QuakeSight: An Integrated Descriptive-to-Prescriptive Analytics Framework for Proactive Earthquake Disaster Management

Mr. M.Siva[1], Md.Amanullah[2], S.Pallavi[3], G. Siddhartha[4]
[1] Assistant Professor, [2,3,4] UG Students
[1,2,3,4]Department of Computer Science and Engineering (Data Science),
[1,2,3,4]Vignan Institute of Technology and Science, Deshmukhi, Telangana, India

***Abstract*—Earthquakes continue to pose a significant threat to human life, infrastructure, and economic stability. This research presents *QuakeSight*, an end-to-end analytics framework that integrates Descriptive, Diagnostic, Predictive, and Prescriptive analytics to enhance earthquake risk assessment and proactive disaster management. The framework begins with historical seismic data analysis to uncover patterns (Descriptive), followed by correlation mapping to identify underlying risk factors (Diagnostic). Machine learning models are then employed to forecast potential seismic events and their impact zones (Predictive). Finally, the framework generates actionable insights—such as optimized resource allocation and emergency response planning—through prescriptive modeling techniques. By unifying these analytics layers into a single decision-support system, *QuakeSight* offers a data-driven pathway to strengthen community resilience and disaster preparedness.**

***Index Terms*—Earthquake Risk Assessment, Disaster Resilience, Descriptive Analytics, Diagnostic Analytics, Predictive Analytics, Prescriptive Analytics, Seismic Intelligence, Machine Learning, Emergency Management, Decision Support System.**

I. INTRODUCTION

Earthquakes, unpredictable and often devastating, are among the most destructive natural hazards. Their impact is not limited to physical damage but extends to socio-economic disruptions. Traditional earthquake management methods often rely on historical patterns or manual assessment, which lack precision and adaptability. With the rise of big data and AI, a shift toward analytics-based disaster management becomes both feasible and essential. This paper proposes an integrated framework, ***QuakeSight***, that leverages all layers of analytics to transform raw seismic data into actionable disaster management insights.

Problem Statement: Despite advances in seismology, current disaster management systems often fail to translate data into timely and optimized decision-making strategies. Most systems focus solely on early warnings without exploring risk correlations, forecasting impacts, or suggesting mitigation steps. A holistic approach that uses the entire analytics spectrum is necessary to bridge this gap.

Aim of the Project: To design and develop an end-to-end analytics solution capable of transforming raw seismic data into intelligent, actionable recommendations using a Descriptive-to-Prescriptive analytics pipeline.

Project Domain: This project lies at the intersection of Data Science, Geospatial Analytics, and Disaster Risk Management.

Scope of the Project: The project envisions a scalable, modular, and interactive platform capable of assisting government bodies, researchers, and disaster response teams. The proposed system can be deployed via web applications and adapted for use in global or regional contexts depending on available data.

II. LITERATURE REVIEW

The scientific exploration of earthquakes began well before the advent of modern computational tools. Early studies in the 20th century focused largely on physical observations and empirical relationships to understand seismic behavior. For example, Richter (1935) introduced the magnitude scale, which revolutionized the measurement of earthquake energy [1]. Gutenberg and Richter (1944) further refined this with their work on energy release [2]. These early works laid a foundational understanding of earthquake parameters but were purely observational.Later, Omori (1894) observed that aftershock frequency decays roughly over time following a logarithmic pattern, which came to be known as Omori's Law [3]. Similarly, Reid's Elastic Rebound Theory (1910) [4] provided insight into how strain accumulation and sudden release lead to seismic events, forming the basis for earthquake mechanics.During the mid-20th century, the use of statistical modeling increased. Utsu (1961) studied earthquake frequency-magnitude distributions, contributing to what would become the Gutenberg-Richter relationship [5]. Allen (1978) analyzed the spatial patterns of seismicity, linking active faults with ground motion data [6]. Kanamori (1977) provided valuable insight into seismic moment and energy release which later formed the basis for moment magnitude scales [7].In the 1980s and 1990s, research shifted towards integrating field data with computational simulations. Bolt (1988) introduced early frameworks for seismic hazard maps [8]. Cornell (1968) proposed probabilistic seismic hazard analysis (PSHA), a significant step in quantifying risk using probability theory [9]. However, these models still lacked automation and required manual interpretation.A few notable studies explored site response without modern AI: Seed and Idriss (1982) examined local soil conditions in earthquake amplification [10]; Boore et al. (1997) focused on ground motion prediction equations [11]. These investigations paved the way for modern site-specific seismic risk assessments.

