EARTHQUAKE PREDICTION

Vijay Kumar Kodari | Sri Krishna Tirthala | Dharma Teja Bandaru

[CS519 APPLIED MACHINE LEARNING I] [Open Machine Learning Project]

Introduction:

Man has always looked to computer systems for help in every area of human endeavours and even in understanding very complex systems as the human ability is somewhat limited in handling the amount of data generated every day. As the size of this data has increased considerably, the need to explore the use of computer system to learn from the data became imperative.

Machine learning, which is an application of artificial intelligence, provides computer systems the ability to learn autonomously from experience and improve as well without being explicitly programmed. Natural disasters cause massive casualties, damages and leave many injure. Human beings cannot stop them, but timely prediction and due safety measures can prevent human life losses and many precious objects can be saved. Earthquake is one of the major catastrophes. Unlike other disasters, there is no specific mechanism for earthquake prediction, which makes it even more destructive.

Though many of them say that it is impossible to make earthquake predictions, few Scientists declare that it is a predictable phenomenon.

Motivation:

Believing that earthquakes could be predicted, from the past experiences we can say that we can avoid life loss and property damage if we have an estimation on severity of earthquake that will occur in near future.

On a yearly basis, there are approximately 14,000—16,000 fatalities due to earthquakes. The count of fatalities per year due to earthquakes when compared to other disasters and accidents might be less. Whereas, an abrupt and a huge earthquake in a city could take away thousands of lives at once. So, a solution to this problem using machine learning would save many lives and properties with no much of the damage.

Problem Definition:

What could be the Solution?

- When the quake will occur.
- Where it will occur.

The above two factors could be able to give a proper intimation to the governing bodies and it could warn them to take the required actions accordingly.

But considering the data we have, the **WHEN** can only be predicted and that would help to some extent.

What are Solution Benefits?

Earthquakes are natural disasters and it could occur in any place around the earth, but we must be worried if the prediction is in an overcrowded place or in the place of historic

sculptures or national monuments. In this case, predictions that are made accurately could give an intimation to evacuate the place and make amendments to the buildings with concrete and redesign them in a way that it's foundation can take that sort of punch.

Solution Use

Lifetime of the usage of solution and probability of solution to be correct is often based on hypothesis. As it is related to nature, sometimes the levels of earthquake might be more that of expected or it might not occur. The solution from the algorithm always depends on the balance of nature, i.e.; if the solution is made when the ecology system is good, the algorithm might not work as expected when the ecology system is not good in future. The solution might be useful for short-term or Long-term depending on the balance in nature and ecology system.

How would I solve the problem?

Understanding the data:

There are three kinds of datasets provided [sample, Training, Test]

In the train.csv, there is huge data i.e.; 600 million rows of data with two columns 'acoustic_data','time_to_failure' .

acoustic_data : is the acoustic signal measured in the laboratory experiment.

time_to_failure: the time until a failure will occur.

In the test data, it is divided into many segments but there is only data regarding the acoustic data with 15000 rows, and we need to predict the time_to_failure.

Processing the data:

As we need to predict time_to_failure only with acoustic_data present in test datasets, few common statistical features like mean, max, min, standard deviation must be calculated using the data in train.csv.

We can apply the regression and bring out the time_to_failure values.

Proposed solution:

Predicting the time remaining before the next laboratory earthquake i.e.; (time_to_failure) is the solution for this project, for which we are using the (acoustic_data) from the training data and are able to accomplish the task. Furtherly, we can also use classification techniques and work on generating seismic signals in (acoustic_data), so that the **WHERE** issue can also be solved as the values depend upon the laboratory locations.