Grenze International Journal of Engineering and Technology, June Issue

Earthquake Forecasting in Kyrgyzstan using Machine Learning: A Random Forest Approach

Ravinder Kumar[1] and Priya sharma[2]
[1-2]Department of Mathematics, Chandigarh University, Mohali, Punjab, India.
Email: ravinderkumar2@gmail.com, priyusharma3020@gmail.com

***Abstract*—This study is focused on earthquake forecasting in Kyrgyzstan using machine learning (ML). Earthquake are among the most devasting natural disasters, often resulting in extensive destruction and significant loss of life. Predicting earthquakes can help in reducing the damage caused by earthquakes and save human lives. ML method analyze massive datasets and generate different models over time with focusing on performance. ML method called Random Forest (RF) to predict earthquakes in Kyrgyzstan region is used to experiences frequent earthquakes. We used data such as earthquake size, depth, location, timing, and distance from fault lines to train the model. The RF method was also able to find out the patterns present in the data and predict places, times for earthquakes. Most importantly, this method was able to identify factors causing earthquakes, tectonic stress, and the activity of a fault line. This research shows the improvement in earthquake prediction by using RF method so that there may be better warning system for Kyrgyzstan. Future research will involve the improvement of the model to include real-time data and expand its application to earthquake-prone regions worldwide.**

***Index Terms*— Machine Learning, Random Forest method, Earthquake Prediction, Seismic data analysis, Magnitude and depth of earthquake.**

I. INTRODUCTION

Earthquake prediction is a branch of seismology focused on determining the time, location, and magnitude of future earthquakes within specified limits. Earthquakes are caused by the release of energy within the Earth's crust, resulting in ground shaking and movements. Like other natural disasters, they lead to severe destruction, loss, and injuries. According to Chaurasia et al. [1], earthquakes occur daily around the world, generating seismic waves that liberate energy. These seismic waves can cause significant damage and loss of life. Asencio-Cortés et al. [2] discussed that Japan, Southern California, Indonesia, Iran, Taiwan, and Turkey frequently experience earthquakes. Similarly, Chin et al. [3] also highlighted these regions. Earthquake can be felt by people if their magnitude is more than 2.5, but minor earthquakes are usually not felt. According to Mousavi et al. [4] and Salam et al. [5], most earthquakes with magnitudes greater than 4.5 cause extensive damage and can sometimes lead to a high loss of life.
In ancient times, people relied on natural signs, such as unusual animal behaviour and atmospheric disturbances to predict earthquakes. The first seismoscope was developed by Chinese scientist Zhang Heng in 132 AD, allowed for the detection of distant earthquakes. At the same time, in Japan, earthquakes were attributed to mythical creatures like Numazu and in ancient Greece, philosopher Thales suggested that underground water

Grenze ID: 01.GIJET.11.2.553_1

movement caused seismic activity. As scientific research progressed, Robert Mallet proposed in 1855 that movements along faults in the Earth's crust caused earthquake. As San Francisco [6] explained, this later influenced the development of the elastic rebound theory after the 1906 earthquake. Between 1960s and 1980s, researchers studied foreshocks and identified seismic gaps to test various earthquake prediction methods. While some approaches showed promise others proved ineffective.

In 1975, China successfully predicted the Haicheng earthquake magnitude 7.3, saving thousands of lives but it failed to predict the Tangshan earthquake magnitude 7.5 the following year which caused massive destruction. As technology advanced, researchers improved earthquake forecasting and by the 21st century, scientists began using artificial intelligence, satellite-based monitoring and geophysical modelling to study earthquake patterns. Countries such as Japan, U.S., Mexico and Taiwan implemented Earthquake Early Warning Systems (EEWS), which send alerts before strong shaking begins, helping to reduce damage, improve disaster preparedness and save lives. Despite these advancements accurately predicting earthquakes remains a challenge. Short-term predictions (within hours to weeks) are highly reliable in theory but extremely difficult to achieve due to the unpredictable nature of seismic activity. Long-term predictions which estimate earthquake probabilities over months, years or even decades, rely on statistical models using historical seismic data fault line movements and geological studies. Although less precise, these predictions help with urban planning and disaster preparedness.

ML, a branch of Artificial Intelligence (AI), helps computers learn from data and make predictions without explicit programming. ML algorithms process large datasets identify patterns and improve accuracy over time. Methods like RF provide new ways to predict earthquakes. Researchers widely use RF in natural hazard forecasting because it manages large unstructured datasets and identifies key predictive factors. By combining multiple decision trees, it enhances accuracy and reliability in predictions. Traditional methods rely on observable signs like ground movements and changes in water levels but struggle to handle complex data. In contrast, modern ML-based methods uncover patterns that older techniques often miss. This study applies the RF technique to predict earthquakes in Kyrgyzstan. By analysing historical earthquake data identifying key predictive factors and developing a model it aims to improve disaster preparedness and response by predicting the magnitude, depth, and location of future earthquakes in the region.

