

ARTIFICIAL INTELLIGENCE

Project Title: *Earthquake Prediction Model*

Phase 1 : Problem Definition and Design Thinking

Problem Definition :

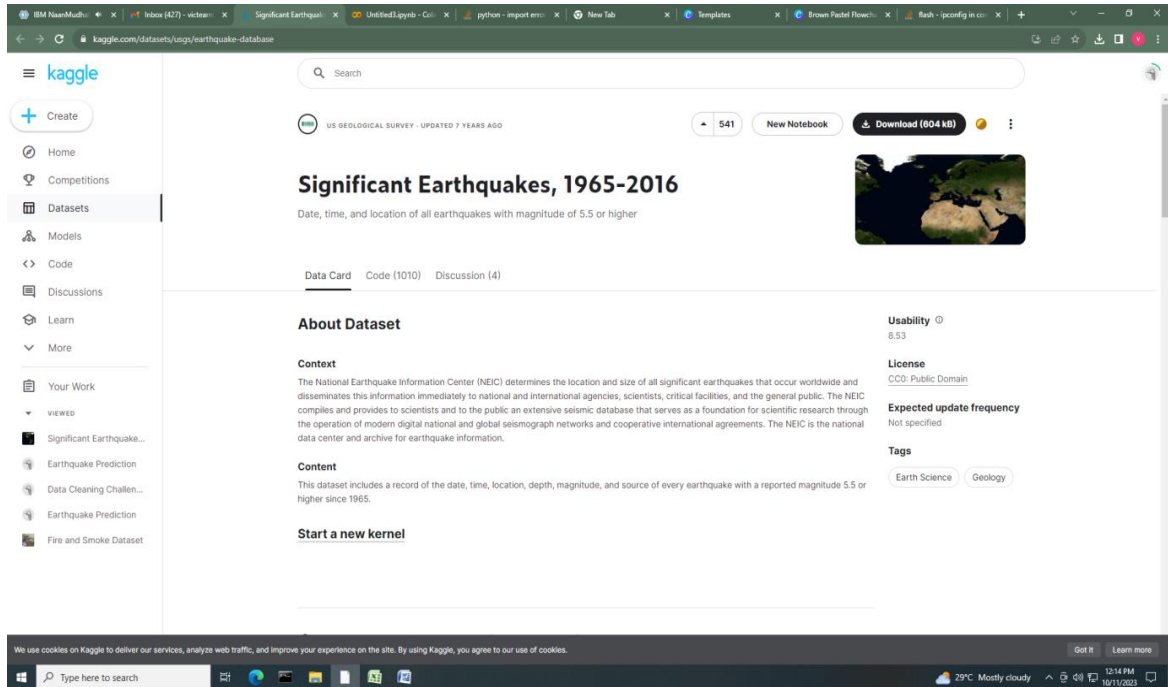
The problem at hand is to develop an earthquake prediction model using a kaggle dataset. The primary objective is to explore and understand the key features of earthquake data, visualize the data on a world map for a global overview, split the data for training and testing, and ultimately construct a neural network model that can predict earthquake magnitudes based on the provided features.

DESIGN THINKING

Data Source

The first step in solving this problem is selecting a suitable kaggle dataset that contains earthquake data. This dataset should include essential features such as date, time, latitude, longitude, depth, and magnitude. The choice of the dataset is crucial as it forms the foundation of our model.

Dataset Source :



Sample Data Snapshot :

