**ML Assignment 4 – Logistic Regression, Decision Tree, KNN**

**Campus Placement Prediction**

The goal of this assignment is to develop and evaluate predictive classifier models to determine a student's placement status based on a variety of academic and extracurricular features. The dataset includes factors such as CGPA, internship experience, project involvement, and other relevant attributes. By building and comparing three distinct machine learning models — Logistic Regression, Decision Tree, and K-Nearest Neighbors (KNN) — you will assess which model best predicts whether a student will be placed or not.

**Dataset Link:**

<https://raw.githubusercontent.com/ArchanaInsights/Datasets/main/campus_placement.csv>

**Machine Learning Model Building Steps:-**

1. **Data Preprocessing:**
   1. **Load the Dataset:** Load the dataset and perform an initial exploration to understand its structure and content.
   2. **Handle Missing Values:** Check for and handle any missing values in the dataset if present.
   3. **Encode Categorical Features:** Convert categorical features into numerical format using appropriate encoding techniques.
   4. **Feature Selection:** Select the relevant features and the target variable for your analysis.
   5. **Data Splitting:** Split the dataset into training and testing sets (e.g., 80% training, 20% testing).
   6. **Feature Scaling**: Standardize or normalize numerical features using parameters derived from the training set, and apply the same scaling to the test set.
2. **Logistic Regression - Model Building and Evaluation:**
   1. **Model Building:** Build a Logistic Regression model using the training dataset. Experiment with different values for the **max\_iter** parameter e.g., LogisticRegression(max\_iter=300).
   2. **Model Evaluation:** Evaluate the model on the testing dataset and report the accuracy score.
3. **Decision Tree - Model Building and Evaluation:**
   1. **Model Building:** Build a Decision Tree model using the training dataset. Experiment with different values for the **max\_depth** parameter to control the tree's depth and prevent overfitting.
   2. **Model Evaluation:** Evaluate the model on the testing dataset and report the accuracy score.
4. **K-Nearest Neighbors (KNN) - Model Building and Evaluation:**
   1. **Model Building:** Build a KNN model using the training dataset. Experiment with different values of k (number of neighbors).
   2. **Model Evaluation:** Evaluate the model on the testing dataset.
   3. **Reporting:** Report the accuracy score for each value of k and provide a brief interpretation of the results. Choose the best k value and justify your choice.
5. **Comparison and Analysis:**
   1. Compare the accuracy scores of the three models.
   2. Discuss which model performed best and why, based on the accuracy scores and other relevant factors.
   3. Reflect on the strengths and weaknesses of each model in the context of this dataset.