II. Literature Review

Mallouhy et al. [7] outline the possibility of predicting major earthquake events by using several ML algorithms. The topic of their article continues developing from previous studies to achieve higher prediction precision ensuring adequate preparation towards disaster response. Ridzwan and Yusoff [8] explains the ML method, which can predict big earthquakes by studying previous data ways in the prediction of correct outcomes and preparations for disaster. Mignan and Broccardo [9] look at how neural networks have been used to predict earthquakes over 25 years. It reviews past studies and discusses the challenges and limits of using neural networks for accurate earthquake predictions. Berhich et al. [10] explore the use of LSTM models in earthquake prediction and explain how such models improve the analysis of seismic data to make more accurate predictions and understand the patterns of earthquakes. Gürsoy et al. [11] analyses how ML and deep learning can be help in the prediction of earthquakes. It reviews algorithms and their prospects for future potential in the forecasting of earthquakes. Banna et al. [12] looks at how artificial intelligence (AI) is used to predict earthquakes. It explains the current AI methods and their strengths in making prediction as well as the challenges and future opportunities in improving earthquake forecasting. Kulkarni et al. [13] explores the application of ML techniques for predicting earthquakes. It discusses how various algorithms can analyze seismic data to improve prediction accuracy, helping with early warning systems and disaster preparedness. Mukherjee et al. [14] introduces a new method for predicting earthquakes. It combines different seismic signals and indicators to make more accurate predictions using an ensemble approach, which helps improve the reliability of earthquake forecasts. A recent study by Mousavi and Beroza [15] explored how ML is being used to improve earthquake science. Researchers are using these powerful computer techniques to better analyze earthquake data. This helps them more accurately detect when earthquakes happen, pinzpoint where they occur and determine their size. Beroza et al. [16] discussed the improvement of earthquake prediction through ML. They reviewed the recent progress in these computational methods discussed current issues that hinder achieving accurate predictions and also presented future steps for improving reliability. Geller [17] in his 1997 study summed up the difficulties in predicting an earthquake He placed a conclusion that short-term prediction on when an earthquake would occur is impossible. It is then suggested that long-term planning on how people would be safe is more recommended to alleviate the effects of earthquakes. Scholz [18] explains in his book how earthquakes occur and how faults in the Earth's crust behave. He describes the processes that lead to earthquakes providing a deeper understanding of the science behind these events. A book by Stein and Wysession [19] explains the basics of earthquakes and what the Earth is made of. It covers how earthquakes happen, how the vibrations (seismic waves) travel through the Earth, and what these vibrations tell us about the Earth's interior. Geib et al. [20] utilized data from multiple remote sensing sensors, integrating it with ML methods to determine the structural types of buildings in seismic the work shows that the fusion of optical and radar data is suitable for automatic classification which therefore facilities the application of seismic vulnerability assessments on large scales. Tyralis et al. [21], conducted a comprehensive review forest, focusing on their applications and advantages for water scientists and practitioners. The study indicated the increasing trend of using RFs in resource management given their flexibility, accuracy and ability to deal with non-linear relationships and high-dimensional data. Kumar and Saini [22-23] investigated the propagation of surface-type seismic waves in a porous layer using Biot’s theory of elasticity, examining the influence of interface irregularity along with parameters such as porosity, initial stress, inhomogeneity, and anisotropy. Saini and Kumar [24] explored the impact of rigidity and parabolic interface irregularity on Love wave propagation in a porous layer over a nonhomogeneous elastic half-space. Kumar et al. [25] studied Rayleigh wave propagation in a semi-infinite elastic medium. Saini and Poonia [26] provided a review of seismic wave propagation behaviour, while Saini and Kumar [27] analysed Love waves in an initially stressed transversely isotropic porous layer over a nonhomogeneous elastic half-space with parabolic interface irregularity.

III. DATASET

The dataset [28] offers an overview of seismic activity in Kyrgyzstan and nearby regions from 2014 to 2024, including all necessary parameters for earthquake analysis such as frequency, latitude, longitude, depth (in km), intensity, and magnitude. This data enables the identification of trends and facilitates predictions based on past seismic events, focusing on earthquakes magnitude, depth and location. It is ideal for geospatial analysis, time series analysis and predictive modelling.