	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	
532	#####	15:55:10	-21.558	170.396 Earthquake	150	6.9 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
533	#####	20:55:57	01.156	-150.126 Earthquake	55	5.9 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
534	#####	01:12:17	-16.054	-177.279 Earthquake	15	6.1 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
535	#####	6:25:55	-60.354	-26.485 Earthquake	20	6.1 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
536	#####	0:06:38	-11.793	121.601 Earthquake	20	6 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
537	#####	4:22:17	-31.241	-177.747 Earthquake	25	5.8 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
538	#####	9:13:28	29.734	142.680 Earthquake	15	5.7 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
539	#####	12:55:32	-56.111	-27.62 Earthquake	100	5.7 MW	ISCEMSU	ISCEGM	ISCEGM	ISCEGM	Automatic							
540	#####	21:42:00	-10.665	-78.228 Earthquake	40	8.1 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
541	#####	8:01:34	-1.571	-15.611 Earthquake	15	7 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
542	#####	18:28:55	-18.319	167.58 Earthquake	35	5.7 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
543	#####	14:21:03	22.232	145.978 Earthquake	10	6.7 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
544	#####	2:39:29	38.796	21.242 Earthquake	25	5.9 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
545	#####	16:24:33	19.102	-67.878 Earthquake	26.5	6 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
546	#####	12:45:14	-15.37	-175.109 Earthquake	15	6.4 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
547	#####	3:02:34	-31.977	-68.301 Earthquake	113	6.1 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
548	#####	15:31:05	52.141	-169.158 Earthquake	29.2	5.6 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
549	#####	11:50:33	-23.822	-67.863 Earthquake	135	5.6 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
550	#####	12:49:44	41.746	144.389 Earthquake	32.5	6 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
551	#####	18:45:09	-15.681	167.256 Earthquake	35	6.3 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
552	#####	2:51:54	17.042	-61.885 Earthquake	79.4	5.8 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
553	#####	6:29:54	48.019	146.857 Earthquake	460	5.9 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
554	#####	7:01:12	-57.964	-25.518 Earthquake	35	5.8 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
555	#####	2:19:13	-14.851	167.051 Earthquake	35	5.9 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
556	#####	5:18:19	-29.561	-70.497 Earthquake	91.8	5.7 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
557	#####	7:32:53	6.6	-82.746 Earthquake	16.7	5.8 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
558	#####	4:56:59	-11.954	167.045 Earthquake	125	6.8 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
559	#####	9:31:18	3.251	128.174 Earthquake	75	5.7 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
560	#####	17:17:43	44.472	151.655 Earthquake	15	5.9 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
561	#####	13:06:35	14.351	-91.963 Earthquake	70	5.7 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
562	#####	18:08:15	-1.424	145.703 Earthquake	25	6.3 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
563	#####	21:07:53	-4.928	144.05 Earthquake	65	6.3 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
564	#####	5:08:05	21.487	94.322 Earthquake	94.3	6.7 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
565	#####	20:52:18	-29.533	80.823 Earthquake	20	5.7 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
566	#####	12:26:55	-26.123	-63.152 Earthquake	578	6.3 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
567	#####	15:30:01	37.30217	-116.408 Nuclear E	1.2	31.61	16	5.62 ML	0.245	10	261	1.476	99	0.86	C1329911	CI	CI	Reviewed
568	#####	8:51:59	-19.59	169.834 Earthquake	230	6.3 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
569	#####	15:50:21	-7.099	148.256 Earthquake	30	6.6 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
570	#####	1:22:18	37.126	141.202 Earthquake	55.3	6.3 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
571	#####	8:18:07	-25.494	-70.55 Earthquake	25	7.7 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							
572	#####	27:16:31	-32.838	-111.783 Earthquake	10	5.6 MW	ISCEMBA	ISCEGM	ISCEGM	ISCEGM	Automatic							

FEATURE EXPLORATION

Once the dataset is acquired, it's essential to dive into feature exploration. This phase involves:

1. Data Inspection:

Carefully examining the dataset to understand its structure, data types, and any missing values.

2. Statistical Analysis:

Calculating summary statistics, including mean, median, standard deviation, and quartiles for each feature. This will help us identify outliers and understand the data's distribution.

3. Correlation Analysis:

Investigating the correlations between features, especially between earthquake magnitude and other variables. Identifying highly correlated features can be beneficial for model development.

VISUALIZATION

Visualization plays a crucial role in gaining insights from the data. In this phase:

1. World Map Visualization:

Creating a world map visualization to display the geographical distribution of earthquakes. This can help identify earthquake-prone regions and patterns.

2. Time Series Plots:

Visualizing the earthquake data over time to detect any temporal trends or seasonality.

DATA SPLITTING

To evaluate our model effectively, we need to split the dataset into two subsets:

1. Training Set:

This set will be used to train our neural network model. It should contain a significant portion of the data, ensuring that the model learns from a diverse range of examples.

2. Test Set:

The test set is crucial for evaluating the model's performance. It should be separate from the training data and used to assess how well the model generalizes to unseen earthquake data.

MODEL DEVELOPMENT

In this phase, we focus on building the earthquake prediction model using a neural network. Key steps include:

1. Data Pre processing:

Preparing the data for model input, which may involve normalization, scaling, or encoding categorical variables.

2. Neural Network Architecture:

Designing the architecture of the neural network. This includes defining the number of layers, neurons, activation functions, and loss functions.

3. Model Training:

Training the neural network on the training set using appropriate optimization techniques, such as stochastic gradient descent (SGD) or Adam.

TRAINING AND EVALUATION

The final phase involves training the model and evaluating its performance:

1. Model Training:

Fit the neural network to the training data and monitor its convergence. Adjust hyper parameters as needed to optimize performance.

