



## MID-TERM PROJECT ASSIGNMENT

### NAVTTC PROGRAM

### ARTIFICIAL INTELLIGENCE (MACHINE & DEEP LEARNING)

Instructor: Azam Rashid

#### Submission:

- Make a Google Collab notebook Or VS Code Editor (Jupyter Extension) to implement this Project Assignment.
- Submit a .ipynb file detailing all the information. No other format will be accepted
- Submission file should be named as Project\_Midterm\_Students-BatchNo.ipynb
- Deadline for this Mid-Term Project Assignment is Monday 16-09-2024.
- Strictly follow the submission deadline.
- Make Submission in the Midterm Project Assignment on Google Form and press the submit button.
- Click here to submit the Project Assignment

**Problem Statement:** Use Heart Disease Dataset to build a classification model to predict Heart disease.

#### **Load the dataset:**

Preprocess the dataset Include necessary data preprocessing steps, such as handling missing values, encoding categorical variables, and scaling features.

Identify and separate the target variable from the dataset. The target variable indicates the presence or absence of Heart disease.

Split the dataset into training and testing sets (e.g., 80% training, 20% testing).

#### **Train and evaluate the following algorithms:**

- Logistic Regression
- Decision Tree Classifier
- Random Forest Classifier
- Gradient Boosting Classifier
- Support Vector Machine (SVM)

**For each algorithm, calculate and compare the following metrics:**

- Gini Coefficient (use AUC score for binary classification)
- Confusion Matrix
- Accuracy
- Precision
- Recall
- F1 Score
- AUC-ROC Curve

**Report Findings:**

Summarize the performance of each model. Discuss which model performed best based on the evaluation metrics.

**Model Saving:**

Save the Best Performance Model on local hard drive or Google Collab.

**Model GUI:**

Make a GUI (Graphic User Interface) Using Streamlit Python Package for giving the categorical features as input for the prediction of Heart Disease.

**Model Prediction Result Storage Through API:**

By Using Fastapi Python Package save the Prediction Results in PostgreSQL DB. Only Authorized User may Save Results.

**Instructions:** You can download the datasets from here:

Heart Disease Dataset: <https://www.kaggle.com/datasets/abdmental01/heart-disease-dataset>