionality and non-linearity of feature spaces and to scan most of the data with characteristics of high precision, which is particularly important in such applications where false alarms need to be minimized. Thus, these models take the lead in terms of adoption in the classification methods field of natural hazard prediction.

VI. Results

This section presents the performance outcomes of the machine learning models applied for natural hazard classification. Each subsection provides a detailed analysis of evaluation metrics, visual interpretation of results, and comparative insights among the models tested.

A. Feature Correlation Analysis

A correlation matrix was produced with the aim of exploring any interconnections between the numerical features in the dataset. The matrix showed that the fatalities and economic damage had very strong correlation coefficients, thus, the conclusions by us were "Disasters with the largest scope seem to both have more money losses and involve many people who are dead." Furthermore, those variables of temporal and geographic nature were found, in some cases, to be moderately associated with given disaster types, with this information possibly suggesting the existence of certain regional or seasonal patterns. Realizing these relations is crucial for both feature selection and model interpretability. Those features that were highly correlated were selected to keep as many information as possible. However, the normalization was implemented to avoid the distraction from the features that may considerably dominate during model training because of their disproportion. The exploration of the correlations has been instrumental in identifying and setting the selected features as well as in the decision on which features to give the highest priority during the process of classification with machine learning models.

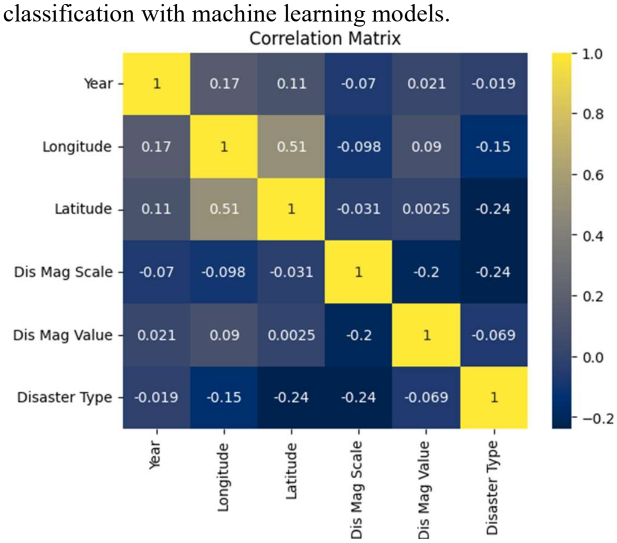

Fig. 5. Correlation Matrix on Disaster type and Its Scale and Magnitude.

B. Evaluation Metrics:

The models were assessed using **Accuracy**, **Precision**, **Recall**, and **F1-Score**, calculated as weighted averages due to class imbalance in the dataset.

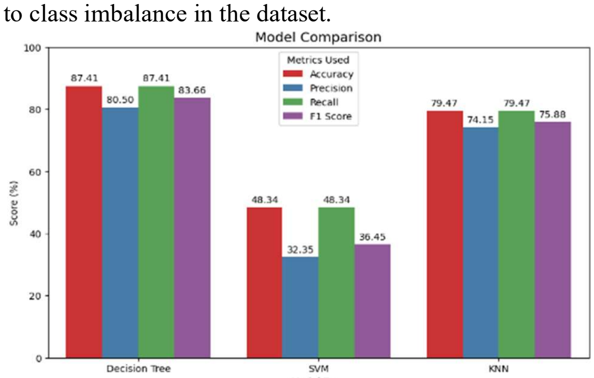

Fig. 6. Model Performance Comparison

C. Comparative Performance:

- Based on all four metrics, Decision Tree turned out to be the best model, and this suggests that it is the most suitable in cases where both interpretability and accuracy are important.
- What the analysts found was that the precision of SVM was so high that its use in systems should focus on minimizing false positives, e.g., public alert systems.
- Even with lower accuracy, KNN is still useful when data structures are more local and/ or when computer simplicity is required.

D. Error Analysis:

- These classes had more errors, mainly because of KNN when their feature distributions overlapped.
- Even all the models sometimes wrongly classified the disaster types with fewer occurrences, which shows the importance of data balancing or synthetic augmentation strategies in the future.
- An ensemble of models or the use of feature selection techniques would help in improving this model if model performance is just the issue of noise.