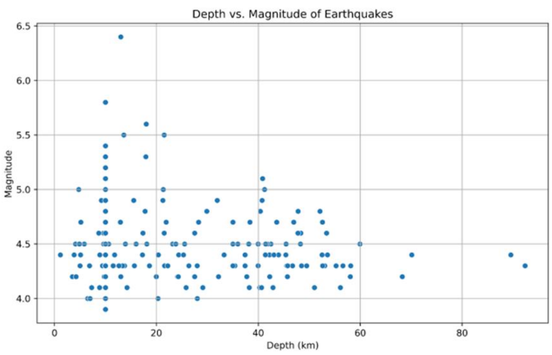

Figure 1 Depth vs magnitude of earthquakes in Kyrgyzstan

Fig.1 shows how the depth of earthquakes correlates with their magnitudes in Kyrgyzstan. The x-axis represents magnitude. The graphical representation above shows that most of the earthquakes occurred in the considered time period are shallow with moderate magnitude. Shallow refers to earthquakes that occur near the earth’s surface typically between 0 and 70 kilometres deep. They are felt more intensively because they occur closer to where people live. The magnitude in the graph represents the size or energy released by an earthquake. In this graph, the earthquakes magnitudes mostly range between 4.0 and 5.5, which are considered moderate earthquakes. A few earthquakes in the graph have magnitudes above 6.0, which are much stronger and a can release a significant amount of energy and the magnitudes points in the graph show that while most earthquakes in Kyrgyzstan are moderates, the presence of some large magnitude events highlights the more destructive earthquakes in the region.

Additionally, the graph shows that most earthquakes in Kyrgyzstan occur at depths less than 40 km. Therefore, this means that the earthquakes are happening in the upper crust part of the earth where the tectonic plates pull and push against each other, causing cracks and movements along faults. As the depth increases beyond 40 km. This suggest that most of the seismic activities within the region tend to occur at the crust because this is the region with strongest tectonic forces.

Fig.2 shows the sizes of earthquakes in Kyrgyzstan. The bottom line (horizontal axis) shows how strong the earthquakes were, from 4.0 to 6.5. The side (vertical axis) shows how many earthquakes happened at each strength.

The histogram indicates that most of the earthquakes have magnitudes between 4.0 and 4.5, with the highest number around 4.3. Moreover, tall bars mean more earthquakes. The tallest bar is around 4.5 magnitude, meaning that most of the earthquakes are around this strong. As strength increases, the bars decrease in size, indicating that stronger earthquakes, such as above 5.0, occur less frequently. Extremely strong earthquakes, above 6.0, almost never happen and this is why the far-right bars are very short.

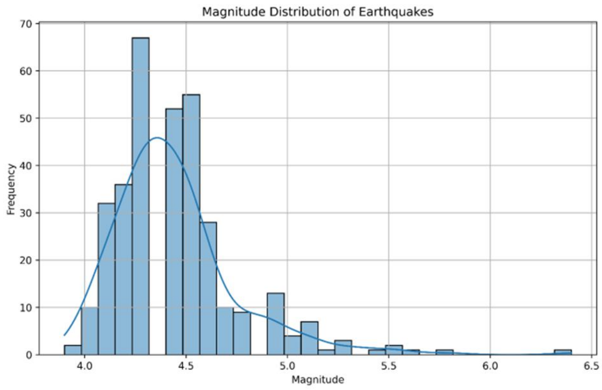

Figure 2 Frequency vs Magnitude of earthquakes in Kyrgyzstan

IV. RANDOM FOREST

Cutler et al. [29] explained that Leo Breiman introduced RF a ML method, in 2001.It works on the principle of creating multiple predictions by many small decision trees to make better overall predictions. In training, it creates a large number of decision trees, with each making its own predictions using a bootstrapped version of the data and a random subset of features at each step. This randomness makes the trees different from each other, which helps prevent the model from overfitting (focusing too much on details in the training data). For classification tasks, RF takes a vote from all the trees to decide the final output and for regression tasks, it averages their predictions. It is useful because it can handle complex data and works well even if the data is messy, noisy or has missing values. RF can also tell which features in the dataset are most important for making predictions which is helpful for understanding the data. While RF is quite accurate, including in classification tasks (like spam detection) and regression tasks (like predicting prices), it is also slow for very large datasets or when creating a lot of trees as mentioned by Ahmed [30]. Also, it is harder to interpret than more simplistic models like a single decision tree. Yet, it is very commonly used in finance, health care, fraud detection and even in predicting earthquakes by analysing patterns within huge datasets for making reliable predictions. In short, it's a very powerful and versatile tool to solve real-world problems.

