RES MILITARIS
REVUE EUROPEENNE D ETUDES EUROPEAN JOURNAL OF MILITARY STUDIES
Social Science Journal

MACHINE LEARNING-BASED MAGNITUDE ESTIMATION FOR EARTHQUAKE EARLY WARNING SYSTEMS

[1]A Vinod Kumar,[2]Shaik Hussain Bi,[3]B. Benarji,[4]Pathan Subhani
[1]Assistant Professor,[2,3]Associate Professor,[4]Student
Department of CSE
G V R & S College of Engineering & Technology,Guntur,AP

ABSTRACT
To help earthquake early warning (EEW) systems make quick decisions, we build a random forest (RF) model for rapid earthquake localization. This system computes the differences in P-wave arrival timings between the first five stations to record an earthquake as a reference station (i.e., the first recording station). The RF model categorises these differential P-wave arrival times and station locations in order to determine the epicentral position. Using a Japanese earthquake catalogue, we train and evaluate the suggested algorithm. The Mean Absolute Error (MAE) of the RF model, which forecasts earthquake sites, is 2.88 km. Importantly, the suggested RF model can learn from little data—10% of the dataset—and a lot fewer recording stations—three—and yet get good results (MAE5 km). The approach provides a potent new tool for quick and precise source-location prediction in EEW since it is accurate, generalizable, and responsive.

1. INTRODUCTION

1.1 BRIEF INFORMATION

Earthquake hypocenter localization is crucial to seismology and is important for a number of applications, including tomography, source characterisation, and hazard evaluation. This emphasises the need of creating reliable seismic monitoring systems for pinpointing the timings and places of the event's genesis. A key but difficult job for creating seismic hazard reduction tools like earthquake early warning (EEW) systems is the quick and accurate classification of active earthquakes. Even though traditional techniques have been extensively used to develop EEW systems, it is still difficult to determine hypocenter locations in real-time because of the little data available during the early stages of earthquakes. Timeliness is one of the many important aspects of EEW, and more work is needed to further enhance the hypocenter location estimates using only data from the first few seismograph stations that are activated by the ground shaking and the first few seconds following the arrival of the P-wave.

In this paper, we present a differential P-wave arrival time and station location-based RF-based approach to find earthquakes. Only P wave arrival timings found at the first few stations are used in the proposed method. For EEW warnings to spread quickly, it must react quickly to earthquake first arrivals. By including the source-station locations into the RF model, our method implicitly takes the effect of the velocity structures into account. We assess the suggested method using a comprehensive Japanese seismic catalogue. Our test findings demonstrate that the RF model can effectively pinpoint earthquake areas with little data, which offers fresh insight on creating effective machine learning.

We use the suggested network to solve a Japanese earthquake detection issue. We base our findings on a comprehensive catalogue provided by the Japan Meteorological Agency, the National Research Institute for Earth Science and Disaster Resilience, and other organisations. Between January 1st, 2009, and November 11th, 2020, the Hi-net seismic network collected 2,235,159 regional seismic events, which are included in this extensive catalogue. We determine the position of the recording stations as well as the arrival times, magnitudes, depths, latitudes, and longitudes for each occurrence. We define qualified events as those meeting the following requirements for further analysis: P-wave arrivals are recorded at a minimum of five stations, the distance to the epicentre is less than 112 km, and the event magnitudes are larger than 0 ML. These standards provide reasonably accurate forecasts while facilitating quick responses to earthquakes. The final catalogue, with a total of 1,692,787 qualified occurrences, exhibits a wide variation of source characteristics and provides an excellent dataset for developing and evaluating the suggested method. The catalog's longitude ranges from 121.86 to 146.48 and its latitude from 23.42 to 46.22. Event depths vary from 0 to 440.78 km, and its magnitude varies from 0.10 ML to 7.59 ML. Note that according to several experiments we have conducted, the intermediate (80-300km) and deep (300+km) events in the training dataset only slightly impact the location accuracy.

