

Geological Earthquake Probability Prediction using data mining technique &
digital implementation
BY

Thaharim Khan
ID: 142-15-3457

Shazzad Hossain Kanon
ID: 142-15-3802

Sadia Afreen
ID: 142-15-4033

This Report Presented in Partial Fulfillment of the Requirements for the Degree of
Bachelor of Science in Computer Science and Engineering

Supervised By

Shah Mohammad Tanvir Siddiquee
Senior Lecturer,
Department of CSE,
Daffodil International University

Co-Supervised By

Ahmed Al Marouf
Lecturer,
Department of CSE,
Daffodil International University



DAFFODIL INTERNATIONAL UNIVERSITY
DHAKA, BANGLADESH
MAY 2018

APPROVAL

This Project/Internship titled “**Geological Earthquake Probability Prediction using data mining technique & digital implementation**”, submitted by **Thaharim Khan, Shazzad Hossain Kanon and Sadia Afreen** to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 7th May, 2018.

BOARD OF EXAMINERS

Dr. Syed Akhter Hossain

Chairman

Professor and Head

Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

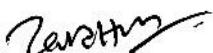


Dr. Sheak Rashed Haider Noori

Internal Examiner

Associate Professor

Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University



Md. Zahid Hasan

Internal Examiner

Assistant Professor

Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University



Dr. Mohammad Shorif Uddin

External Examiner

Professor and Chairman

Department of Computer Science and Engineering
Jahangirnagar University

DECLARATION

We hereby declare that, this project has been done by us under the supervision of **Shah Mohammad Tanvir Siddique** Daffodil International University. I also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

SUPERVISED BY:



Shah Mohammad Tanvir Siddique

Senior Lecturer

Department of CSE

Daffodil International University

CO-SUPERVISED BY:

Ahmed Al Marouf

Lecturer

Department of CSE

Daffodil International University

SUBMITTED BY:

Thaharim Khan

ID: 142-15-3457

Department of CSE

Daffodil International University

Shazzad Hossain Kanon

ID: 142-15-3802

Department of CSE

Daffodil International University

Sadia Afreen

ID: 142-15-4033

Department of CSE

Daffodil International University

ACKNOWLEDGEMENT

In the very beginning, we would like to convey our gratitude and thankfulness to the almighty for giving us the proficiency and capability along with all the blessings by virtue of which, it has been possible to implement this thesis and execute all the stages as per the planning efficiently.

In addition to this, it was an honor for us to have Shah Mohammad Tanvir Siddique, Senior Lecturer, Department of CSE, Daffodil International University and Ahmed Al Marouf, Lecturer, Department of CSE, Daffodil International University as our mentor for their constant support during my study. They have inspired us a lot to work in this project, his motivation encouraged us extremely to our project.

Besides, we really appreciate and would like to express our cordial gratitude to Prof. Dr. Syed Akhter Hossain, professor and head, Department of CSE, for his kind help and constant collaboration to successfully execute our thesis and also to other faculty member and staff of CSE department of Daffodil International University.

We would like to thank the authority of Daffodil International University for providing us a good environment and facilities such as computer laboratory to complete this project with software which we need during process.

We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work. Furthermore, we would like to thank to the parents who never doubted, constantly supported, and always loved us, to our team who have never any single problem to contribute to this project and all other of our friends and classmates at Daffodil International University who have supplementary support to my study efforts

ABSTRACT

Environment issue is a major ordeal for data mining. Be that as it may, more often than not here data mining utilized, for example, prescient information like land slide, seismic tremors, spatial information and so on. In our life we are centered on the spatial information, for example, logical information. Learning revelation in databases (KDD) has been characterized as the non-paltry procedure of finding substantial, novel, conceivably valuable and at last reasonable examples from information. A Database-Oriented Framework have presented for spatial information mining which depends on the ideas of neighborhood diagrams and ways and numerous utilizations of DBMS and data mining calculations. This research concentrated on geological information and spatial information by growing new systems for point design investigation, forecast in space– time information, and examination of moving article information. In a short we can decide the plausible time of fiasco like quake by estimating the circumstance of the plates utilizing topographical information, spatial information through our calculation. A clarified reference list is incorporated for ease in finding other helpful direction.

TABLE OF CONTENTS

Contents	Page
Acknowledgement	iii
Abstract	iv
List of Figures	vii
Chapter 1: Introduction	1-4
1.1 Introduction	1
1.2 Motivation	1
1.3 Rationale of the Study	2
1.4 Research Questions	2
1.5 Expected Output	3
1.6 Report Layout	4
Chapter 2: Background	5-7
2.1 Introduction	5
2.2 Related Works	5
2.3 Research Summary	6
2.4 Scope of the Problem	7
2.5 Challenges	7
Chapter 3: Research Methodology	8-13
3.1 Introduction	8
3.2 Research Subject and Instrumentation	10
3.3 Data Collection Procedure	10
3.4 Statistical Analysis	11
3.5 Implementation Requirements	13

Chapter 4: Experimental Results and Discussion	14-22
4.1 Introduction	14
4.2 Experimental Results	14
4.3 Descriptive Analysis	18
4.4 Summary	21
Chapter 5: Summary, Conclusion, Recommendation and Implication for Future Research	23-24
5.1 Summary of the Study	23
5.2 Conclusions	23
5.3 Recommendations	24
5.4 Implication for Further Study	24
References	25
References Appendices	
Appendix A: Research Reflection	26
Appendix B: Related Issues	26
Plagiarism Checker Screenshot	27

LIST OF FIGURE

List	page
Figure 2.2.1 Earthquake prediction	06
Figure 3.3.1 Data collection procedure	10
Figure 3.4.1 Statistical overview in weka	11
Figure 4.2.1 Data set	15
Figure 4.2.2 Training data	15
Figure 4.2.3 Training data	16
Figure 4.2.4 Training data	17
Figure 4.2.5 Training data	17
Figure 4.3.1 Android application implementation	19
Figure 4.3.1 Android application implementation	20
Figure 4.4.1 Magnitude vs Predicted magnitude	21

CHAPTER 01

INTRODUCTION

1.1 Introduction

Earthquake prediction is a big deal for our environment. We can save many of our resources if we can be informed about earthquake before

This research focused on geographical data and spatial data by developing new techniques for point pattern analysis, prediction in space–time data, and analysis of moving object data.

