The Cavalcade of Immaculate and Disparate Algorithms for Detecting Distracted Earthquakes Employing Machine Learning.

Nazeer Shaik[1], Akshaya G[2], Manoj A[3], Aravind B[4], Deepthi T[5]

[1]Assistant Professor, Dept. of Computer Science, Srinivasa Ramanujan Institute of Technology, Anantapur, Andhra Pradesh, India - 515701.

[2,3,4,5]UG Student, Dept. of Computer Science of Engineering, Srinivasa Ramanujan Institute of Technology, Anantapur, Andhra Pradesh, India - 515701.

Abstract

An Earthquake is the sudden shaking of the surface of the earth resulting from sudden release of energy in the Lithosphere. Earthquake detection using machine learning is a promising approach to improving the speed and accuracy of earthquake detection. Machine learning algorithms use statistical models to analyze seismic data and identify patterns associated with earthquakes. The process of earthquake detection using machine learning involves collecting seismic data from seismometers and extracting relevant features from the data. The features extracted from the seismic data are used to train machine learning models such as decision tree, KNeighborsClassifier , Support Vector Machines , and Random Forest Classifier to identify patterns in the seismic waves associated with earthquakes. Once the models are trained, they can be used for earthquake detection.

Keywords: Machine Learning, Earthquake, Random Forest Classifier, Decision tree, Support Vector Machine, KNeighborsClassifier.

I. Introduction

Earthquakes generally occur because of plate tectonics. While it is important to understand the nonlinear dynamics of earthquake generation process. Earthquake detection is an essential area of research that deals with identifying and analysing seismic activity, which can cause significant destruction and loss of life. Traditional earthquake detection methods rely on seismometers, which can be expensive and not always reliable. However, with the advancements in machine learning, earthquake detection using machine learning algorithms has gained popularity. Machine learning algorithms can identify patterns in large amounts of seismic data and detect seismic activity in real-time. This has the potential to improve the speed and accuracy of earthquake detection, allowing for earlier warnings and more effective disaster response. In this context, machine learning algorithms are trained on seismic data and use various techniques to detect earthquakes.

However, the use of machine learning for earthquake detection also presents some challenges. One of the main challenges is the selection of appropriate features and the accuracy of the model. The success of the approach depends on the quality and quantity of seismic data and the ability to identify and extract relevant features from the data.

Overall, earthquake detection is a critical component of disaster preparedness, and the use of machine learning algorithms has the potential to improve the accuracy and speed of earthquake detection, providing a crucial step towards mitigating the impact of earthquakes on society.

II.Literature Survey

Ali G. Hafez, Ahmed Abdel Azim, M. Sami Soliman & Hideki Yayama [1] have proposed a technique called P-wave picking which is one of the important steps for earthquake parameter determination. Although manual inspection of P-wave arrival timing is the most accurate method for detection but, using automated algorithm is a must to facilitate this continuous task. This algorithm generates a daily report of all events recorded by this sub-network. This algorithm can detect very small events starting from microearthquakes due to the use of multiresolution analysis (MRA) of discrete wavelet transform (DWT). Results show a high rate of successful detections of 94.6% with low false alarm rate.

Lomax, A. Michelini and D. Josipovic's [2] have suggested an investigation of rapid earthquake characterization using single station waveforms and a convolutional neural network. Effective early-warning, response and information dissemination for earthquake sand tsunamis require rapid characterization of an earthquake's location, size and other parameters. This characterization is mainly provided by real-time seismogram analysis using established, rule-based, seismological procedures. With the advent of powerful machine learning tools to make predictions from large data sets.

Omar M. Saad, Ali G. Hafez, and M. Sami Soliman [3] has proposed Deep Learning Approach for Earthquake Parameters Classification in Earthquake Early Warning System. Magnitude determination of earthquakes is a mandatory step before an earthquake early warning (EEW) system sends an alarm. Beneficiary users of EEW systems depend on how far they are located from such strong events. Therefore, determining the locations of these shakes is an important issue for the tranquillity of citizens as well. The proposed algorithm depends on a convolutional neural network (CNN) which can extract significant features from waveforms that enabled the classifier to reach a robust performance in the required earthquake parameters. The classification accuracies of the suggested approach for magnitude, origin time, depth, and location are 93.67%, 89.55%, 92.54%, and 89.50%, respectively.

