EELU X Machine Learning Project

Submitted by: Dr. Mayer Ali
Spring 2025

Challenge Description:

The goal of this project is to analyze data and uncover insights that can help improve outcomes. Students are required to clean the data, analyze trends, visualize findings, derive actionable insights and prediction.

Implementation in Steps:

- Step 1: Load the dataset and explore its structure.
- Step 2: Handle missing values and perform basic data cleaning.
- Step 3: Conduct exploratory data analysis (EDA) using visualizations.
- Step 4: Modeling and Prediction.
- Step 5: Present findings in a well-structured report.

Supporting Material:

(Dataset):https://www.kaggle.com/datasets

Student Deliverables:

- Jupyter Notebook
- Presentation to Present your Thoughts

Group Collaboration:

- Each group consists of **(4-5) students**
- **Menofia-Assuit-Fayoum-Suhag (5-6) students**

Evaluation Criteria:

1. Data Cleaning and Preprocessing (1.5 point)

- Handling Missing Values: Missing values are handled appropriately (e.g., imputation, removal, etc.).
- Data Quality: The dataset is cleaned of errors, inconsistencies, and irrelevant data.
- Documentation: The cleaning process is well-documented in the Jupyter Notebook, with clear explanations of the steps taken.

2. Exploratory Data Analysis (EDA) (1.5 point)

- Visualizations: Effective use of visualizations (e.g., histograms, scatter plots, box plots) to explore trends and patterns.
- Insightfulness: The EDA uncovers meaningful trends and patterns in dataset.
- Depth: The analysis goes beyond surface-level observations to explore correlations, outliers, and potential relationships.
- Clarity: Visualizations are clear, well-labeled, and easy to interpret.

3. Modeling and Prediction (1.5 point)

- Algorithm Selection: The most suitable algorithm is chosen based on the problem type and dataset characteristics.
- Model Selection & Tuning: Appropriate models are chosen, and hyperparameters are optimized.
- Performance Evaluation: Metrics like accuracy, RMSE, or F1-score are used to assess model effectiveness.

4. Report and Presentation (1.5 point)

- Structure: The report is well-organized, with a logical flow (e.g., introduction, methodology, findings, conclusions).
- Clarity: The report and presentation are clear, concise, and easy to follow.

5. Timely Submission (1 point)

- The project is uploaded to the model before the deadline.

6.Individual Discussion (3 point)

- Each student engages in a one-on-one discussion to explain their work and findings.