In a short we can determine the probable time of earthquake by measuring the situation of the plates using geographical data, spatial data. An annotated reference list is included for ease in finding other useful guidance.

1.2 Motivation

An earthquake cause many damages as we are not informed about it before so we cannot reduce the destruction .This research of predication is about earthquake so that we can get some information about it and take some precaution.

Prediction is only possible with data mining. And it never give us the result about 100% but with it we can get only probable result. People are before never get ready about it but now they can predict about it. Weather it happens or not.

With this methodology the office that provide information about weather get information about it before and they can inform people earlier.

1.3 Objectives

- To deal with data mining and knowledge base discovery database.
- Create strong connection between IOT and Data mining.
- Reduce the disastrous effect by getting information about disaster before.
- Specify a privilege for every people lead a better life and reduce the after effect of disaster.
Which may help the national economy.

1.4 Rationale of the Study

Data mining is used for predication type of work. For this in our research we are dealing with this. All type of prediction is become easier with this. This never give us 100% result but we can guess about something with this type of prediction. The surface of the Earth is in continuous slow motion. This is plate tectonics--the motion of immense rigid plates at the surface of the Earth in response to flow of rock within the Earth. The plates cover the entire surface of the globe. Since they are all moving they rub against each other in some places, sink beneath each other in others or spread apart from each other. At such places the motion isn't smooth-the plates are stuck together at the edges but the rest of each plate is continuing to move, so the rocks along the edges are distorted. As the motion continues, the strain builds up to the point where the rock cannot withstand any more bending. With a lurch, the rock breaks and the two sides move. An earthquake is the shaking that radiates out from the breaking rock.

There is many reason behind this shaking of rock. When there is amount of combustible elements under the soil then they are mainly release heat when they are rubbing with each other. When this heat come continuously to the rock then at one time the rock recede from its place and causes shaking.

Temperature is another reason behind this Earthquake. Before earthquake a temperature release from surface which is almost fusty .There is logic behind this, when you rubbing your two hands with one another for few minutes after that you can feel a temperature in your hand and observe a

little shake when stop rubbing . Whenever the Rocks are shaking they release a temperature which come out from the surface and after that we feel the shaking.

Another reason is the nearest place from the epicenter. The place which is close to the epicenter is one of the targeted place which is can be affected by earthquake. Epicenter is this type of place from where earthquake originated

1.5 Expected Outcome

This is a predication type work with data mining. And with using of IOT we can be informed about it before not exactly but get a probable idea about Earthquake. Data mining is the process of sorting through large data sets to identify patterns and establish relationships to solve problems through data analysis. Data mining tools allow enterprises to predict future trends. In our project Data mining is base, without it all type of prediction is almost impossible .The Internet of Things (IOT) is a system of interrelated computing devices, mechanical and digital machines, objects, animals or people that are provided with unique identifiers and the ability to transfer data over a network without requiring human-to-human or human-to-computer interaction. In other work like android apps is design in the platform of android as android phone is available very much so everyone can very easily get about any types of very easily. So, in the expected outcome will be like an android apps from where everyone can get about information about the probable time of earthquake and get ready for defending the disaster.

1.6 Report layout

Here, in this paper, we have tried to divide this entire paper into five segments for the precise implementation and augmented explanation to make the concept more convenient in term of understanding. The initial segment has been decorated with the basic understanding and rationale of the study to portrait the expected outcome and the motivation behind the concept to assist comprehending easily. Later the second segment or chapter has been made with the Background for this concept and the summary of our research has been highlighted along with other relevant researches to get assistance from. Also the challenges and the field of errors faced during the research and implementation has also been added. The third one is the Research Methodology where we have tried to express the idea about the methodology of the research work and the stages we have needed to go through including Instrumentation for the researching work, stages of Data Collection, Procedure for Data Preprocessing and Post Processing, Statistical analysis and the requirements for the execution. In the fourth chapter, the ultimate implementation of the experiments and extensive discussions has been made regarding the methodologies for the comprehensive possibilities to finish the task successfully. And finally the gist of our research work and all the studies regarding the earthquake and the implication for further study has been displayed with precise and legitimate explanation to utilize the study and technique for further usability to enrich the earthquake predictability and diminish the bad effects of this unpredictable calamity to a great extent.

CHAPTER 02

BACKGROUND

2.1 Introduction

Earthquake is one of the most horrible disaster for human life. A shaking of few minute. Causes a huge harm. So if this can be predicted it will save many of life. One more .Research happen on it. But they only find only 3 matters for predicting this hazard. But our goal is to find some more reasons and those type of reason with which it can be possible to predict about the upcoming earthquakes pretty efficiently.

2.2 Related Works

This type of prediction has been done before by one team but our target is to find better result than them. On their paper they use 3 factors these are

- Ground water levels
- Chemical changes in Ground water
- Radon Gas in Ground water wells.

On this 3 factors they prepare their data file and ready report like.