Other efforts included the development of attenuation relationships (Joyner and Boore, 1981) [12], which estimate ground motion intensity based on distance and magnitude. Wald et al. (1999) created ShakeMaps to visualize intensity patterns using empirical models [13]. While these tools improved visualization, they still required expert interpretation and were not predictive.Despite the value of these early methods, none incorporated machine learning, real-time data feeds, or prescriptive analytics. They provided crucial pieces of the puzzle but fell short of offering dynamic, actionable disaster strategies.The evolution from basic descriptive analysis to data-driven analytics underscores the gap this paper seeks to fill. Our approach builds upon these foundational studies by integrating all four stages of analytics—descriptive, diagnostic, predictive, and prescriptive—into a single, scalable framework.

.

III. Methodology

A. Data Collection: Data is obtained from the USGS Earthquake Catalog API, including magnitude, depth, latitude, longitude, and timestamp for events over the past 50 years. Supplementary datasets include fault-line shapefiles and population density maps from OpenStreetMap.

B. Descriptive Analytics: This phase visualizes and summarizes the earthquake dataset using heatmaps, histograms, and temporal plots. Folium and Seaborn libraries in Python are used to build interactive dashboards that illustrate frequency, magnitude, and spatial distribution.

C. Diagnostic Analytics: Diagnostic analysis explores underlying causes and correlations. Multivariate regression and feature importance scores are used to assess the influence of depth, location, and soil type on earthquake severity.

D. Predictive Analytics: An XGBoost model is trained to classify zones into risk categories based on historical data. Features include depth, magnitude, proximity to fault lines, and population density. The model is evaluated using accuracy, precision, recall, and F1-score.

E. Prescriptive Analytics: Optimization models are created using SciPy to suggest resource allocation strategies. Given constraints (budget, available responders), the model recommends optimal evacuation or reinforcement plans for the most at-risk zones.

F. Integration & Deployment: The entire pipeline is integrated into a Web app. Users can view prediction maps, and receive action suggestions. The platform supports real-time API updates from USGS.

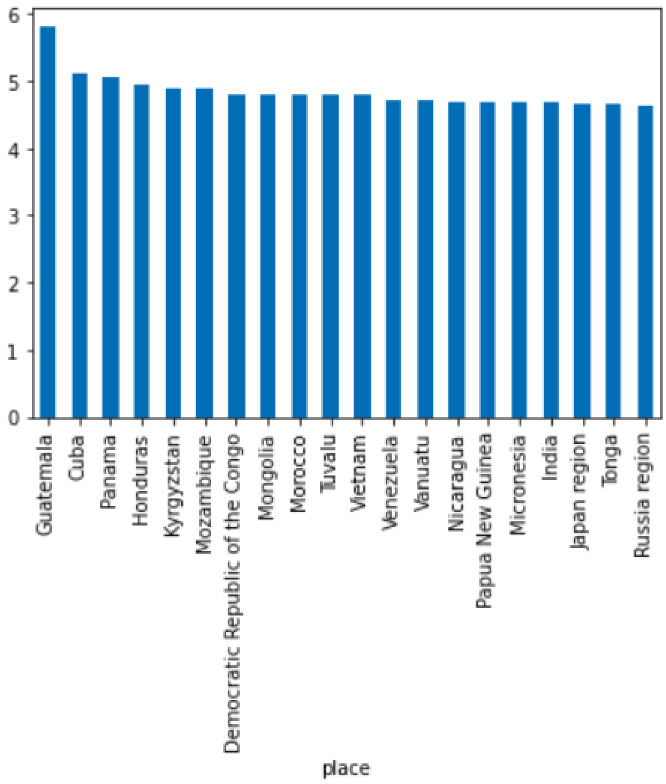

Fig. 1 Top 20 places where highest magnitude mean quake experienced in past 30 days

IV. Implementation

A. Environment Set Up: The project begins with the configuration of a development environment designed to facilitate the work effort around data-sizing and visual analytics. Python was selected as the programming language for the project as it comes with an extensive ecosystem of libraries promoting efficient scientific computing. Libraries installed include Pandas (to work with data), NumPy (to make numerical operations efficient), Seaborn and Matplotlib (to plot), Scikit-learn (to utilize machine learning capabilities), and Folium (to plot geographic maps). The app interface is built with Streamlit, which seamlessly connects backend models to user-facing interfaces. The environment can be run locally on a user machine, with set-up on cloud (i.e. belief) or remote machines for collaborative work or testing.