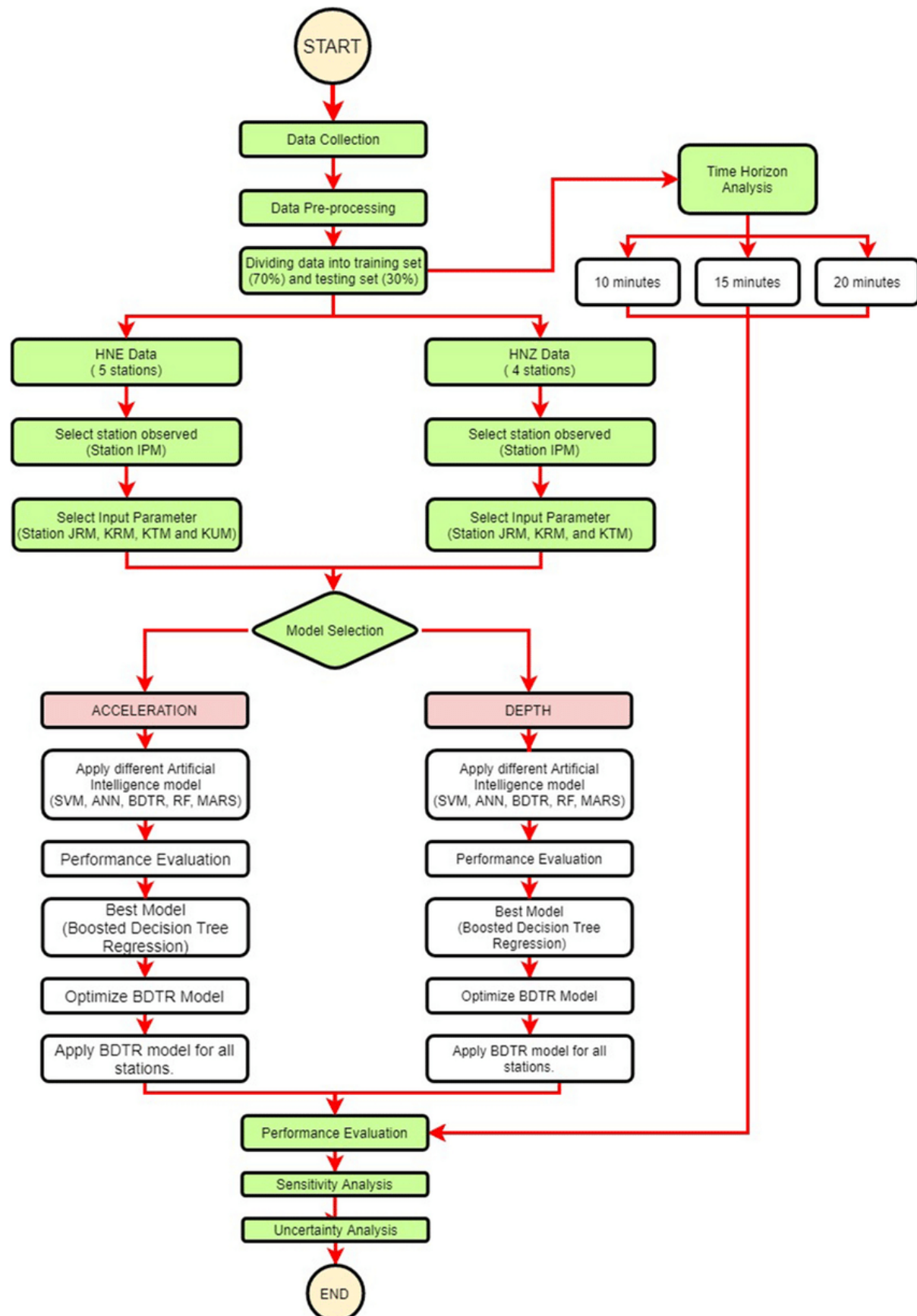
2. Model Evaluation:

Assess the model's performance on the test set using appropriate evaluation metrics, such as mean squared error (MSE) or root mean squared error (RMSE).

3. Fine-Tuning:

If the model's performance is not satisfactory, consider fine-tuning the architecture or exploring advanced techniques like hyper parameter tuning or different neural network architectures.

Flow Chart:



CONCLUSION

In summary, the project aims to develop an earthquake prediction model by following a systematic approach, from data acquisition to model evaluation. The key is to combine data analysis, visualization, and machine learning techniques to create a robust model capable of predicting earthquake magnitudes. Regular monitoring and iterative improvements will be essential throughout the project's lifecycle to achieve the best possible results.