**DEPRATMENT OF COMPUTER SCIENCE & ENGINEERING**

**BAPATLA ENGINEERING COLLEGE: BAPATLA**

**SUBJECT: IDA/ 14CS401**

**UNIT – IV**

**CLASSIFICATION**

1. Define Machine Learning?
2. Define Supervised Machine Learning? Give an example of Supervised Machine Learning Algorithm? What are the different categories in them?
3. Define Un-Supervised Machine Learning? Give an example of Un-Supervised Machine Learning Algorithm?
4. Define the following terms
   1. Input Variable/Feature, or Independent Variable, or Predictor Variable
   2. Output Variable/Feature, or dependent Variable, or Class Variable
   3. Continuous Variable
   4. Categorical Variable
5. What is Regression? Give an example dataset suitable for it?
6. What is Classification? Give an example dataset suitable for it? What is class label of a classification problem?
7. Define training and test set of a dataset? What is the purpose of them?
8. In R language, how we divide a data set in to training and test sets?
9. Explain the process of training a machine learning algorithm? How we get test set accuracy?
10. What is confusion matrix? Explain the process of constructing it?
11. What is the model of a machine learning algorithm? Explain the process of constructing a model for a machine learning algorithm?
12. Explain the process of prediction in machine learning? What function is used to predict the target variable in R?
13. Define Underfit and Overfit models?
14. Define generalization of a model? What is the intuition of a generalized model?
15. Define logistic regression? How it works? What function is used to construct a model based on logistic regression?
16. When Logistic regression is applied? How to classify the target variable in logistic regression?
17. How to improve the accuracy of the logistic regression model?
18. What is decision tree? Example the process of construction of a decision tree?
19. How to construct a decision tree in R? Explain about CP in a decision tree?
20. What is the difference between the pure and impure node in a decision tree? How they are used while constructing a decision tree?
21. Explain the process of pruning of a decision tree in R?
22. Explain the process of predicting target variable by using decision tree model in R language?
23. Write the difference between Classical decision trees and conditional inference trees in R language?
24. What is random forest? Explain the algorithm steps?
25. How to determine variable importance for random forest model?
26. Define the following term
    1. Sensitivity
    2. Specificity
    3. Positive predicted value
    4. Negative predicted value
    5. Accuracy