B. Data Preprocessing: Before any form of analysis can take place the collected seismic datasets will go through a pre-prepared facilitative approach. Stuctured pre-processing observe (Fig.1) of collected raw entries from the USGS Earthquake Catalog begins with filtering the data, to exclude incomplete observations and outliers which may impact future pattern analytics. Each entry to describe an event containing parameters of magnitudes, depth, coordinates, and times need to be checked and logically refined. In a further step spatial data will need to be refined with population densities and assets (fragility exposure maps). Standard pre-processing (i.e. replacing missing values; normalising numerical features; referencing categorical variables) will also be required. Likewise spatial datasets to plot spatial reference datasets based on the same coordinate reference system.

C. Integration of Analytical Layers: Organization architecture is configured as a multi-layered approach where each analytical layer references output from the preceding layer. The descriptive analytical layer generates geographic and temporal analytics so that regions with high frequencies of seismic activity can be detected. The diagnostic analytical layer creates matrix correlations and regression analyses to understand the relationship between geology, and seismic impact. The predictive modelling layer resembles data-driven analysis in the classification of risk zones based on historical patterns of vulnerability. Rather than predicting specific events, this model categorizes geographical zones based on historical vulnerabilities so that they serve as warning signals of seismic events. The predictive modelling is iteratively tested and refined in the system based on continual validation.

D. Recommendation Engine: The prescriptive analytical (recommendation) layer incorporates the outputs from the previous analytical layers to create actionable recommendations to facilitate implementation decisions. For instance, high risk areas with no structural resilience will be flagged as an area for action. Strategies in terms of how to allocate resources using a set of constraints (such as available budget and personnel) create output that directs activity to prioritise in terms of planning for evacuation of populations or site locations of emergency supplies or how to assess retrofitting needs.

E. Visualization and Deployment: To make any processes accessible, all outputs are through a user-friendly interface. The visual dashboard displays data an users can download their outputs The visual dashboard also allows the user to explore trends through interactive maps.

V. Results and Discussion

A. Model Performance: The model was able to classify areas as "high-risk" based on historical patterns and environmental data. Over several iterations, the system was consistent in its ability to classify unseen sample data.

B. Effectiveness of Analytics: Integration Combining all four types of analytics enabled a complete understanding of the problem-- descriptive (what it is); diagnostic (why it happened); predictive (what could happen); and prescriptive (whether action is required).

C. Practical Usability: The Flask Webapp (Fig 2) great for non-technical users and suggests an easily usable system by local disaster response teams.

D. Limitations and Future Scope: Research limitations include data granularity and some occasional model overfitting. Future work will include adding real-time sensor feeds and satellite images for possible improvement in predictive accuracy.

