

Note: The project will be weighted at 20% of your total grade, follow the Rubric Cube that provides a general guideline for evaluating the project's different components.

Project Title: Earthquake Prediction and Location Classification

Project Description:

Objective: The goal of this project is to develop a machine learning model that:

1. Predicts the time
of the next earthquake and the location of the next earthquake (Longitude and Latitude).
2. Classifies its location based on historical earthquake data.

Dataset: Use the provided dataset containing information about earthquake occurrences from 1915 to 2023 in Turkey. The dataset includes variables such as date of occurrence, latitude, longitude, depth, and magnitude values (MD, ML, Mw, Ms, Mb).

Tasks:

1. Time Prediction (Regression):

Preprocess the date variable and convert it into relevant time-related features (year, month, day).

2. Split the dataset into training and testing sets, considering the temporal aspect.
3. You can choose an appropriate regression model to predict the time of the next earthquake, or one of deep neural networks models (MLP, ANFIS, CMLP, RNN, or CNN).
4. Evaluate the performance of the model using relevant regression metrics (e.g., Mean Absolute Error, Mean Squared Error, Prediction Rate), and performance metrics relevant for classification.
5. For Location Classification (Multiclass Classification):
 - Preprocess the latitude and longitude variables.
 - Divide the geographic area into distinct regions or classes (e.g., using clustering techniques or predefined regions).

- Implement a classification model (e.g., Random Forest, Support Vector Machine, or Neural Network) to predict the location class of the next earthquake.
- Evaluate the classification model using appropriate metrics (e.g., accuracy, precision, recall).

6. Combine Prediction and Classification:

- Integrate the two models to provide a comprehensive prediction that includes both the time and location of the next earthquake.
- Assess the combined model's overall performance using a suitable evaluation metric.

Deliverables:

1. Python script containing the code for data preprocessing, model training, and evaluation.
2. Written report explaining the methodology, model choices, and results.
3. Visualization of the prediction results (e.g., time series plots, classification maps).
4. Justify your selection for the ML model and Upload your final data with your project.
5. You can compare between two or three models.
6. The performance of the model should be high (don't submit low performance model).
7. Write appropriate references that you used to complete your project.
8. If you want to use different tool (Exe. Matlab, etc.) this will be accepted.

Criteria	Excellent (5)	Good (4)	Satisfactory (2)	Needs Improvement (0)
Time Prediction (Regression)				
Data Preprocessing 1 Point	Detailed and accurate preprocessing of the date variable with proper feature engineering.	Adequate preprocessing, but some potential improvements could be made.	Basic preprocessing, missing some key steps.	Incomplete or incorrect preprocessing.
	Selects and implements an advanced regression	Chooses a suitable model with	Selects a basic model but lacks	Chooses an inappropriate model without

Criteria	Excellent (5)	Good (4)	Satisfactory (2)	Needs Improvement (0)
Model Choice and Implementation 3 Points	model, providing clear reasoning.	reasonable justification.	detailed reasoning.	clear justification.
Model Evaluation 1 Point	Uses relevant regression metrics and thoroughly interprets the results.	Applies appropriate metrics but lacks in-depth interpretation.	Uses metrics but misinterprets results.	Limited or no use of relevant metrics.
Location Classification (Classification)				
Data Preprocessing 1 Point	Proper preprocessing of latitude and longitude variables with geographic region classification.	Adequate preprocessing, but may miss some geographic considerations.	Basic preprocessing, missing key steps for geographic region classification.	Incomplete or incorrect preprocessing.
Model Choice and Implementation 3 Points	Chooses and implements a robust classification model with clear justification.	Selects a suitable model but could provide more detailed reasoning.	Chooses a basic model with limited justification.	Chooses an inappropriate model without clear reasoning.
Model Evaluation 1 Point	Uses appropriate classification metrics and provides a comprehensive interpretation.	Applies relevant metrics but lacks in-depth interpretation.	Uses metrics but misinterprets results.	Limited or no use of relevant metrics.
Combination of Prediction and Classification				
Integration of Models 3 Points	Successfully integrates regression and classification models to provide a cohesive prediction.	Integrates models but with some inconsistencies or gaps.	Attempts integration, but the combination is unclear or ineffective.	Fails to integrate models effectively.
Overall Performance 2 Points	Demonstrates a strong understanding of the strengths and	Understands the overall performance but could provide	Identifies some aspects of performance but	Fails to assess the combined model's

Criteria	Excellent (5) weaknesses of the combined model.	Good (4) more nuanced insights.	Satisfactory (2) lacks a comprehensive understanding.	Needs Improvement (0) performance adequately.
Presentation and Documentation				
Code Readability and Comments 1 Point	Code is well-structured, easy to follow, and includes clear comments.	Code is readable with some room for improvement in comments.	Code is somewhat messy or unclear, lacking comments.	Code is disorganized or unreadable with minimal comments.
Report Quality 2 Points	Provides a well-written report with a detailed explanation of the methodology and results.	Offers a good report but may lack detail or organization.	Presents a basic report with limited detail.	Provides a poorly written or incomplete report.
10 minutes Presentation and Questions 2 Points	Clarity (e.g. explains ideas well, integrates with report, clear introduction and conclusion, obvious transitions, doesn't use jargon, demonstrates knowledge of key points). responds well to all questions	presentation is coherent, with clear introduction, transitions, language use, and conclusion; speaker demonstrates intimate knowledge of the subject. Respond to most of the questions	Presentation is coherent for the most part, but missing 1 or 2 important elements Partially answered the questions	Presentation lacks coherence/ Couldn't answer the questions