Tutorial-3

1. Consider the following database of a car insurance company shown below:

Name	AgeGroup	CarType	CrashRisk
Ben	30-40	Family	Low
Paul	20-30	Sports	High
Bill	40-50	Sports	High
James	30-40	Family	Low
John	20-30	Family	High
Steven	30-40	Sports	High

(a) Using the above training data, show how a decision tree is learned for the class label Attribute CrashRisk. Assume the algorithm uses the Gini- index measure.

(b) Using the above classifier, determine the class label of the following data:
{Peter, 20-30, Sports} and{Bob, 40-50, Family}

2. Build a Decision Tree using the training data in the table given below. Divide the Height attribute into ranges as follows: (0,1.6], (1.6,1.7], (1.7, 1.8], (1.8, 1.9], (1.9, 2.0], (2.0, 5.0]

Gender	Height	Class
F	1.6 m	Short
M	2 m	Tall
F	1.9 m	Medium
F	1.88 m	Medium
F	1.7 m	Short
M	1.85 m	Medium
F	1.6 m	Short
M	1.7 m	Short
M	2.2 m	Tall
M	2.1 m	Tall
F	1.8 m	Medium
M	1.95 m	Medium
F	1.9 m	Medium
F	1.8 m	Medium
F	1.75 m	Medium

3.

(a) In a text mining application, 20 documents are retrieved for a given query. 7 of the retrieved documents are relevant. The total number of relevant documents in the database is 30. When 30 documents are retrieved for the same query, 10 are found to be relevant. Plot the recall Vs. precision for this text retrieval system.

(b) A classifier is tested with a number of test data. The classifier output and the correct class are shown below. Draw the confusion matrix for the classifier.

Srl No.	Classifier Output	Correct Class
1	C1	C2
2	C1	C1
3	C1	C3
4	C2	C2
5	C2	C2
6	C2	C2
7	C3	C1
8	C3	C1
9	C3	C1

4 . Normalize the following data 13, 15, 16, 16, 19, 20,20, 21, 22,22, 25,25,25,25,25, 30,33,33,35 . Use min-max, z-score and decimal-scaling normalization techniques.

Using a Bayesian Classifier, classify the tuple (Adam, M, 1.95) as short, medium or tall. The training data for the classifier is given below:

Name	Gender	Height	Output
Kristina	F	1.6m	Short
Jim	M	2m	Tall
Maggie	F	1.9m	Medium
Martha	F	1.88m	Medium
Stephanie	F	1.7m	Short
Bob	M	1.85m	Medium
Kathy	F	1.6m	Short
Dave	M	1.7m	Short
Worth	M	2.2m	Tall
Steven	M	2.1m	Tall
Debbie	F	1.8m	Medium
Todd	M	1.95m	Medium
Kim	F	1.9m	Medium
Amy	F	1.8m	Medium
Wynette	F	1.75m	Medium