H Hang Zhang 1 2, Jun Zeng 1, Chunchi Ma 1 3, Tianbin Li 1, Yelin Deng 1, Tao Song [4] has proposed a Multi-Classification of Complex Microseismical Waveforms Using Convolutional Neural Network. In this study, a micro seismic multi-classification (MMC) model is proposed based on the short time Fourier transform (STFT) technology and convolutional neural network (CNN). The real and imaginary parts of the coefficients of micro seismic data are inputted to the proposed model to generate three classes of targets. micro seismic data recorded under different geological conditions are also tested to prove the generality of the model, and a micro seismic signal with $Mw \geq 0.2$ can be detected with a high accuracy. The proposed method has great potential to be extended to the study of exploration seismology and earthquakes.

Khawaja Asim, Abdul Basit, Francisco Martinez-Alvarez and Talat Iqbal [5] has detected Earthquake magnitude in Hindukush region using machine learning techniques.

In the research, four machine learning techniques including pattern recognition neural network, recurrent neural network, random forest, and linear programming boost ensemble classifier are separately applied to model relationships between calculated seismic parameters and future earthquake occurrences. Here, several performance measures can be done with parameters and accuracy can be estimated.

III. Proposed Work:

1.Random Forest Classifier

- A random forest is a machine learning technique that's used to solve regression and classification problems. It utilizes ensemble learning, which is a technique that combines many classifiers to provide solutions to complex problems.
- A random forest algorithm consists of many decision trees. The 'forest' generated by the random forest algorithm is trained through bagging or bootstrap aggregating. Bagging is an ensemble meta-algorithm that improves the accuracy of machine learning algorithms.
- The (random forest) algorithm establishes the outcome based on the predictions of the decision trees. It predicts by taking the average or mean of the output from various trees. Increasing the number of trees increases the precision of the outcome.
- A random forest eradicates the limitations of a decision tree algorithm. It reduces the over fitting of datasets and increases precision. It generates predictions without requiring many configurations in packages (like Scikit-learn).

Features of a Random Forest Algorithm:

- It's more accurate than the decision tree algorithm.
- It provides an effective way of handling missing data.
- It can produce a reasonable prediction without hyper-parameter tuning.
- It solves the issue of over fitting in decision trees.
- In every random forest tree, a subset of features is selected randomly at the node's splitting point.

2.Decision Tree

Overall, the decision tree algorithm can help improve the accuracy of earthquake detection by identifying the most important features for classification, allowing for the creation of more complex models that capture the nuances of the seismic data. This can ultimately lead to faster and more accurate detection of earthquakes, which is critical for early warning and response systems.

The decision tree algorithm is a method used to identify important features in earthquake detection using Artificial Neural Networks (ANN). It works by creating a tree-like structure that splits the seismic data into branches based on the most informative features, helping to distinguish between earthquake and non-earthquake events. This algorithm can lead to more accurate detection of earthquakes by identifying patterns in the seismic data that may be difficult for humans to recognize. By using both ANN and decision tree algorithms, more complex models can be developed, which can improve the accuracy of earthquake detection and provide faster warnings to people in affected areas.

Decision Tree is a Supervised learning technique that can be used for both classification and Regression problems, but mostly it is preferred for solving Classification problems. It is a tree-structured classifier, where internal nodes represent the features of a dataset, branches represent the decision rules, and each leaf node represents the outcome.

3.Support Vector Machine

Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. However, primarily, it is used for Classification problems in Machine Learning.

