EXERCISE 3: CLASSIFICATION PROCESS USING WEKA TOOL

Objectives:

a) Perform classification on the Iris.arff dataset using the Naive Bayes classifier algorithm.

b) Conduct classification on the weather.nominal.arff dataset using the J48 Tree classifier algorithm.

c) Execute classification on the soybean.arff dataset using the RandomForest Tree classifier algorithm.

Procedure:

a) Classification on Iris.arff dataset using Naive Bayes Classifier Algorithm

1. Setup:
   - Open Weka Tool on your computer.
   - Load Iris.arff dataset.
   - Explore dataset attributes and structure.
   - Preprocess dataset (handle missing values, convert nominal to numeric, etc.).
2. **Classification:**
   - Choose Naive Bayes classifier algorithm from 'Classify' tab.
   - Set classifier options and parameters.
   - Split dataset into training and testing sets using cross-validation or holdout method.
   - Apply Naive Bayes classifier to training dataset.
3. **Evaluation:**
   - Evaluate classifier's performance using various metrics: accuracy, precision, recall, and F1-score.
4. **Interpretation and Reporting:**
   - **Generate the following reports:**
     - ✓ Confusion Matrix
     - ✓ ROC Curve
     - ✓ Precision-Recall Curve
   - **Interpretation:**
     - ✓ Analyze confusion matrix for classification accuracy and misclassifications among classes (setosa, versicolor, virginica).
     - ✓ Evaluate ROC curve and precision-recall curve to assess classifier's performance across different thresholds.

b) Classification on weather.nominal.arff dataset using J48 Tree Classifier Algorithm

1. **Setup:**
   - Open Weka Tool on your computer.
   - Load weather.nominal.arff dataset.
   - Explore dataset attributes and structure.
   - Preprocess dataset if necessary (handle missing values, convert nominal to numeric, etc.).
2. **Classification:**
   - Choose J48 Tree classifier algorithm from 'Classify' tab.
   - Set classifier options and parameters.
   - Split dataset into training and testing sets using cross-validation or holdout method.
   - Apply J48 Tree classifier to training dataset.
3. **Evaluation:**
   - Evaluate classifier's performance using various metrics: accuracy, precision, recall, and F1-score.
4. **Interpretation and Reporting:**
   - **Generate the following reports:**
     - ✓ Confusion Matrix
     - ✓ Decision Tree Visualization
     - ✓ Classification Summary
   - **Interpretation:**
     - ✓ Analyze confusion matrix for classification accuracy and misclassifications among different weather conditions.
     - ✓ Review decision tree visualization to understand rules and splits used by the classifier.
     - ✓ Provide a classification summary highlighting key performance metrics like accuracy, precision, recall, and F1-score.

c) Classification on soybean.arff dataset using RandomForest Tree Classifier Algorithm

1. **Setup:**
   - Open Weka Tool on your computer.
   - Load soybean.arff dataset.
   - Explore dataset attributes and structure.
   - Preprocess dataset if necessary (handle missing values, convert nominal to numeric, etc.).

2. **Classification:**
   - Choose RandomForest Tree classifier algorithm from 'Classify' tab.
   - Set classifier options and parameters.
   - Split dataset into training and testing sets using cross-validation or holdout method.
   - Apply RandomForest Tree classifier to training dataset.
3. **Evaluation:**
   - Evaluate classifier's performance using various metrics: accuracy, precision, recall, and F1-score.
4. **Interpretation and Reporting:**
   - **Generate the following reports:**
     - ✓ Confusion Matrix
     - ✓ Feature Importance Plot
     - ✓ Out-of-Bag Error Plot
   - **Interpretation:**
     - ✓ Analyze confusion matrix for classification accuracy and misclassifications among different soybean classes.
     - ✓ Review feature importance plot to identify influential features in classifying soybean types.
     - ✓ Assess out-of-bag error plot to understand model's generalization performance during training.

Video References:

1. **https://youtu.be/NsjH880c-AU**
2. **https://youtu.be/bPrTeUAS6_I**
3. **https://youtu.be/kOLFj-QF32A**
4. **https://youtu.be/tlBf2FpXGeY**
5. **https://youtu.be/_CYIZilARgw**

Web references:

1. **https://www.tutorialspoint.com/weka/weka_classifiers.htm**
2. **https://www.geeksforgeeks.org/how-to-use-classification-machine-learning-algorithms-in-weka/**