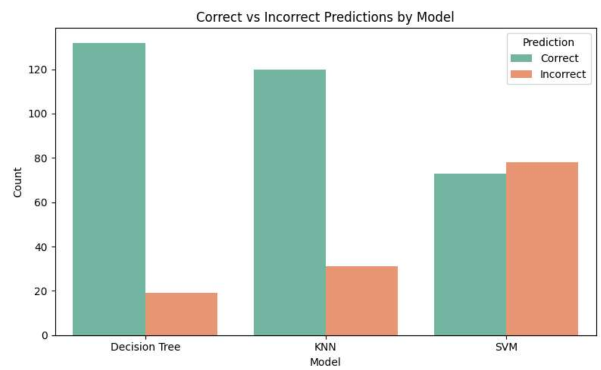

Fig. 7. Error analysis

E. Summary of Findings:

- A Decision Tree classifier counts as the study's most effective and understandable method of multi-class natural hazard prediction.
- An SVM guarantees high precision and ability for generalization while KNN ensures modest performance with simplicity.
- From the results, it is deduced that a combination of various models will be able to not only exploit the unique advantages of each individual classifier but also significantly improve the predictive capability and system reliability.

VII. Conclusion and Future Scope

This work focused on the utilization of machine learning algorithms in the identification of natural disaster events with semi-structured, disaster-related data. Three models, namely Decision Tree, K-Nearest Neighbors (KNN), and Support Vector Machine (SVM), were designed and applied to a pre-processed dataset that mostly contained attributes such as fatalities, economic losses, and geographic data. Decision Tree was the most accurate out of these classifiers and it also turned out to be very suitable for data of a complex rule-based nature and with varying categories. The values of the error function and the accuracy of the models were used to further study the generalization abilities of models as well as exceptions that can come with the assumption of overlapping class distributions and imbalanced data.

The findings emphasize the potential of AI in building early warning systems and decision-support tools for disaster risk management. By automating hazard classification, such models can assist policymakers, emergency response teams, and humanitarian organizations in timely resource allocation and risk assessment. Moreover, incorporating interpretability mechanisms into these models—especially tree-based classifiers—can enhance trust and transparency in high-stakes disaster scenarios.

However, this research also acknowledges several limitations and paves the way for future work. The dataset, while comprehensive, may not fully represent rare or extreme hazard events due to class imbalance. Addressing this through synthetic data generation (e.g., SMOTE, GANs) could improve the robustness of the models. Additionally, integrating temporal trends, satellite imagery, or real-time sensor data may lead to more context-aware and dynamic prediction systems. Future studies could also explore ensemble learning, deep learning architectures, or hybrid AI models to further enhance predictive performance.

In conclusion, AI-driven approaches hold immense promise in transforming natural hazard forecasting. With the right combination of data, model interpretability, and real-time integration, they can significantly improve the accuracy, efficiency, and responsiveness of disaster management strategies in the coming years.

REFERENCES

[1] Hanspal, Manindra Singh, and Bijayananda Behera. "Role of Emerging Technology in Disaster Management in India: An Overview." International Journal of Disaster Risk Management 6, no. 2 (2024): 133-148.

[2] Gupta, Tanu, and Sudip Roy. "Applications of artificial intelligence in disaster management." In Proceedings of the 2024 10th International Conference on Computing and Artificial Intelligence, pp. 313-318. 2024.

[3] Qin, Feng, Chengrong Huang, and Zhenjie Lin. "Big data and artificial intelligence-driven natural disaster prediction and prevention: Technological advances and prospects." Geographical Research Bulletin 3 (2024): 381-398.

[4] Chen, Wei, and Jiing Fang. "Optimizing ai-driven disaster management through llms." (2024).

[5] Chakraborty, Anusheema, and P. K. Joshi. "Mapping disaster vulnerability in India using analytical hierarchy process." Geomatics, Natural Hazards and Risk 7, no. 1 (2016): 308-325.

[6] Joshi, Jagdish Chandra, Prabhjot Kaur, Bhupinder Kumar, Amreek Singh, and P. K. Satyawali. "HIM-STRAT: a neural network-based model for snow cover simulation and avalanche hazard prediction over North-West Himalaya." Natural Hazards 103, no. 1 (2020): 1239-1260.

[7] Tiwari, Anuj, Mohammad Shoab, and Abhilasha Dixit. "GIS-based forest fire susceptibility modeling in Pauri Garhwal, India: a comparative assessment of frequency ratio, analytic hierarchy process and fuzzy modeling techniques." Natural hazards 105 (2021): 1189-1230.

