

Worksheet

1

Name:

Roll No: Class: Section:

Date:

CLASSIFICATION AND CLUSTERING

Unit-9

1. Define the following.
 - a. False Positive
 - b. Multiclass classification
 - c. F1 Score
2. Fill in the blanks:
 - a. is the process of categorizing a dataset into classes or groups.
 - b. In logistic regression the dependent variable is always
 - c. is used to validate the performance of a classification model.
 - d. If the value is low, it means there are more false alarms than the actual ones.
3. Out of false positive or false negative, which one is dangerous for the patient? Why? Give reasons.
4. Complete the following table:

	Type of data used (labelled/ unlabelled/ continuous/ discrete/ordinal etc.)	Type of learning	Prediction (numeric/ textual/ both)	Application
Clustering				
Regression				

5. Why is K means clustering called centroid-based clustering?