School of Computer Science and Engineering, VIT Chennai.

BCSE209P Machine Learning

Lab-7 **Naïve Bayes and Decision Tree**

Faculty : Dr. R. Jothi

**Due Date : 03/03/2025**

Submit your python code (Jupyter notebook): with output for all the questions.

Q1. Suppose you want to build Naïve Bayes classifier for predicting whether a cricket match will be played in the given weather conditions or not. Here the weather conditions are described by features outlook, temperature, humidity and wind. The target is play with two class labels *Yes* and *No*. (Dataset: play.csv)

1. Implement Naïve Bayes classifier which must
   * Print the class prior probabilities form the training set
   * Classify the test sample <Rain, Cool, High, Strong>. Need to print likelihood for <Rain, Cool, High, Strong> and also class conditional probabilities for “YES” and “NO” classes.
2. Use sklearn CategoricalNB to validate the results obtained above.

Q2. Use the same dataset play.csv, and

1. Compute information gain for all the attributes and display them.
2. Find which attribute will become the root node of the decision tree.
3. Scikit DecisionTreeClassifier
   1. Train using DecisionTreeClassifier using tennis dataset;
   2. Classify the test sample <Rain, Cool, High,Weak>.
   3. Draw the decision tree for a max depth 2

graph = Source( tree.export\_graphviz(id3\_model, out\_file = None))

SVG(graph.pipe(format='svg'))

1. Check whether root node that you identified is same as the one returned by Scikit DecisionTreeClassifier.

Q3. Use sklearn/scikit libraries to implement Naïve Bayes and Decision Tree algorithms for the following prediction problems and report the performance.

* 1. Diabetes Prediction (diabetes.csv)
  2. Iris species classification (iris.csv)