1.2 PURPOSE

Regionalizing earthquake epicentres or predicting their specific hypocenter positions have both been accomplished using clustering techniques based on convolution neural networks. In the latter instance, the model for swarm event localisation is trained using three-component waveforms from several stations. In this paper, we present a differential P-wave arrival time and station location-based RF-based approach to find earthquakes. Only P-wave arrival timings found at the first few stations are used in the proposed method. Its quick reaction to earthquake first arrivals is essential for effectively spreading EEW notifications. By including the source-station locations into the RF model, our method implicitly takes the effect of the velocity structures into account. We assess the suggested method using a comprehensive Japanese seismic catalogue. Our test findings demonstrate that the RF model can effectively pinpoint earthquake areas with little input, which offers fresh insight on creating effective machine learning.

1.3 SCOPE

The next subject uses a similarly effective machine learning methodology. Contrary to the P-arrival timings in this study, the magnitude prediction is mostly reliant on P-wave amplitude and so needs more waveform data for a forecast. The suggested framework may enhance several currently used deep learning or deterministic methods-based efficient magnitude estimate techniques. We also look at the model's performance for bigger earthquakes since EEW systems are primarily concerned with larger earthquakes (e.g., those over M4). Another test is run, but this time we only choose catalogue events that are M4 or M5 or above. The MAEs for the M4 and M5 events for this group of data are 4.950 km and 4.271 km, respectively. Because there are fewer training samples for stronger earthquakes, the error is significantly greater for M4 and M5 events. Particularly, the majority of M4 incidents are found near to the coast (offshore), where station density is often low. There are a few possible solutions to the problem of inadequate training data sets, such as weighting the objective function, expanding the training dataset, or doing synthetic tests. We collect the difference in time between the origin time and the fifth station's P-wave arrival time in order to calculate the total time needed to determine the earthquake site. The P trip time to five neighbouring stations is typically within 5 s due to the average station spacing of 24 km. Additionally, there is likely a 1 second delay in data transmission and the selection algorithms require an extra 1 second after the first P arrival to validate a pick. Additionally, it takes the RF model 0.107 seconds to estimate the earthquake's location. Thus, 7.107 seconds is the anticipated total time needed by the suggested model. The deep learning method for earthquake localization, on the other hand, requires two seconds of data following the arrival time in addition to 0.179 seconds to pinpoint the earthquake. As a result, the deep learning strategy [4] takes 8.179 s in total. The location of stations in seismic monitoring might become more denser in the future, making the suggested approach more appropriate.

1.4 MOTIVATION

A series of observed waves (arrival times) and the locations of seismograph stations that are activated by ground shaking may be used to solve the localization issue. The recurrent neural network (RNN), one of several network designs, is capable of accurately extracting information from a series of input data, which makes it the best choice for managing a collection of seismic stations that are triggered sequentially following the seismic wave propagation patterns. This approach has been researched to enhance the effectiveness of real-time earthquake detection and source characteristic categorization. For earthquake monitoring, several machine learning-based methods have also been suggested. For the earthquake detection issue, comparisons of conventional machine learning techniques, such as closest neighbour, decision tree, and support vector machine, have also been done. The accuracy of these approaches may be impacted by a common problem in the aforementioned machine learning-based frameworks: the selection of input characteristics often necessitates expert knowledge. Regionalizing earthquake epicentres or predicting their specific hypocenter positions have both been accomplished using clustering techniques based on convolution neural networks.

2. LITERATURE SURVEY

The most crucial stage of the software development process is the literature review. Determine the time factor, economics, and corporate strength prior to building the tool. The following stages are to decide which operating system and language were utilised to construct the tool if these requirements have been met. Once the programmers begin creating the tool, they need a lot of outside assistance. This assistance was gathered from senior programmers, books, or websites. The aforementioned factors were taken into account before constructing the suggested system.

