Assignment – I

Maximum Marks: 15

1. Each student is required to do a thorough Exploratory Data Analysis (EDA) in Python for the data provided against their roll number.

Odd Roll Number	Classification Data 1 Regression Data 1
Even Roll Number	Classification Data 2 Regression Data 2

2. **The student must provide a full report on the EDA along with codes**.
3.
   - In *MS word report* classification/regression prediction results should be clearly mentioned along with the chart of EDA.
   - Further all the codes should be pasted in word file.
   - Finally, the word file should be converted into pdf file and submit to MS-TEAM.
   - **File name which needs to be uploaded must have file name: 20BCP013_Hardik_Inani (Example)**

4. After EDA, you must apply the ML Model as per detail mentioned below **(against their roll number)**:

Odd Roll Number	**Logistic Regression** **ANN (For Regression)** **SVM (For Classification)** **KNN (For Regression)**
Even Roll Number	**Navie Bayes** **SVM (For Regression)** **KNN (For Classification)** **Linear Regression**

5. During the training of a machine learning (ML) model, conduct the essential processes, such as data transformation and encoding.
6. Compare the results of both classifiers and regressors using the appropriate parameters and charts, and write a full report with accompanying codes.
7. Only one report comprising the results of both the EDA and the ML model should be submitted.
8. **Clearly mention your roll number, name, branch and assignment name as Assignment 1 before submission.**

Note: -

- **Classification Data 1: - Drugs Classification**
- **Classification Data 2: - Mushroom Classification**
- **Regression Data 1: - Air Quality Prediction**
- **Regression Data 2: - Possum Age Prediction**

Data is attached with assignment in MS Team. Further students are encouraged to tune hyperparameters using newly developed optimization algorithms and prediction results of tenfold cross validation should be submitted.

https://www.kaggle.com/datasets/prathamtripathi/drug-classification

https://archive.ics.uci.edu/ml/datasets/mushroom

https://www.kaggle.com/datasets/abrambeyer/openintro-possum

https://archive.ics.uci.edu/ml/datasets/air+quality