Assignment: Prediction Model with Decision Trees

Objective:
Utilize machine learning classification techniques, specifically decision trees, to predict whether students will pass or fail based on student performance data.

Tasks:

1. **Dataset Exploration:**
   - Download the dataset from Avans+ Connect+.
   - Explore the dataset using Pandas for data cleaning, manipulation, and exploration.
   - Examine the attributes, identify data types, and handle any missing or inconsistent values.
2. **Data Preprocessing:**
   - Convert any textual or categorical attributes into a numerical format suitable for machine learning algorithms.
   - Split the dataset into features (independent variables) and the target variable (pass/fail). Download and read Dataset description carefully.
3. **Building Decision Trees:**
   - Utilize the scikit-learn library to build a decision tree using the **DecisionTreeClassifier** function.
   - Implement cross-validation techniques to assess the model's generalization performance.
4. **Visualization:**
   - Create a visual representation of the decision tree using appropriate functions within scikit-learn.
5. **Model Evaluation:**
   - Evaluate the performance of the decision tree model using a confusion matrix.
   - Interpret the results and discuss the accuracy, precision, recall, and F1-score.
6. **Presentation Preparation:**
   - Prepare a Jupyter Notebook documenting the entire process.
   - Include code snippets, visualizations, and explanations for each step.
   - Clearly present the confusion matrix and discuss the implications of the model's predictions.
7. **Class Presentation:**
   - Each student will present their findings during the class on 4/12/2023.
   - Highlight the dataset characteristics, preprocessing steps, decision tree model creation, and evaluation metrics.

Dataset Information:

- The dataset is available on Connect+.
- It includes a mix of words, phrases, and numeric attributes

Submission Deadline:

- Jupyter Notebooks /project (two format: PDF + Jupyter) should be submitted at end of the class on 4/12/2023.