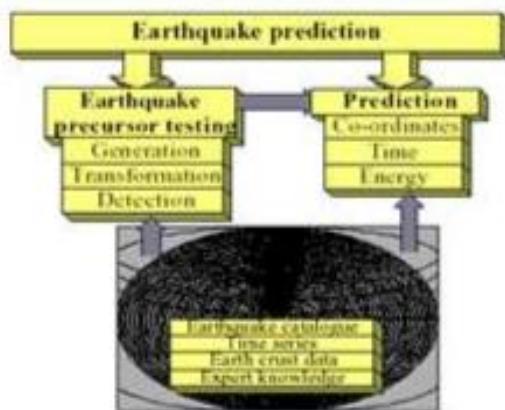


Figure2.2.1: Earthquake prediction

2.3 Research Summary

This research features the data mining methods connected to mine for surface changes after some time (e.g. Seismic tremor break). The data mining systems help analysts to anticipate the adjustments in the force of volcanoes. This paper utilizes prescient factual models that can be connected to zones, for example, seismic action, the spreading of flame. The essential issue in this class of frameworks is undetectable flow as for seismic tremors. The space-time designs related with time, area and extent of the sudden occasions from the power limit are detectable. This paper features the perceptible space time tremor designs from undetectable elements utilizing information mining methods, design acknowledgment and gathering anticipating. Along these lines this paper gives knowledge on how data mining can be connected in finding the outcomes of seismic tremors and subsequently alarming the general population.

2.4 Scope of the Problem

On previous work they are doing the work with simply the data which cannot be identify or collected by anyone they only give reason and use only seismic data with which they predict the total process. And using all the data mining technique, with this they can get result about 50%. In our project we use normal temperature and severe growing temperature which we termed as one of the most easily predictable factor. Combustible elements are another reason for ground shaking we are taking this values of our targeted areas and use them in various data mining algorithm. Distance from epicenter is also considered in our project. Epicenter helps us to know how long it will take for next. All this data help to get a better result than before as all the data are mined well and process by the data mining algorithms.

2.5 Challenges

This research is related with geological work so our main challenges that we faced during the time of data collection. We are targeted 10 areas and tried to analysis on these areas data. But these data are not available in internet, not only internet we cannot find a complete data file from anywhere. So we have to collect data separately for each attribute. There is chance of not getting the original data but we careful about it and tried to collect all the data from trusted source, so that we can get a better output from data mining.

CHAPTER 03

RESEARCH METHODOLOGY

3.1 Introduction

The various data mining techniques are-

- 1) Statistics
- 2) Clustering
- 3) Visualization
- 4) Association
- 5) Classification & Prediction
- 6) Outlier analysis
- 7) Trend and evolution analysis

These techniques are used to get perfect prediction result. All these tools are used for data. Which is collected for prediction.

Statistics

- Data cleansing i.e. the removal of erroneous or irrelevant data known as outliers.
- EDA Exploratory data analysis e.g. frequency counts histograms.
- Attribute redefinition e.g. bodies mass index.
- Data analysis is a measure of association and their relationships between attributes Interestingness of rules, classification, prediction etc.

Visualization

Enhances EDA, make patterns visible in different views.

Clustering (cluster analysis)

Clustering is a process of grouping similar data. The data which are not part of clustering are called as outliers. How to cluster in different conditions-

- Class label is unknown: Group related data to form new classes, e.g., cluster houses to find distribution patterns.
- Clustering based on the principle: maximizing the intra-class similarity and minimizing the interclass similarity. It provides subgroups of population for further analysis or action very important when dealing with large databases.

Association (correlation and causality)

- Mining Classification and Prediction association rules finds the interesting correlation relationship among large databases.
- Finding models (functions) that describe and distinguish classes or concepts for future prediction e.g., classify countries based on climate, or classify cars based on gas mileage.
- Presentation: decision-tree, classification rule, neural network.
- Prediction: Predict some unknown or missing numerical values.

Outlier analysis

A data object that is irrelevant to general behavior of the data it can be considered as an exception but is quite useful in fraud detection in rare events analysis.

Trend and evolution analysis

- Trend and deviation: regression analysis.
- Sequential pattern mining, periodicity analysis.
- Similarity-based analysis.

- Papers main intention is on clustering & visualization technique for predicting the earthquakes.

3.2 Research Subject and Instrumentation

Our project is predictive and our goal is to save human life and resources with the help of some Predictive analysis or methodology. Various type of data are used for this prediction like latitude, longitude, distance from epicenter, combustible elements under the surface, population density. These data are related with earthquake. We are just predicting and trying to give a probable answer about 60-70% correct.

Instrumentation

- Weka 3.8
- Microsoft office excel
- Android Studio

3.3 Data Collection Procedure

For Earthquake prediction we need geological data which we cannot find from anywhere. We have to collect all that data set from internet. Following some procedure-

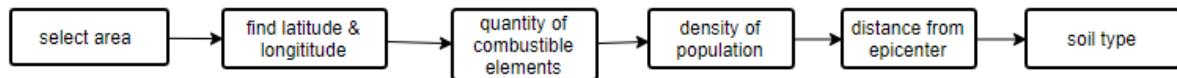


Figure 3.3.1: Data collection procedure

Data collection methods

- Interviews
- Open observations
- Record reviews
- Diaries

Instrument to collect data

- Observations
- Interviews
- Verbal reporting
- Questionnaires
- Tests

3.4 Statistical Analysis

Correlation coefficient	0.0202
Mean absolute error	1.3114
Root mean squared error	1.5275
Relative absolute error	105.8879 %
Root relative squared error	106.382 %
Total Number of Instances	75

Figure 3.4.1: Statistical overview in weka

The data sets which we used can give us the results with all these criteria.

Correlation coefficient

Correlation coefficient formulas are used to find how strong a relationship is between data. The Formulas return a value between -1 and 1, where:

- Indicates a strong positive relationship.
- -1 indicates a strong negative relationship.

A result of zero indicates no relationship at all

Mean Absolute Error

Mean Absolute Error is a model evaluation metric used with regression models. The mean absolute error of a model with respect to a test set is the mean of the absolute values of the individual prediction errors on over all instances in the test set. Each prediction error is the difference between the true value and the predicted value for the instance.