[8] Hubig, Nina, Philip Fengler, Andreas Züfle, Ruixin Yang, and Stephan Günnemann. "Detection and prediction of natural hazards using large-scale environmental data." In Advances in Spatial and Temporal Databases: 15th International Symposium, SSTD 2017, Arlington, VA, USA, August 21–23, 2017, Proceedings 15, pp. 300-316. Springer International Publishing, 2017.

[9] Ward, Philip J., Veit Blauhut, Nadia Bloemendaal, James E. Daniell, Marleen C. de Ruiter, Melanie J. Duncan, Robert Emberson et al. "Natural hazard risk assessments at the global scale." Natural Hazards and Earth System Sciences 20, no. 4 (2020): 1069-1096.

[10] Nitsche, Maximilian, S. Karthik Mukkavilli, Niklas Kühl, and Thomas Brunschwiler. "AB2CD: AI for Building Climate Damage Classification and Detection." In Proceedings of the AAAI Symposium Series, vol. 2, no. 1, pp. 115-123. 2023.

[11] Leinonen, Jussi, Ulrich Hamann, Ioannis V. Sideris, and Urs Germann. "Thunderstorm Nowcasting With Deep Learning: A Multi-Hazard Data Fusion Model." Geophysical Research Letters 50, no. 8 (2023): e2022GL101626.

[12] Katsuumi, Arisa, Yuxin Cong, and Shinya Inazumi. "AI-Driven Prediction and Mapping of Soil Liquefaction Risks for Enhancing Earthquake Resilience in Smart Cities." Smart Cities 7, no. 4 (2024): 1836-1856.

[13] Bari, Lazima Faiah, Iftekhar Ahmed, Rayhan Ahamed, Tawhid Ahmed Zihan, Sabrina Sharmin, Abir Hasan Pranto, and Md Rabiul Islam. "Potential use of artificial intelligence (AI) in disaster risk and emergency health management: a critical appraisal on environmental health." Environmental health insights 17 (2023): 11786302231217808.

[14] Esparza, Miguel, Bo Li, Junwei Ma, and Ali Mostafavi. "AI Meets Natural Hazard Risk: A Nationwide Vulnerability Assessment of Data Centers to Natural Hazards and Power Outages." arXiv preprint arXiv:2501.14760 (2024).

[15] Kumar, Deepak, Manoj Thakur, Chandra S. Dubey, and Dericks P. Shukla. "Landslide susceptibility mapping & prediction using support vector machine for Mandakini River Basin, Garhwal Himalaya, India." Geomorphology 295 (2017): 115-125.

[16] Albahri, A. S., Yahya Layth Khaleel, Mustafa Abdulfattah Habeeb, Reem D. Ismael, Qabas A. Hameed, Muhammet Deveci, Raad Z. Homod, O. S. Albahri, A. H. Alamoodi, and Laith Alzubaidi. "A systematic review of trustworthy artificial intelligence applications in natural disasters." Computers and Electrical Engineering 118 (2024): 109409.

[17] Khan, Amina, Sumeet Gupta, and Sachin Kumar Gupta. "Multi-hazard disaster studies: Monitoring, detection, recovery, and management, based on emerging technologies and optimal techniques." International journal of disaster risk reduction 47 (2020): 101642.

[18] Choudhary, Ankit, Pushkar Matliwala, Ms. Rachna Jain, Harsh Vyas, and Pritesh kumar Vijay Solanki. "NATURAL DISASTER PREDICTION SYSTEM." International Journal of Novel Research and Development 9, no. 3 (2024): e522-c522

[19] Himeur, Yassine, Bhagawat Rimal, Abhishek Tiwary, and Abbes Amira. "Using artificial intelligence and data fusion for environmental monitoring: A review and future perspectives." Information Fusion 86 (2022): 44-75.

[20] Arora, Aman, Alireza Arabameri, Manish Pandey, Masood A. Siddiqui, U. K. Shukla, Dieu Tien Bui, Varun Narayan Mishra, and Anshuman Bhardwaj. "Optimization of state-of-the-art fuzzy-metaheuristic ANFIS-based machine learning models for flood susceptibility prediction mapping in the Middle Ganga Plain, India." Science of the Total Environment 750 (2021): 141565.

[21] Goswami, Saptarsi, Sanjay Chakraborty, Sanhita Ghosh, Amlan Chakrabarti, and Basabi Chakraborty. "A review on application of data mining techniques to combat natural disasters." Ain Shams Engineering Journal 9, no. 3 (2018): 365-378.