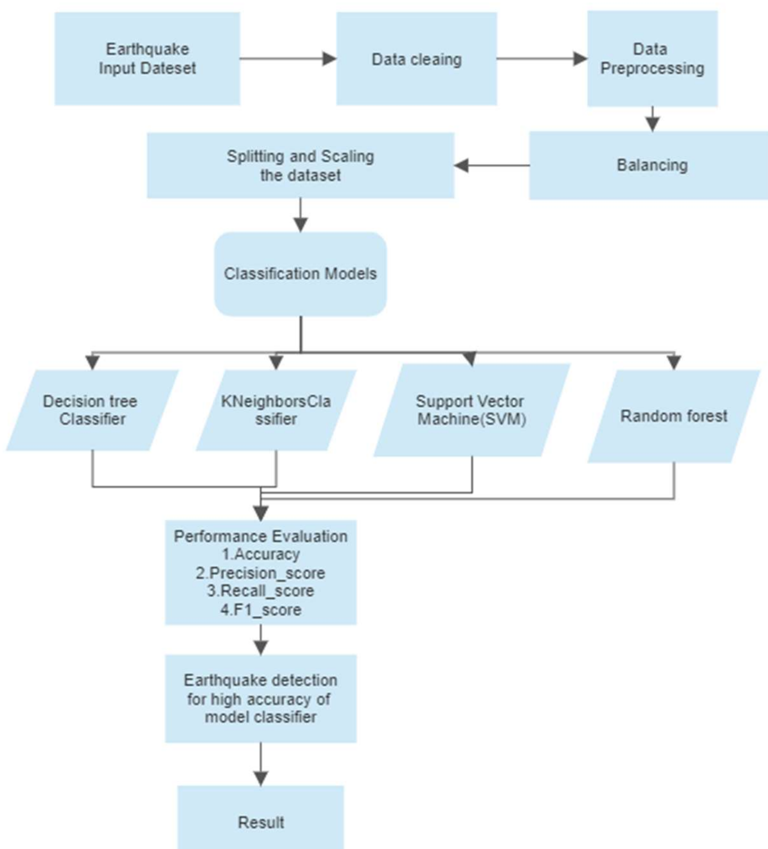
SVC works by finding the optimal boundary, or hyperplane, between two classes of data points, in this case, earthquake and non-earthquake events. The algorithm identifies the data points closest to the boundary, known as support vectors, and uses them to construct the boundary.

4.K-Nearest Neighbour

K-Nearest Neighbours (KNN) algorithm is a machine learning algorithm that can be used for various tasks, including classification and regression. In the context of earthquake detection, KNN can be used as a predictive model to identify seismic events based on their similarity to previously recorded earthquakes.

In earthquake detection, KNN can be used as a pattern recognition tool to identify seismic events that are similar to known earthquakes based on their location, magnitude, and other features. This can help in early warning systems and rapid response to earthquakes. For example, if a new seismic event occurs, KNN can be used to quickly identify if it is similar to known earthquakes in the area, and if so, issue an alert to nearby populations to take appropriate action.

IV. Workflow:



The Schematic Representation of Workflow Process

I.Methodology:

The main steps in earthquake detection using machine learning algorithms with a dataset from CSV are as follows:

- **Data Loading**: Load the earthquake dataset from a CSV file using a library such as Pandas in Python.
- **Data Pre-processing**: Pre-process the data to remove any missing values, handle categorical variables, and normalize the data. This may involve techniques such as imputation, one-hot encoding, and scaling.
- **Feature Selection**: Identify the most relevant features that can be used to detect earthquakes. This may involve statistical analysis and feature engineering techniques.
- **Model Selection**: Choose a suitable machine learning algorithm for the task of earthquake detection. This may involve evaluating multiple algorithms and selecting the one that performs the best.
- **Model Training**: Train the selected machine learning algorithm using the pre-processed data and the selected features.
- **Model Evaluation**: Evaluate the performance of the trained model using various metrics such as accuracy, precision, recall, and F1 score.

II.System Requirements

1. Hardware Requirements

MINIMUM (Required for Execution)		MY SYSTEM (Development)
System	Pentium IV 2.2 GHz	i3 Processor 5^{th} Gen
Hard Disk	20 Gb	500 Gb

Ram	1 Gb	4 Gb

2. **Software Requirements**

Operating System	Windows 10/11
Development Software	Python 3.10
Programming Language	Python
Domain	Machine Learning
Integrated Development Environment (IDE)	Visual Studio Code, Google Colab

V. Results And Analysis:

Following results were observed after executing the python code:

In this section, results are analysed using classification technique Decision tree, SVM, K-Neighbour's classifier and Random Forest Algorithms. The classification technique implementation was performed using Matplotlib Library. Here the dataset contains various attributes of earthquake, depth, magnitude, longitude, latitude, here we divided the data into two parts which consists of training data and other is validating data,

With the given training data, we trained a Decision tree, SVM model, KN classifier model and Random Forest model. Then using validating data, we validated the models and plotted a confusion matrix and found out the 93.07% accuracy through SVM model, 92.30% accuracy through K-Neighbours classifier model and 94.23% accuracy through Decision Tree classifier model and 97.69% accuracy for the Random Forest classifier.