1) A smartphone seismic network that goes beyond earthquake early warning

V. A. I. Huvenne, T. P. Le Bas, and others, R. B. Wynn

Tens of thousands to hundreds of thousands of people are still killed and injured each year by large earthquakes that strike metropolitan areas, causing long-lasting social and economic catastrophes. The Earthquake Early notice (EEW) system gives seconds to minutes of notice, enabling people to relocate to safe areas and automate the slowing down and shutting down of transportation and other equipment. Only a few countries have conventional seismic and geodetic networks, which are used by the few EEW systems in operation worldwide. Traditional networks are significantly less common than smartphones, which include accelerometers that can also be used to detect earthquakes. We discuss the creation of a novel seismic system called MyShake that uses the sensors on individual or private smartphones to gather information and study earthquakes. We demonstrate that cellphones can capture magnitude 5 earthquakes at a distance of little more than 10 km, and we create an earthquake detection feature that can distinguish earthquakes from common tremors. Then, at a central point where a network detection algorithm verifies that an earthquake is occurring and instantaneously determines its position and magnitude, our proof-of-concept system gathers earthquake data. Then, a warning of impending earth shaking may be sent out using this information. MyShake may be used to improve EEW in areas with established networks and might be the sole EEW option in areas without them. The seismic waveforms captured might also be utilised to create quick microseism ograms, research building effects, and perhaps even image shallow ground structure and earthquake rupture kinematics.

2) Recurrent neural networks for intelligent real-time earthquake detection.

T. L. Chin, K. Y. Chen, and D. Y. Chen

One of the most seismically active regions in the world is Taiwan, which is situated where the Philippine Sea Plate and the Eurasian Plate converge. Around the island, devastating earthquakes have sometimes caused significant damage. Early earthquake warning (EEW) is crucial for preventing serious loss, and one of its most important functions is the quick and accurate identification of earthquakes. The commencement of the earthquake waves is often detected using criterion-based algorithms in conventional earthquake detection techniques. At the moment, those criteria's levels are often determined experimentally, which might lead to an excessive number of false alarms. Of course, false alerts might result in unnecessary fear and damage the system's confidence. In this article, a real-time EEW system is created using recurrent neural network (RNN) models. The created system is made to recognise when an earthquake event occurs and how long the P-wave and S-wave last. Using seismograms captured in Taiwan between 2016 and 2017, it was practised on and put through testing. According to the simulation findings, the suggested method performs better in terms of processing speed and detection accuracy than the conventional criterion-based schemes.

3) Develop detecting skills: Increasing earthquake detection precision

T. L. Chin, C. Y. Huang, S. H. Shen, and Y. C. Tsai are the authors.

High-speed computer networks are used by earthquake early warning systems to send earthquake information to population centres prior to the arrival of catastrophic earthquake waves. This little (10 s) lead time will enable emergency actions, such as shutting down gas pipeline valves, to be initiated in order to lessen the possibility of a catastrophe and fatalities. But the high incidence of false alarms in such a system comes at a high price in terms of lost services, unwarranted worry, and declining confidence of such a warning system. At the moment, the algorithm used to decide whether to provide a warning when an earthquake is about to occur is often based on experimentally selected characteristics and heuristically defined thresholds, and thus has a high false alarm rate. In this study, we tested the performance of three cutting-edge machine learning methods, including the K-nearest neighbour (KNN), classification tree, and support vector machine (SVM), versus a more conventional criterion-based approach. For these tests, we used seismic data gathered by an experimental strong motion detection network in Taiwan. We found that the machine learning methods display greater detection accuracy with a much lower false alarm rate.