Root Mean Square Error (RMSE)

Root Mean Square Error (RMSE) is the standard deviation of the residuals (prediction errors). Residuals are a measure of how far from the regression line data points are; RMSE is a measure of how spread out these residuals are. In other words, it tells you how concentrated the data is around the line of best fit. Root mean square error is commonly used in climatology, forecasting, and regression analysis to verify experimental results.

Root relative squared error

The root relative squared error is relative to what it would have been if a simple predictor had been used. More specifically, this simple predictor is just the average of the actual values. Thus, the relative squared error takes the total squared error and normalizes it by dividing by the total squared error of the simple predictor. By taking the square root of the relative squared error one reduces the error to the same dimensions as the quantity being predicted.

Relative absolute error

The relative absolute error is very similar to the relative squared error in the sense that it is also relative to a simple predictor, which is just the average of the actual values. In this case, though, the error is just the total absolute error instead of the total squared error. Thus, the relative absolute error takes the total absolute error and normalizes it by dividing by the total absolute error of the simple predictor.

3.5 Implementation Requirements

For implementing this project we use

- Weka
- Android.

Now-a-days android is most Usable OS and easily understandable for everyone. At first we use weka for creating a trained data set which we collect for our research. We can get also a data model from weka. With this trained data and the model we are trying to implement an app where we give some seismic area to user. These areas are arranged alphabetically. When user click any area then that area identified through Latitude and longitude and after that comparing with the other attributes a there create a pattern. When pattern is ready then this pattern is matched with the model which is created by the trained data set. Finally a new model create through these two data set which we use for prediction. The final data set is used for prediction which gives us the value of magnitude. With which we can find either there is a probability of an earthquake or not.

CHAPTER 04

EXPERIMENTAL RESULT AND DISCUSSION

4.1 Introduction

The result that we get from the data set is mainly a magnitude. This magnitude tells us about the intensity of an Earthquake. We can also convert the magnitude for prediction. We get the magnitude from the attributes which we used in our data set. And all the attributes are related as all of them are geological attribute. Whenever we identify a place from lat-long then all that information we want can easily get from google. Because with lat-long we get the exact location. After that the attributes we used such as the soil type, density of population, combustible elements all are varying from place to place. The main attribute is distance from tectonic plate .For using this attribute we need to compare all the previous data with all of our attributes and get a model of data which we can use for prediction known as magnitude. Finally with this magnitude we can determine about the probable time of earthquake.

4.2 Experimental Result

Our result is very simple just to get a magnitude for a place which tells us about the Probability. We need to train our data set for this experiment.

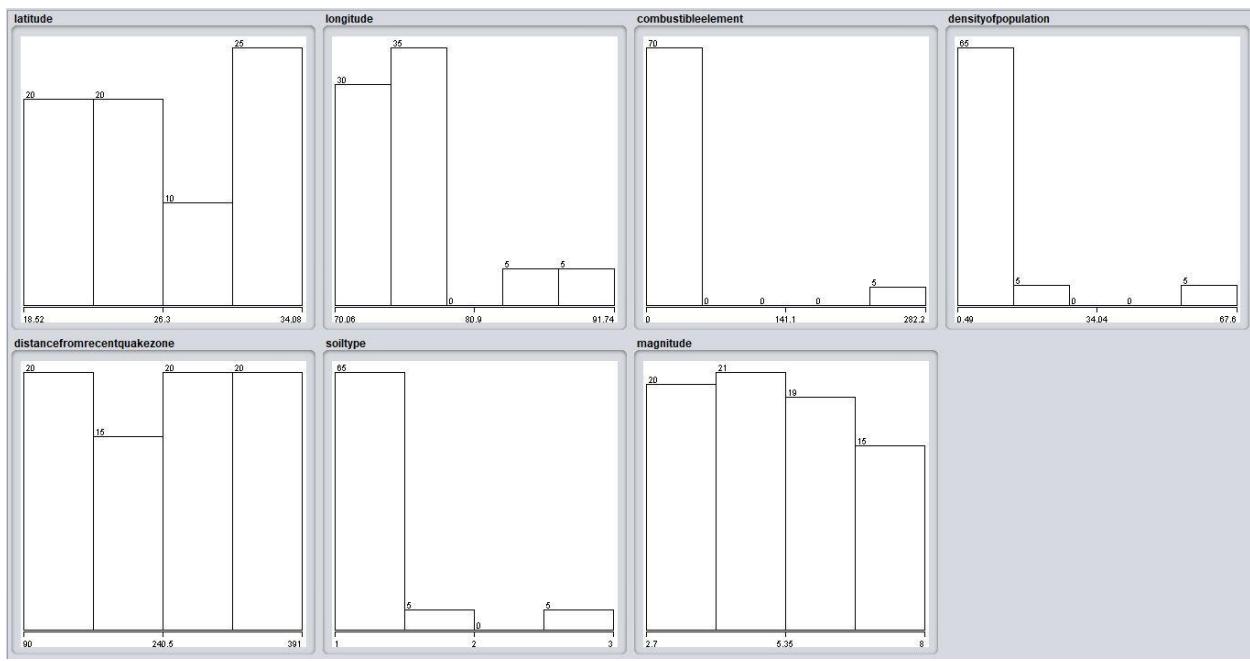


Figure 4.2.1: Data set

Here in figure 4.2.1 we take 15 places into our consideration and find all the attributes of those places and finally get the magnitude. This is the main terminology which give us the data model for prediction. And we can predict with this magnitude only.



Figure 4.2.2: Training data set

Here in figure 4.2.2 we trained our data of distances from tectonic plate. We consider that distance to x-axis and get magnitude against it from y-axis.