By comparing the Decision tree SVM, Neighbour's classifier, Random Forest algorithms Random Forest gives the better accuracy rate of 97.

S. No	Algorithms	Accuracy	Precision	Recall	F1 Score
1	Decision Tree Classifier	94.23%	94.17%	94.23%	94.18
2	K-Neighbours Classifier	92.30%	93.43%	92.30%	92.38%
3	SVM	93.07%	94.53%	93.07%	93.23%
4	Random Forest Classifier	97.69%	97.73%	97.69%	97.69%

Table.1: The Comparisons of the above Algorithms on Various Factors

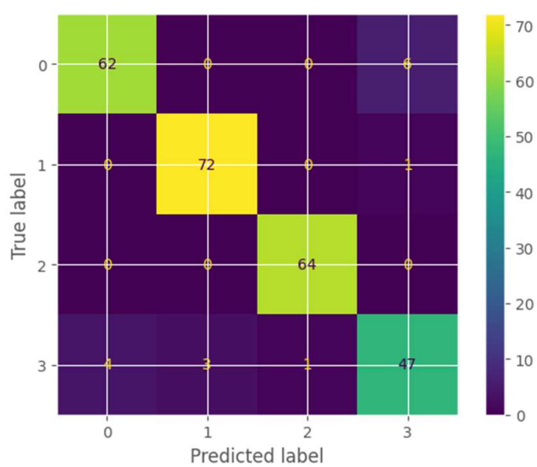


Fig.1: Confusion matrix for Decision Tree

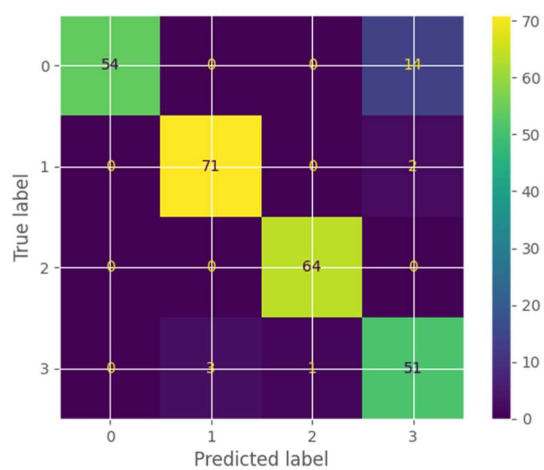
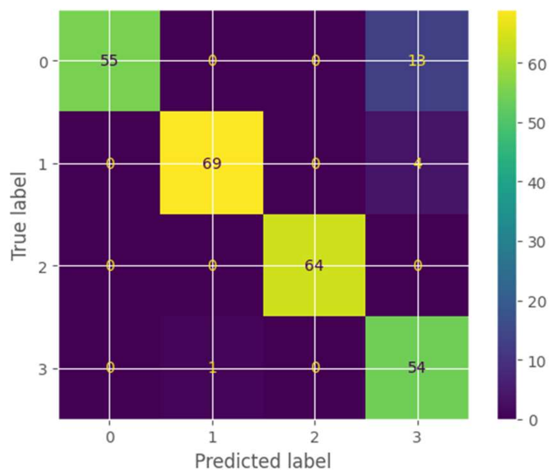