3. SYSTEM ANALYSIS

3.1 EXISTING SYSTEM

To reduce seismic risks, earthquake early warning (EEW) systems are mandated to notify earthquake locations and magnitudes as soon as possible before the destructive S wave arrival. Instead of using seismic phase selections, deep learning approaches have the capacity to extract information about earthquake cause from whole seismic waveforms. With the goal of concurrently detecting earthquakes and estimating their source characteristics from continuous seismic waveform streams, we created a revolutionary deep learning EEW system. As soon as a small number of stations pick up earthquake signals, the system calculates the position and size of the quake. Meanwhile, by continuously collecting data, the system evolves its solutions. We use the technique to analyse the first week of aftershocks from the 2016 M 6.0 Central Apennines, Italy earthquake. As early as 4 s after the earliest P phase, it is possible to confidently predict the locations and magnitudes of earthquakes, with typical error ranges of 8.5-4.7 km and 0.30–0.27, respectively.

Disadvantages of Existing System

- To enhance the effectiveness of real-time earthquake detection and source categorization, a current system approach is not explored.
- Neither the regionalization of earthquake epicentres nor the accurate location prediction of their hypocenters have been achieved using clustering techniques based on convolution neural networks.

3.2 PROPOSED SYSTEM

Using the differential P-wave arrival timings and station locations, the system suggests an RF-based way to find earthquakes (Figure 1). Only Pwave arrival times found at the first few stations are used in the proposed method. For EEW warnings to spread quickly, it must react quickly to earthquake first arrivals. By including the source-station locations into the RF model, our method implicitly takes the effect of the velocity structures into account. The suggested system tests the proposed algorithm using a comprehensive Japanese seismic catalogue. Our test findings demonstrate that the RF model can effectively pinpoint earthquake areas with little input, which offers fresh insight on creating effective machine learning.

Advantages of Proposed System

- The number of stations plays a key role in determining the data accessibility and forecast precision. An increasing need for simultaneous recording at additional stations reduces the quantity of qualifying events since the suggested RF model depends on the arrival timings of P waves recorded at various stations.
- The locations of seismograph stations that are activated by ground trembling and a series of observed waves (arrival times) may be used to solve the localization issue. The recurrent neural network (RNN) is one kind of network design that is particularly good at accurately extracting information from a series of input data, making it the best choice for managing a set of seismic stations that are triggered sequentially in accordance with the routes taken by seismic waves as they propagate.

3.5 HARDWARE REQUIREMENTS

The physical computer resources, sometimes known as hardware, are the most typical set of specifications given out by any operating system or software programme. The following sections go into detail about the different hardware requirements.

- System : CORE i3 Processor.
- Hard Disk : 100 GB.
- RAM : 4 GB.

3.6 SOFTWARE REQUIREMENTS

Software requirements are concerned with specifying the software resources and prerequisites that must be installed on a computer to provide the best possible performance of a programme. These prerequisites must be installed individually before the programme can be installed since they are often not included in the software installation package.

- Operating system : Windows 7 Ultimate(min)
- Coding Language: Python
- Front-End : Python, Django-ORM
- Designing : HTML,CSS.
- Data Base : MySQL (WAMP Server).

4. SYSTEM DESIGN

4.1 SYSTEM ARCHITECTURE

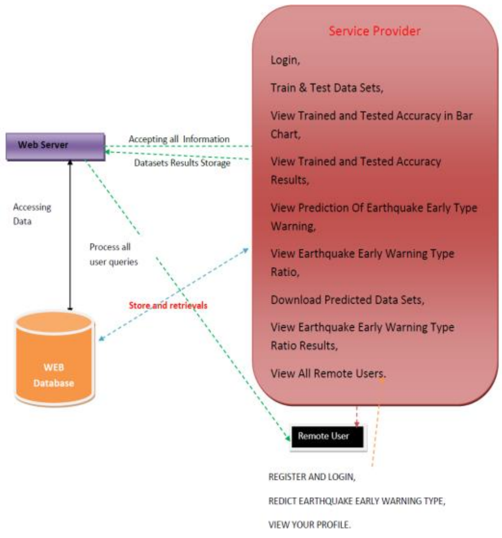


Fig: 4.1 System Architecture

4.2 MODULES

The step of implementation is when the theoretical design is translated into a programmatically-based approach. The application will be divided into a number of components at this point and then programmed for deployment. Python is used for the application's front end, while for the back end data base, Kaggle data was used. The following modules make up the bulk of the application. They are listed below:

Service Provider

The Service Provider must provide a valid user name and password to log in to this module. He can do certain actions such as log in, train and test data sets, and Check out the trained and tested accuracy in the bar chart. Results of Trained and Tested Accuracy, View Earthquake Early Warning Type Ratio, Download Predicted Data Sets, and View Earthquake Early Warning Type Prediction View All Remote Users and Earthquake Early Warning Type Ratio Results.

View and Authorize Users

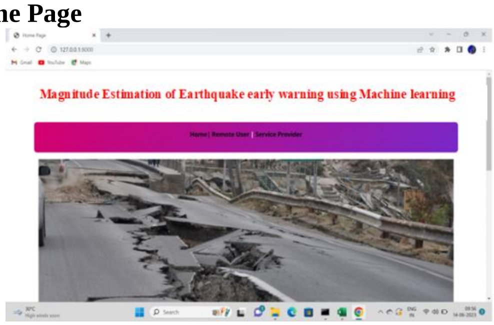
The list of people who have registered may be seen by the administrator in this module. This allows the administrator to access information about the user, including their name, email address, and home address.

Remote User

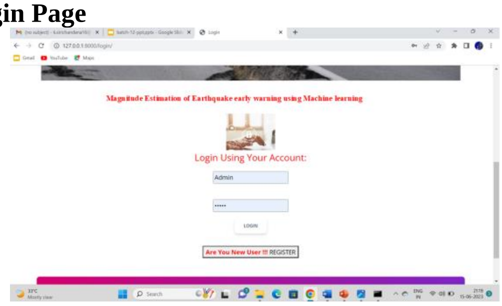
There are n numbers of users present in this module. Before doing any activities, the user should register. Once a user registers, the database will record their information. After successfully registering, he must log in using an authorised user name and password. After successfully logging in, the user may do a number of actions, including REGISTER AND LOGIN, REDICT EARTHQUAKEEARLY WARNING TYPE, and VIEW YOUR PROFILE.