Figure 4.2.3: Training data set

Here in Figure 4.2.3 we trained our data density of population and get magnitude against it. Here the density is in X-axis we get magnitude from Y-axis

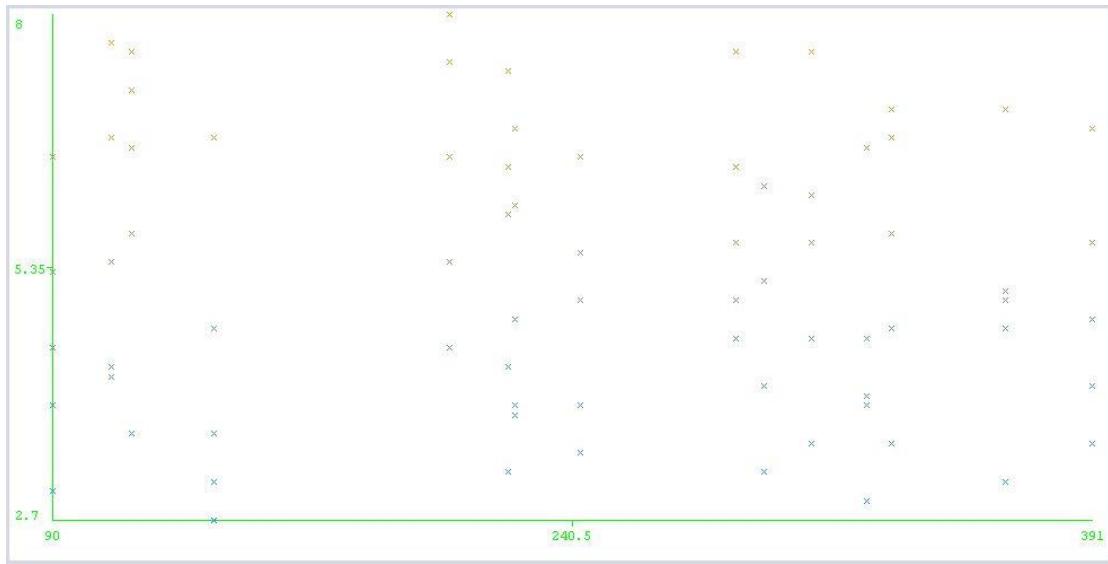


Figure 4.2.4: Training data set

Here in figure 4.2.4 we trained our data of magnesium and get magnitude against it. Here the value of magnesium is in X-axis we get magnitude from Y-axis.

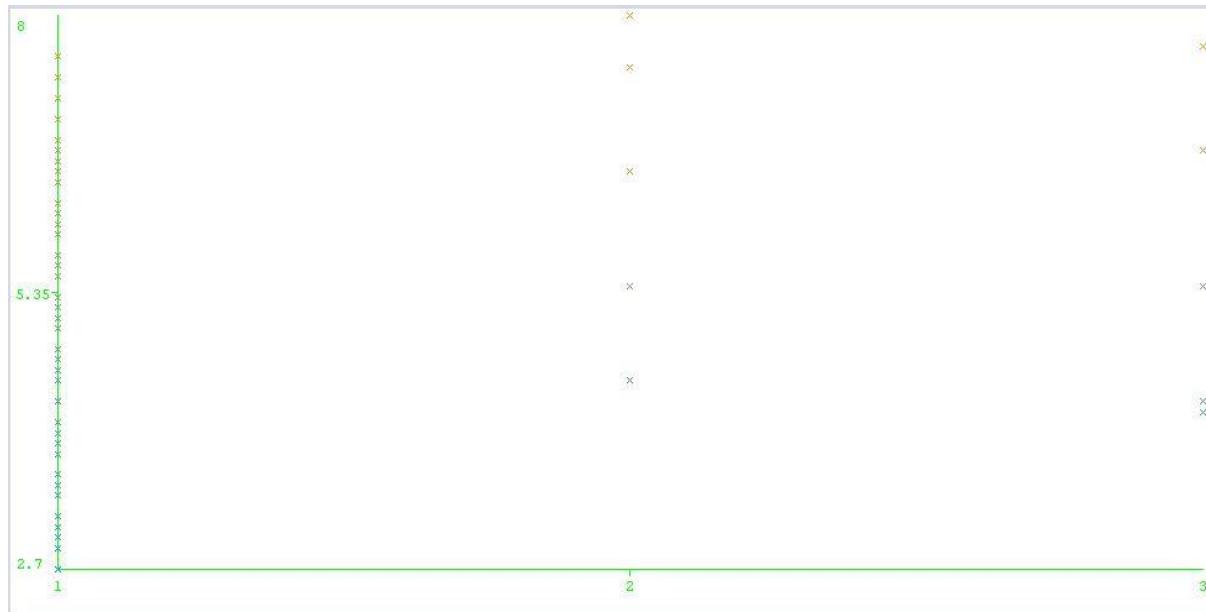


Figure 4.2.5: Training data set

Here in figure 4.2.5 we trained our data of distance from tectonic plate against the value of density of population. Here the value of distance is in X-axis and value of density in Y-axis.

4.3 Descriptive analysis:

We are using for our data set 6 attribute these are-

- Latitude
- Longitude
- Quantity of Magnesium (combustible element)
- Density of population
- Distance from nearest tectonic plate
- Soil type

Here all the attributes are related. Latitude and Longitude are used to get the exact location of that area which we choose, with this we can also get exact situation of the tectonic plate which's movement is the main reason for earthquake.

Combustible elements are also responsible as they are rubbing with each other and create a huge temperature which is also responsible for shaking rock. With the Latitude and Longitude we get that information about where the fault is and what is that current situation.

Population also create a pressure on surface. When there is a pressure on the layer of the surface then there is a probable chance of getting the combustible elements come closer and create a temperature. Which is the reason behind the shaking of rock.