Fig.2: Confusion matrix for KNN classifier



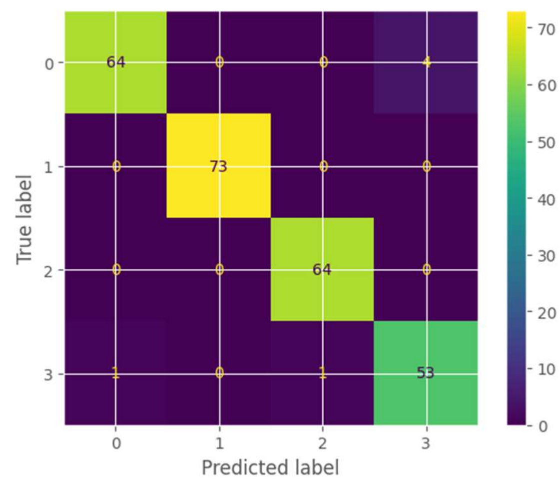


Fig.3: Confusion matrix for SVM

Fig.4: Confusion Matrix for Random Forest classifier

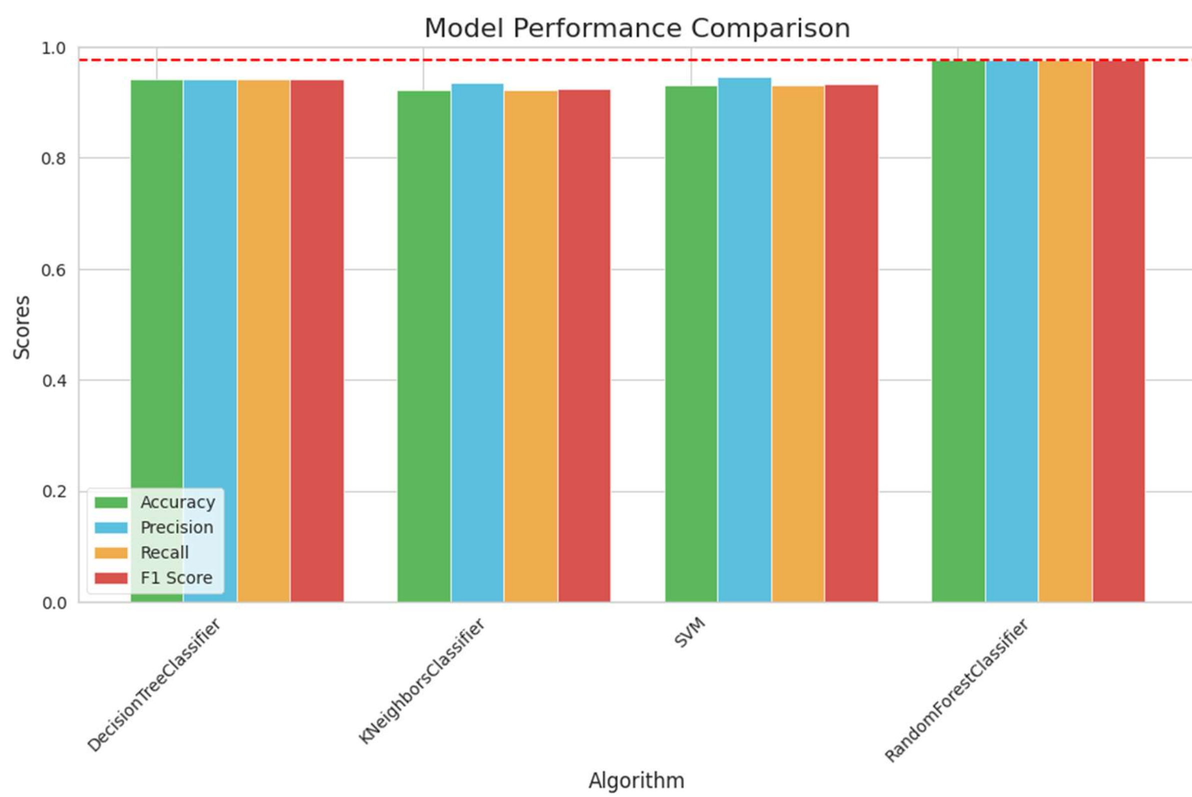


Fig.5: The Schematic Representation of Data Visualization of Corresponding Algorithms

Output:

It emphasizes that the use of earthquake detection and random forest algorithms to predict earthquake magnitudes. It suggests that using a random forest algorithm for earthquake detection gives high accuracy, and that prediction can be done using attributes such as longitude, latitude, depth, cdi, mmi, sig, dmin, nst, and gap.