5.4 OUTPUT SCREENS

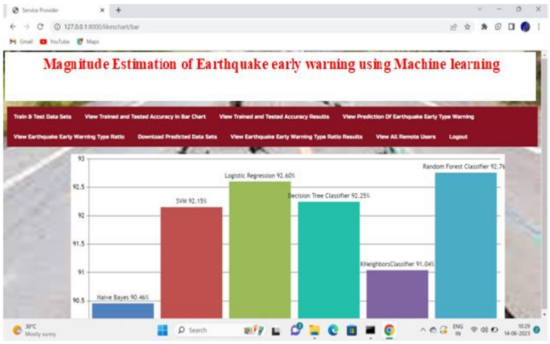
Home Page



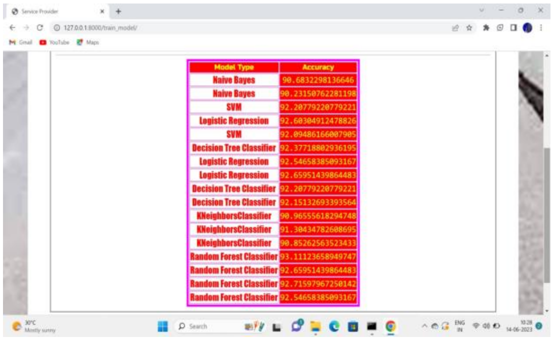
Login Page



Bar Chat

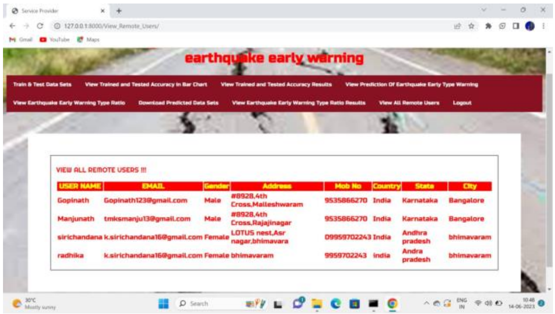


Trained and Test Accurate Results



Model Type	Accuracy
Naive Bayes	90.6832298136646
Naive Bayes	90.23150762281198
SVM	92.20779220779221
Logistic Regression	92.60304912478826
SVM	92.09486166007905
Decision Tree Classifier	92.37718802936195
Logistic Regression	92.54658385093167
Logistic Regression	92.65951439864483
Decision Tree Classifier	92.20779220779221
Decision Tree Classifier	92.15132693393564
KNeighborsClassifier	90.96555618294748
KNeighborsClassifier	91.30434782608695
KNeighborsClassifier	90.85262563523433
Random Forest Classifier	93.11236358949747
Random Forest Classifier	92.65951439864483
Random Forest Classifier	92.71597967250142
Random Forest Classifier	92.54658385093167

All Remote Users



7. CONCLUSION

We pinpoint the epicentre of the earthquake in real-time by comparing the arrival times of P-waves at several seismic sites. To solve this regression issue, random forest (RF) has been suggested, with its output being the difference in latitude and longitude between the earthquake and the seismic stations. Case studies in the Japanese seismic region show extremely promising results and suggest its immediate relevance. We collect data from seismic sensors in the area for all occurrences with at least five measurable P-wave arrival timings. We then created a machine learning model by separating the retrieved events into a training dataset and an evaluation dataset. Furthermore, the suggested technique only requires three seismic stations and 10% of the available dataset for training, but still achieves promising performance, demonstrating the adaptability of the proposed algorithm in real-time earthquake monitoring in more difficult regions. Despite the fact that the random forest technique has trouble training an appropriate model owing to the sparse distribution of various networks throughout the globe, one may employ a large number of synthetic datasets to make up for the dearth of ray pathways in a specific region due to a lack of catalogue and station dispersion.

8. REFERENCES

[1] Q. Kong, R. M. Allen, L. Schreier, and Y.-W. Kwon,"Myshake: A smartphone seismic network for earthquake early warning and beyond," Science advances, vol. 2, no. 2, p. e1501055, 2016.

[2] T.-L. Chin, K.-Y. Chen, D.-Y. Chen, and D.-E. Lin, "Intelligent real-time earthquake detection by recurrent neural networks," IEEE Transactions on Geoscience and Remote Sensing, vol. 58, no. 8, pp. 5440–5449, 2020.

[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 Transactions on Geoscience and Remote Sensing, vol. 57, no. 11, pp. 8867–8878, 2019.

[4] O. M. Saad, A. G. Hafez, and M. S. Soliman, "Deep learning approach for earthquake parameters classification in earthquake early warning system," IEEE Geoscience and Remote Sensing Letters, pp. 1–5, 2020.

[5] X. Zhang, J. Zhang, C. Yuan, S. Liu, Z. Chen, and W. Li, "Locating induced earthquakes with a network of seismic stations in oklahoma via a deep learning method," Scientific reports, vol. 10, no. 1, pp. 1–12, 2020.

[6] L. Breiman, "Random forests," Machine learning, vol. 45,no. 1, pp. 5–32, 2001.

[7] S. M. Mousavi, W. L. Ellsworth, W. Zhu, L. Y. Chuang, and G. C. Beroza, "Earthquake transformeran attentive deep-learning model for simultaneous earthquake detection

and phase picking," Nature Communications, vol. 11, no. 1, pp. 1–12, 2020.
[8] S. M. Mousavi and G. C. Beroza, "A Machine-Learning Approach for Earthquake Magnitude Estimation," Geophysical Research Letters, vol. 47, no. 1, p.e2019GL085976, 2020.