We measure distance of an area from nearest tectonic plate. Because where there is a plate there is most probable chance for earthquake .the value of the distance create an impact on earthquake.

Soil type is only an attribute for which we used non-numeric value. Because soil type can tells us how much combustible elements can it preserve or not.

We assume a number against the soil type for our calculation. The types are –

- alluvial-1
- loamy-4
- clay-1(both clay and alluvial are same type)

- black soil-3
- lateritic-2

All these values are correlated and create a magnitude which tells us what is the Probability of earthquake.

From here we take all the data set and try to implement in Android there we just use all of our data and try to find the probability. There we give those place which we use into our data set. Whenever we select any place then the latitude and the longitude is given to the processing part and in our data file we also keep latitude and the longitude, and others attribute. After calculating with others attribute we get the probable magnitude of that place. Here we mention some screenshots for clarify.

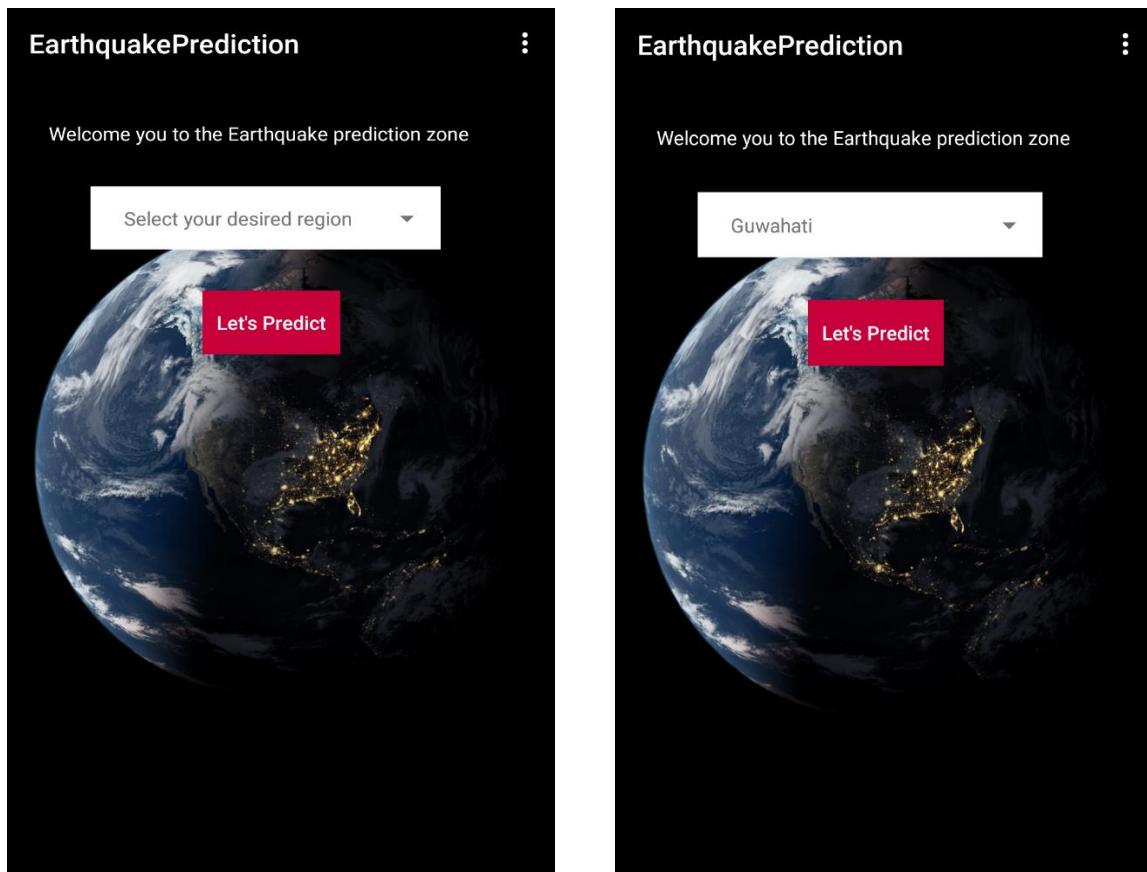


Figure 4.3.1: Android application implementation

This is the first page of our app. There is an option of selecting region there is only those region which we use in our data set. Our data file contain the others 6 attribute according to their name.