By using machine learning technique random forest, we can accurately predict the magnitude of an earthquake using various attributes of the earthquake. This could be a valuable tool for disaster management agencies and first responders who need to quickly assess the potential impact of an earthquake and take appropriate action to minimize damage and loss of life.

```
enter the logitude 159.0270
enter the latitude -54.1325
enter depth 10.000
enter cdi 2
enter mmi 5
enter sig 733
enter dmin 0.371
enter nst 127
enter gap 45.0

low magnitude earthquakes
```

VI. Conclusion:

Based on the evaluation of four machine learning algorithms, namely Decision Tree Classifier, KNeighborsClassifier, SVC, and Random Forest Classifier, it can be concluded that all models have performed well in detecting earthquakes. Among all the algorithms, the Random Forest Classifier algorithm achieved the highest accuracy score of 0.976923, which is the best performance compared to the other models. The Decision Tree Classifier, KNeighborsClassifier, and SVC algorithms also performed well, with accuracy scores of

0.942308, 0.923077, and 0.930769, respectively. Moreover, all the models have performed well in terms of precision, recall, and F1-score. These performance metrics are essential for earthquake detection as it requires accurate and reliable predictions to avoid false positives and false negatives. In conclusion, the Random Forest Classifier algorithm is the best model for earthquake detection based on the given dataset. However, other algorithms such as Decision Tree Classifier, K-Neighbour's Classifier, and SVC can also be used as they have also shown good performance.

VII. References

1. A. G. Hafez, A. A. Azim, M. S. Soliman, and H. Yayama, "Realtime P-wavepicking for earthquake early warning system using discrete wavelet transform," NRIAG J. Astron. Geophys., vol. 9, no. 1, pp. 1–6, Jan. 2020.
2. A. Lomax, A. Michelini, and D. Josipovic's, "An investigation of rapid earthquake characterization using single-station waveforms and a convolutional neural network," Seismol. Res. Lett., vol. 90, no. 2A, pp. 517–529, Mar. 2019.
3. Omar M. Saad, Ali G. Hafez, and M. Sami Soliman, "Deep Learning Approach for Earthquake Parameters Classification in Earthquake Early Warning System" 1545- 598X © 2020 IEEE.
4. Hang Zhang 1 2, Jun Zeng 1, Chunchi Ma 1 3, Tianbin Li 1, Yelin Deng 1, Tao Song "Multi-Classification of Complex micro seismic Waveforms Using Convolutional Neural Network: A Case Study in Tunnel Engineering" Sensors 2021, 21(20), 6762.
5. Khawaja Asim, Abdul Basit-"Earthquake magnitude Detection in Hindukush region using Machine Learning techniques" pp.1-17, 2017.
6. Mousavi, S. M., Zhu, W., Shen, Y. and Beroza, G. C. (2019) CRED: A deep residual network of convolutional and recurrent units for earthquake signal detection. Scientific Reports 9, 10267.
7. Perol, T, Gharbi, M. and Denolle, M. (2018) Convolutional neural network for earthquake detection and location. Science Advances 4, 1700578.

8. Kuyuka, H. S. and Susumua, O. (2018) Real-time classification of earthquake using deep learning. Procedia Computer Science 140: 298– 305.
9. Wang, J. and Teng, T. (1995) Artificial neural network-based seismic detector. Bulletin of the Seismological Society of America 85: 308–319.
10. Madureira, G. and Ruano, A. E. (2009) A neural network seismic detector. IFAC Proceedings Volumes 42(19): 304–309.
11. Yoon, C. E., OReilly, O., Bergen, K. J. and Beroza, G. C. (2015) Earthquake detection through computationally efficient similarity search, Science Advances 1, e1501057.
12. Benz, H. M., McMahon, N. D., Aster, R. C., McNamara, D. E. and Harris, D. B. (2015) Hundreds of earthquakes per day: The 2014 Guthrie, Oklahoma, earthquake sequence, Seismological Research Letters 86: 1318–1325.