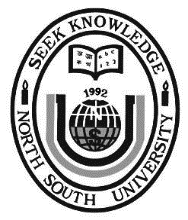
**CSE 445: Machine Learning** 

**Sec – 4, Faculty - ITN**

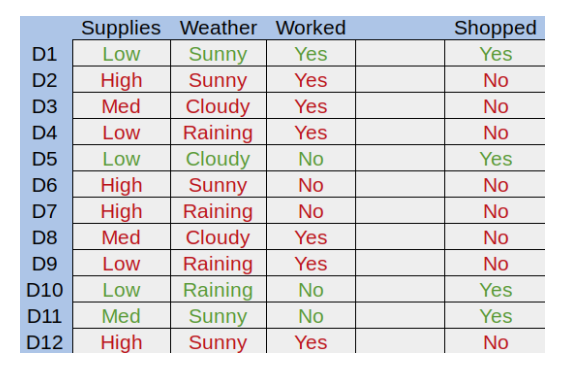
Final, Marks: 60

SET1

Time: 80 Minutes

1. Suppose the following data has been given to you-

8



Find out the root node of this dataset using the decision tree algorithm. Show detailed calculations.

10

1. The following dataset is given to you where X1 and X2 are input features and Y1 is the actual

label and Y2 is the predicted results

|  |  |  |  |
| --- | --- | --- | --- |
| X1 | X2 | Y1 | Y2 |
| 1 | a | 1 | -1 |
| 2 | b | -1 | -1 |
| 3 | c | 1 | 1 |
| 4 | d | 1 | 1 |
| 5 | e | -1 | 1 |
| 6 | f | -1 | 1 |
| 7 | g | -1 | -1 |
| 8 | h | 1 | -1 |
| 9 | i | 1 | 1 |

In this dataset complete the following steps.

a) Add an initial weight to the dataset [1]

b. Compute the error rate [2]

c. compute predictor weight [2]

d. Update weight vector [2]

e. Calculate the normalized weight [3]

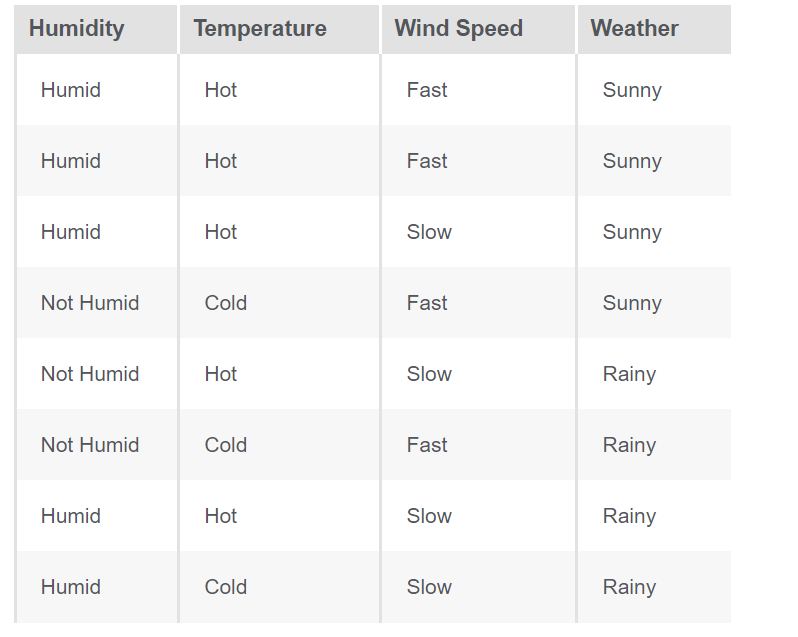
Show all the detailed calculations. Then finally show the updated dataset

5

1. With example, explain 3 differences between the hierarchical clustering and K means clustering
2. Write the formula of polynomial kernel. Explain how this kernel can resolve the nonlinearity issue of the dataset. Give appropriate figures if necessary.

5

1. The following dataset has been given to you



8+2=10

State the Naïve bayes theorem and explain all the parameters. Now if we have an instance where the humidity is humid, temperature is cold and the wind speed is fast then what should be the predicted weather? Use Naïve Bayes’ theorem to do the calculation

11\*2=22

MCQ:

1. Which of the following is finally produced by Hierarchical Clustering?

a) final estimate of cluster centroids

b) tree showing how close things are to each other

c) assignment of each point to clusters

d) all of the mentioned

1. Which of the following are the applications of clustering?
2. Identifying consumer segments and their properties to position products appropriately
3. Identifying patterns of crime in different regions of a city and managing police enforcement based on frequency and type of crime
4. Looking at social media behaviour to find out the types of online communities that exist
5. All of the above
6. Which of the following is required by K-means clustering?

a) defined distance metric

b) number of clusters

c) initial guess as to cluster centroids

d) all of the mentioned

1. Which of the following option is/are correct regarding the benefits of the ensemble model?

1. Better performance

2. Generalized models

3. Better interpretability

a) 1 and 3

b) 2 and 3

c) 1 and 2

d) 1, 2 and 3

5) For which of the following does normalizing your input features do not influence the predictions?

a) decision tree (with usual splitting method)

b) Lasso

c)neural network

d) soft-margin support vector machine

6) The major voting process is considered to be?

a) Sampling

b) Bagging

c) High Variance

d) None of these

7) What do you mean by generalization error in terms of the SVM?

a) How far the hyperplane is from the support vectors

b) How accurately the SVM can predict outcomes for unseen data

c) The threshold amount of error in an SVM

8) What do you mean by a hard margin?

a) The SVM allows a very low error in classification

b) The SVM allows a high amount of error in classification

c) None of the above

9) The effectiveness of an SVM depends upon:

a) Selection of Kernel

b) Kernel Parameters

c) Soft Margin Parameter C

d) All of the above

10 ) Which of the following are real world applications of the SVM?

a) Text and Hypertext Categorization

b) Image Classification

c) Clustering of News Articles

d) All of the above

11) Why would we use a random forest instead of a decision tree?

a) For lower training error.

b) To reduce the variance of the model.

c) better approximate posterior probabilities.

d) For a model that is easier for a human to interpret.