V. RESULT

The earthquake prediction Table 1 show stores information about expected future earthquakes such as their size in magnitude and depth. This data helps to predict where the next earthquake might be and its magnitude which allows researches and emergency service to prepare better for earthquakes. This information on help to make good plans for safety and minimize damage during future events, such as their size in magnitude and depth. This data helps to predict where the next earthquakes might be and its magnitude.

TABLE I: PREDICTED PARAMETER

Parameter	Prediction value
Predicted next earthquakes magnitude	4.43
Predicted next earthquake depth	44.34
Predicted location	74 km NW of Murghab, Tajikistan

VI. CONCLUSION

This study shows how ML, especially the RF algorithm, can be used to predict earthquakes by analysing seismic data from Kyrgyzstan and nearby areas. RF is useful because it can handle large and complex datasets to find important patterns and make accurate predictions. By using past earthquake data, the model was able to predict

key details about future earthquakes, like their size (magnitude), depth, and location. This information can help communities prepare better and reduce the risks associated with earthquakes. The study is important because it combines ML with knowledge from geologists and scientists to make the predictions more reliable. It further emphasizes the need for real-time data, bigger datasets from different regions and combining multiple ML methods to improve accuracy. While traditional earthquake prediction methods are useful, ML offers a faster and more precise way of analysing large amounts of data.
In conclusion, using advanced techniques like RF, along with teamwork between different experts can improve earthquake prediction. These efforts can lead to better early warning systems, safer city planning and more effective disaster preparation ultimately helping save lives and reduce the damage caused by earthquakes.

FUTURE WORK

Future research in earthquake prediction should focus on improving accuracy and practicality. Using real-time data, like live seismic wave measurements or satellite observations, can help models detect changes quickly and provide early warnings. Expanding datasets by including more regions and longer time periods will make predictions more reliable. Combining RF with other ML methods, such as neural networks can enhance predictive accuracy. Collaboration with geologists and seismologists is essential to link data with real-world earthquake science. Ultimately, these advancements aim to create better early warning systems help communities prepare and reduce the impact of earthquakes.

REFERENCES

[1] K. Chaurasia, S. Kanse, A. Yewale, V.K. Singh, B. Sharma, and B.R. Dattu. “Predicting damage to buildings caused by earthquakes using Machine Learning techniques”. In 2019 IEEE 9th International Conference on Advanced Computing (IACC), (pp. 81-86), 2019.

[2] Asencio-Cortés, F. Martínez-Álvarez, A. Troncoso, and A. Morales Esteban, “Medium–large earthquake magnitude prediction in Tokyo with artificial neural networks,” Neural Computer. Appl., vol. 28, no. 5, pp. 1043–1055, 2017, Doi: 10.1007/s00521-015-2121-7.

[3] T. L. Chin, C. Y. Huang, S. H. Shen, Y. C. Tsai, Y. H. Hu, and Y. M. Wu, “Learn to Detect: Improving the Accuracy of Earthquake Detection,” IEEE Trans. Geosci. Remote Sens., vol. 57, no. 11, pp. 8867–8878, 2019, doi: 10.1109/TGRS.2019.2923453

[4] S. M. Mousavi, Y. Sheng, W. Zhu, and G. C. Beroza, “STanford Earthquake Dataset (STEAD): A Global Data Set of Seismic Signals for AI,” IEEE Access, vol. 7, pp.

[5] M. A. Salam, L. Ibrahim and D. S. Abdelminaam, “Earthquake prediction using hybrid Machine Learning techniques.” International Journal of Advanced Computer Science and Applications, vol. 12(5), pp. 654-665, 2021.

[6] D. R.Dean,. The San Francisco earthquake of 1906. Annals of science, vol.50(6), pp.501-521,1993.

[7] R. Mallouhy, C. Abou Jaoude, C. Guyeux, and A. Makhoul. “Major earthquake event prediction using various Machine Learninga algorithms.” In 2019 International Conference on Information and Communication Technologies for Disaster Management (ICT-DM) (pp. 1-7). IEEE, 2019.

[8] N.S.M. Ridzwan and S. H. M. Yusoff. “Machine Learning for earthquake prediction: a review”. Earth Science Informatics, vol. 16(2), pp. 1133-1149, 2023, 2017-2021.

[9] Mignan and M. Broccardo.” Neural network applications in earthquake prediction: Meta-analytic and statistical insights on their limitations”. Seismological Research Letters, vol. 91(4), pp. 2330-2342,2020, 1994-2019.