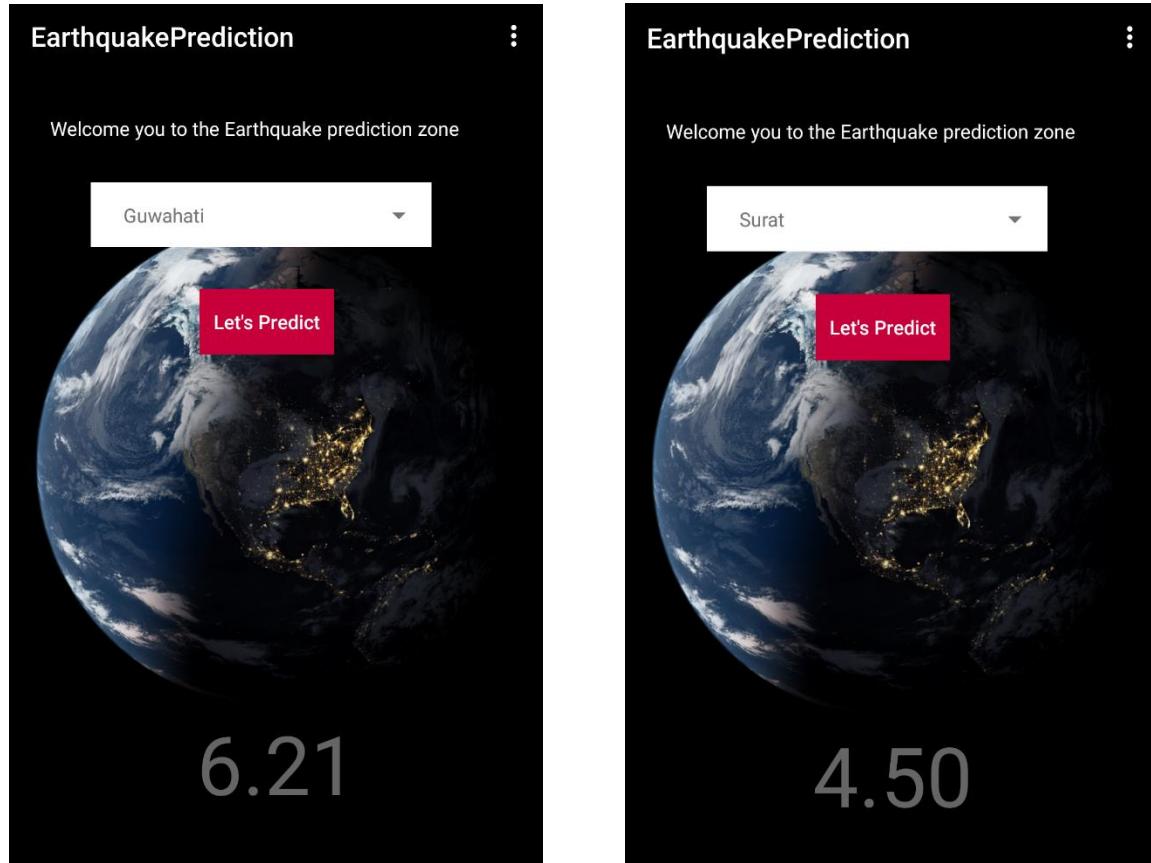


Figure 4.3.2: Android application implementation

Here in figure 4.3.2 Jamnagar is selected. Jamnagar is already into our data file. Jamnagar lies between 22.13° N and 69.42° E in the region of Gujrat popularity referred to as Saurashtra. The Gulf of Kutch lies to the north of Jamnagar district and the Arabian Sea on the west. In the east lies the Rajkot district and the Junagadh district is to the south. The total area of Jamnagar city is 122 sq. kilometer with total population of 600411 as per the 2011 census.

In this figure 4.3.3 we see the magnitude of earthquake. We get this probability calculating all the attribute. Magnitude and Intensity measure different characteristics of earthquakes. Magnitude measures the energy released at the source of the earthquake. Magnitude is determined from

measurements on seismographs. Intensity measures the strength of shaking produced by the earthquake at a certain location. Intensity is determined from effects on people, human structures, and the natural environment.

4.4 Summary

Earthquake prediction is a branch of the science of seismology concerned with the specification of the time, location, and magnitude of future earthquakes within stated limits, and particularly "the determination of parameters for the next strong earthquake to occur in a region. Earthquake prediction is sometimes distinguished from earthquake forecasting, which can be defined as the probabilistic assessment of general earthquake hazard, including the frequency and magnitude of damaging earthquakes in a given area over years or decades. Prediction can be further distinguished from earthquake warning systems, which upon detection of an earthquake, provide a real-time warning of seconds to neighboring regions that might be affected.

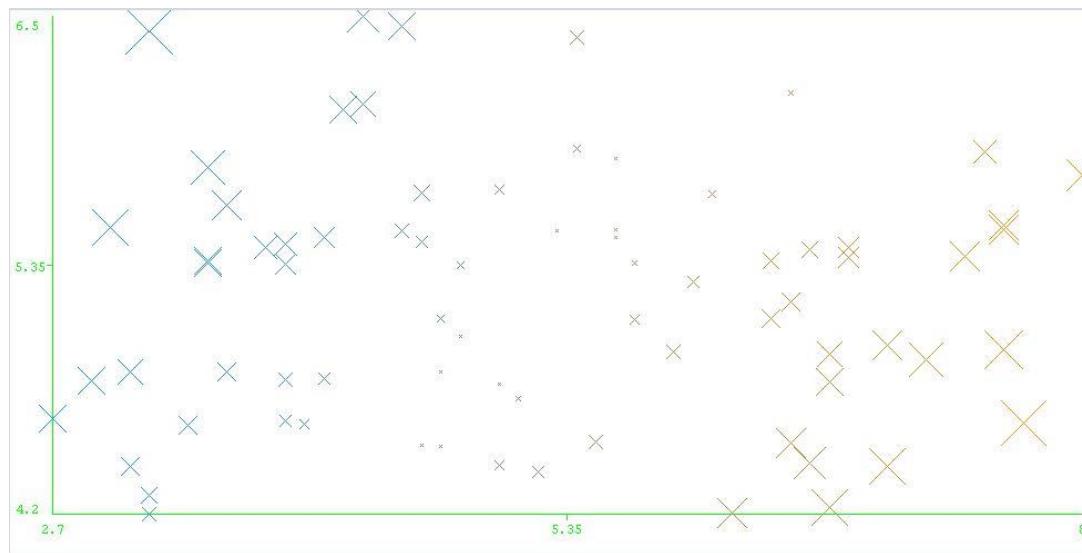


Fig 4.4.1: Magnitude vs Predicted magnitude

Earthquake prediction is an immature science. It has not yet led to a successful prediction of an earthquake from first physical principles. Research into methods of prediction therefore focus on

empirical analysis, with two general approaches: either identifying distinctive precursors to earthquakes, or identifying some kind of geophysical trend or pattern in seismicity that might precede a large earthquake. Precursor methods are pursued largely because of their potential utility for short-term earthquake prediction or forecasting, while 'trend' methods are generally thought to be useful for forecasting, long term prediction (10 to 100 years' time scale) or intermediate term prediction (1 to 10 years' time scale).

