IIT MANDI CS-660 DATA MINING FOR DECISION MAKING

ASSIGNMENT 2: Classification using Naïve Bayes and Decision Tree Classifiers

(Group with 5-8 students; Each Group >= 5 students)

Prof. Varun Dutt **Due: Before 11:59 PM August 20, 2025**

Readings:

- Class notes and slides from week 2 and 3
-
- Weka website: http://www.cs.waikato.ac.nz/ml/weka/

Objectives:

- Classification of a dataset using Naïve Bayes and Decision Tree Classifiers.
- Try classification using WEKA.

Assignment:

1. For the data shown below: [10]

NAME of training pattern	Attributes			Class
	HABIT	EATS	FOOTWEAR	
T1	Gabby	Baked	Clogs	Student
T2	Gabby	Roasted	Sandals	Professor
T3	Gabby	Baked	Sandals	Student
T4	Quiet	Fried	Sandals	Professor
T5	Gabby	Fried	Clogs	Student
T6	Quiet	Baked	Sandals	Student
T7	Gabby	Fried	Sandals	Professor
T8	Quiet	Fried	Clogs	Student

Use the following Bayesian Estimator for the prior and likelihood.

Now, classify the test patterns below:

Name	Attributes			Class
	Eats	**Habits**	**Footwear**	?
R1	Roasted	Quiet	Clogs	?
R2	Fried	-	Sandals	?
R3	-	Gabby	-	?
R4	Roasted	-	-	?
R5	-	-	Clogs	

Use the Bayes rule and please show your working in your submission.

2. Now, based upon the .ARFF file created in the last assignment in WEKA, please develop a prediction for the above five test patterns. For the purpose of using the WEKA tool, you may load the five test patterns one-by-one in WEKA and derive the prediction for each of them. Please report at least one error measure as well as the confusion matrix for each of the five test patterns. Do the probability estimates for the classes of the five test patterns derived in WEKA converge to those found by hand above? If "yes," then why "yes" and if "no," then why "no"? [5]

3. Build a suitable decision tree for the dataset shown below. [15]

NAME of training pattern	Attributes			Class
	HABIT	EATS	FOOTWEAR	
T1	Gabby	Baked	Clogs	Student
T2	Gabby	Roasted	Sandals	Professor
T3	Gabby	Baked	Sandals	Student
T4	Quiet	Fried	Sandals	Professor
T5	Gabby	Fried		Student
T6	Quiet	Baked	Sandals	Student
T7	Gabby	Fried	Sandals	Professor
T8	Quiet	Fried	Clogs	Student

Now, please derive suitable classification rules from the developed decisiontree. Do all the rules give a single class for a set of attributes? If "yes," then why "yes" and if "no," then why "no"?

Now, using the decision tree, please classify the test patterns below:

Name	Attributes			Class
	Eats	**Habits**	**Footwear**	?
R1	Fried	Gabby	-	?
R2	-	-	Sandals	?
R3	-	Quiet	-	?
R4	Roaster	-	-	?
R5	Fried	-	Sandals	

Please ensure that the names of all team members are on your submission!

1. Please submit a Word or PDF document with your responses to the above questions. In addition, please submit your ARFF files for the training and test sets in question
2. Only one group member needs to submit the assignment to LMS.