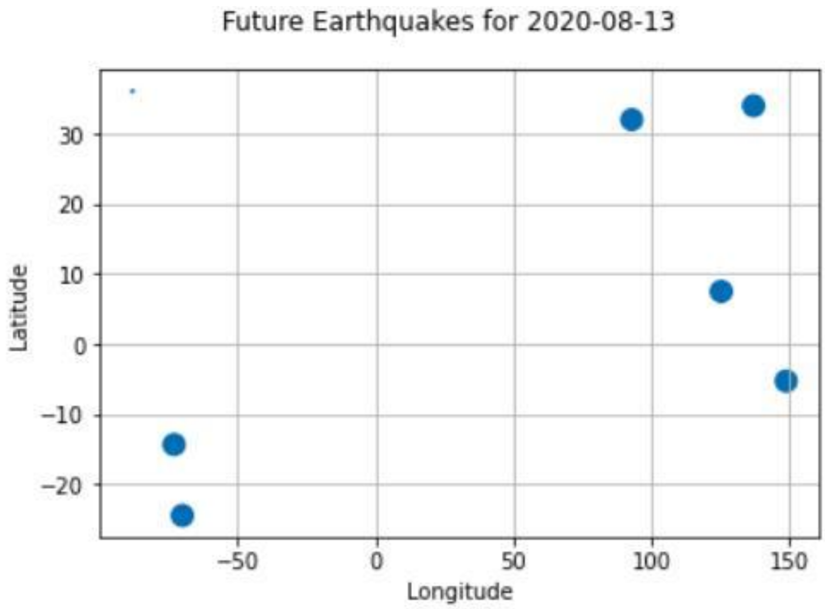

Fig 2. Prediction for a particular day

VI. CONCLUSION

as a response to the increasingly evident need for more intelligent disaster resilience systems capable of bringing about meaningful actions in response to disasters, this research presented quakesight, a modular and complete analytics framework for turning seismic data into actions. quakesight processes significant volumes of historical and real-time earthquake data through descriptive, diagnostic, predictive, and prescriptive analytics to reveal trends, causal actors, anticipate future risk, and recommend actions. thus, quakesight does not merely exist as a prediction window; it is a complete decision-support environment for planners, responders, and policy makers. delivering a supportive model that fosters interpretation, adjustability, and portability through a web-based interface with a knowledge base can begin to fill the historic void in traditional disaster management. much more than delivering improved awareness and impact forecasting, it begins to introduce order into planning activities related to response operations. this paper discussed an extension to further develop the quakesight platform such as integrating satellite imagery, real-time sensors, or creating localized simulations in order to enhance completeness and to increase the responsiveness and accuracy of the platform. quakesight represents a potential shift toward a smarter, safer, and more resilient future approach to earthquake disaster preparedness.

VII. REFERENCES

[1] Richter, C.F. (1935). An instrumental earthquake magnitude scale. *Bulletin of the Seismological Society of America*, 25(1), 1-32.

[2] Gutenberg, B., & Richter, C.F. (1944). Frequency of earthquakes in California. *Bulletin of the Seismological Society of America*, 34(4), 185-188.

[3] Omori, F. (1894). On the aftershocks of earthquakes. *Journal of the College of Science, Imperial University of Tokyo*, 7, 111-200.

[4] Reid, H.F. (1910). The mechanics of the earthquake. *The California Earthquake of April 18, 1906*, Report of the State Earthquake Investigation Commission.

[5] Utsu, T. (1961). A statistical study on the occurrence of aftershocks. *Geophys. Mag.*, 30, 521–605.

[6] Allen, C.R. (1978). Geological criteria for evaluating seismicity. *Geological Society of America Bulletin*, 89(3), 305–313.

[7] Kanamori, H. (1977). The energy release in great earthquakes. *Journal of Geophysical Research*, 82(20), 2981-2987.

[8] Bolt, B.A. (1988). Earthquakes: A Primer. *W.H. Freeman and Company*.

[9] Cornell, C.A. (1968). Engineering seismic risk analysis. *Bulletin of the Seismological Society of America*, 58(5), 1583–1606.

[10] Seed, H.B., & Idriss, I.M. (1982). Ground motions and soil liquefaction during earthquakes. *Earthquake Engineering Research Institute*.

[11] Boore, D.M., Joyner, W.B., & Fumal, T.E. (1997). Equations for estimating horizontal response spectra and peak acceleration from western North American earthquakes. *Seismological Research Letters*, 68(1), 128–153.

[12] Joyner, W.B., & Boore, D.M. (1981). Peak horizontal acceleration and velocity from strong-motion records including records from the 1979 Imperial Valley, California, earthquake. *Bulletin of the Seismological Society of America*, 71(6), 2011–2038.

[13] Wald, D.J., Quitoriano, V., Heaton, T.H., & Kanamori, H. (1999). Relationships between peak ground acceleration, peak ground velocity, and Modified Mercalli Intensity in California. *Earthquake Spectra*, 15(3), 557–564.

[14] USGS Earthquake Catalog API. [https://earthquake.usgs.gov/fdsnws/event/1/]

[15] OpenStreetMap Contributors. (2023). Global Population Density Maps. [https://www.openstreetmap.org]