CHAPTER 05

SUMMARY, CONCLUSION, RECOMMENDATION AND IMPLICATION FOR FUTURE RESEARCH

5.1 summary of the study

Earthquake prediction research has been led for more than 100 years with no conspicuous triumphs. Cases of leaps forward have neglected to withstand investigation. Broad ventures have neglected to discover solid forerunners. Hypothetical work proposes that blaming is a non-direct process which is profoundly touchy to unmeasurably fine points of interest of the condition of the Earth in an extensive volume, not simply in the prompt region of the hypocenter. Any little quake in this way has some likelihood of falling into a substantial occasion. Solid issuing of cautions of impending extensive earthquake has all the earmarks of being appears to be effectively impossible.

5.2 Conclusion

The problem of earthquake prediction is based on data extraction of pre-cursory phenomena and it is highly challenging task various computational methods and tools are used for detection of pre-cursor by extracting general information from noisy data. By using common frame work of clustering we are able to perform multi-resolution analysis of seismic data starting from the raw data events described by their magnitude spatiotemporal data space. This new methodology can be also used for the analysis of the data from the geological phenomena e.g. we can apply this clustering method to volcanic eruptions

5.3 Recommendations

Though we mention some related work but its amount is very few and directly there has no work like us based on Bangladeshi Data. We understand all their research process and work style after that we started to fix our research goal. After a hard effort by doing all of the work step by step finally, we are at a stage what can be said it is our expected research goal. So, for making this kind of research work it need a tremendous work for guiding us through the right path of research. We have experienced some different problems that were inside the starting of our research. We have also stuck with the mastering of the large field of Data mining and Machine Learning. In this research Work our supervisor Shah Mohammad Tanvir Siddique sir helped us a lot and guided us for making this research project successful

5.3 Implication for further study

The future of data mining lies in predictive analytics. The technology innovations in data mining since 2000 have been truly Darwinian and show promise of consolidating and stabilizing around predictive analytics. Nevertheless, the emerging market for predictive analytics has been sustained by professional services, service bureaus and profitable applications in verticals such as retail, consumer finance, Telecommunications, travel and leisure, and related analytic applications. Predictive analytics have successfully proliferated into applications to support customer recommendations, customer value and churn management, campaign optimization, and fraud detection. On the product side, success stories in demand planning, just in time inventory and market basket optimization are a staple of predictive analytics. Predictive analytics should be used to get to know the customer, segment and predict customer behaviour and forecast product demand and related market dynamics. Finally, they are at different stages of growth in the life cycle of technology innovation

REFERENCE

W.Dzwinel (2005) et al Non multidimensional scaling and visualization of earth quake cluster over space and feature space, nonlinear processes in geophysics 12 pp1-12.

C.Lomnitz (1994). Fundamentals of Earthquake prediction

B.Gutenberg & C.H. Richter (1996) Earthquake magnitude, intensity, energy & acceleration bulseism soc. Am 36, 105-145

C.Brunk, J.Kelly & Rkohai (1997) “Mineset An integrate system for data access, Visual Data Mining & Analytical Data Mining”, proceeding of the 3rd conference on KDD

Andenberg M.R (1973) Cluster Analysis for application, New York, Acedamic, Press

S.Sathyabama, K.Thyagarajah, D. Ayyamuthukumar “Predicting the Earthquake using Bagging Method in Data Mining”,

Kristína Machová, František Barčák, Peter Bednár “A Bagging Method using Decision Trees in the Role of Base Classifiers”,

Witold Dzwinel, David A. Yuen, Krzysztof Boryczko, Yehuda Ben-Zion, Shoichi Yoshioka, Takeo Ito “Cluster Analysis, Data-Mining, Multi-dimensional Visualization of Earthquakes over Space, Time and Feature Space”,

Anon, (2015). [online] Available at: Bbc.co.uk, (2015). BBC - GCSE Bitesize: Case studies.

APPENDICES

Appendix A: Project Reflection

The purpose of this appendix is to provide an introduction to **Project Reflection**. The group research project was a challenging and enjoyable experience typical of the course as a whole. We have had little exposure to group work at university so it was a nice change to be part of an effective and dynamic team. I think we complemented one another quite well both in bringing together interdisciplinary perspectives and in balancing the work at hand.

The experience taught us that planning and crafting responses takes a longer time in teams than on your own. The extensive effort required was ultimately a good thing. When working alone, you can end up with a result that is identical to your initial plans. In our group, we were constantly developing and refining one another's ideas. It was fascinating just how productive our group meetings were. The time seemed to fly and yet we always got a lot done and managed to help another along the way towards the endpoint of having a substantive policy.

Appendix B: Related Diagrams

The purpose of this appendix is to provide an introduction to Related Diagrams (RD) and pointers to some of the foundational references. An influence diagram B is a compact graphical representation of a decision.

PLAGIARISM CHECKER SCREENSHOT

The screenshot shows the Plagiarism Checker interface. On the left sidebar, there are links for 'Upload', 'Papers', 'Payments', 'Free', 'Earn money', 'RATE US' (with a 5-star rating), and 'CONTACT US'. The main content area displays a file named 'earthquake prediction.docx' uploaded 5 minutes ago. A circular progress bar indicates a 'Similarity' of 21%. Below the bar, three categories are shown: Paraphrase (1%), Improper Citations (0%), and Matches (26). A red warning message at the bottom states 'HIGHEST PLAGIARISM RISK' with three red stars above it. A purple footer bar includes a 'View detailed report' button, a question about English words, and a sign-in prompt for 'priimtinas'.

Search...

earthquake prediction.docx
5 minutes ago

21%
Similarity

1% Paraphrase 0% Improper Citations 26 Matches

HIGHEST PLAGIARISM RISK

View detailed report

Activate Windows
Go to PC settings to activate Windows.

Is this an English word and is it correct? priimtinas Yes No