[10] Berhich, F. Z. Belouadha and M.I. Kabbaj. “LSTM-based models for earthquake prediction”. In Proceedings of the 3rd International Conference on Networking, Information Systems & Security (pp. 1-7),2020.

[11] G. Gürsoy, A. Varol and A. Nasab. “Importance of Machine Learning and deep learning algorithms in earthquake prediction”: a review. In 2023 11th International Symposium on Digital Forensics and Security (ISDFS) (pp. 1-6). IEEE, 2023.

[12] M. H. Al Banna, K.A. Taher, M. S. Kaiser, M. Mahmud, M. S. Rahman, A. S. Hosen and G. H. Cho. “Application of artificial intelligence in predicting earthquakes: state-of-the-art and future challenges”. IEEE Access, vol. 8, pp. 192880-192923,2020.

[13] M. Kulkarni, C. Mulay, S. Marathe and P. Itankar. “Earthquake prediction using ML”. International Journal of Advance Research, Ideas and Innovations in Technology, vol. 7(3), 2021.

[14] S. Mukherjee, P. Gupta, Pagar,N. Varshney, and M. Chhetri, M. . Research Article A Novel Ensemble Earthquake Prediction Method (EEPM) by Combining Parameters and Precursors, 2022.

[15] S. M. Mousavi and G. C. Beroza.” Machine Learning in earthquake seismology”. Annual Review of Earth and Planetary Sciences, vol. 51(1), pp. 105-129,2023.

[16] G. C. Beroza, M. Segou, and S. Mostafa Mousavi. “Machine Learning and earthquake forecasting—next steps”. Nature communications, vol. 12(1), pp. 4761, 2021.

[17] R. J. Geller. "Earthquake prediction: a critical review". Geophysical Journal International, vol. 131(3), pp. 425-450,1997.
[18] H. Scholz. "The mechanics of earthquakes and faulting". Cambridge university press,2019.
[19] S. Stein and M. Wysession. "An introduction to seismology, earthquakes, and earth structure". John Wiley & Sons,2009.
[20] Geiß, P.A. Pelizari,M. Marconcini,W. Sengara, M. Edwards, T. Lakes and H. Taubenböck "Estimation of seismic building structural types using multi-sensor remote sensing and Machine Learning techniques". ISPRS journal of photogrammetry and remote sensing, vol. 104, 2015.
[21] H. Tyralis G. Papacharalampous and A. Langousis . "A brief review of RFs for water scientists and practitioners and their recent history in water resources". Water, vol. 11(5), pp. 910, 2019.
[22] R. Kumar and A. Saini. "Influence of Parabolic Irregularity, Inhomogeneity, Initial Stress, and Anisotropy on Love Wave Propagation". Mechanics of Solids, vol. 59(3), pp. 1443-1458, 2024.
[23] R. Kumar and A. Saini. "Effect of anisotropy, inhomogeneity and porosity on love wave propagation through fluid-saturated porous layers in irregular layered media". The European Physical Journal Plus, vol. 139(12), pp. 1-20, 2024.
[24] Saini and R. Kumar. "Effect of Rigidity and Parabolic Irregularity on Love Wave Propagation in Transversely Isotropic Fluid-Saturated Porous Layer Lying over a Nonhomogeneous Half-Space". Mechanics of Solids, vol.59(2), pp. 1094-1107, 2024.
[25] R. Kumar, A. Goyal and A. Saini. "Effect of inhomogeneity and triangular irregularity on propagation of shear waves in an anisotropic porous layer". Mechanics of Solids, vol. 59(2), pp. 1036-1049, 2024.
[26] Saini andR. K. Poonia. "Study of seismic wave propagation". In AIP Conference Proceedings (Vol. 2735, No. 1). AIP Publishing, 2023.
[27] Saini and R. Kumar. "Analysis of Love Waves in an Initially Stressed Transversely Isotropic Porous Layer Over Heterogeneous Half Space with Parabolic Irregularity". Journal of Vibration Engineering & Technologies, pp. 1-14, 2024.
[28] Dataset- Earthquake Data for Kyrgyzstan (2014-24).
[29] A.Cutler, D. R. Cutler and J.R. Stevens . "RFs". Ensemble Machine Learning: Methods and applications. pp. 157-175, 2012.
[30] N. S.Ahmed. "ML Models for Pavement Structural Condition Prediction": A Comparative Study of RF (RF) and extreme Gradient Boosting (XGBoost). Open Journal of Civil Engineering, vol. 14(4), pp. 570-586,2024.