Machine Learning Assignment-7

Ans. 1) d

Ans. 2) d

Ans. 3) b

Ans. 4) a

Ans. 5) c

Ans. 6) c

Ans. 7) b

Ans. 8) c

Ans. 9) Gini index: p(A)(1-p(A)) + p(B)(1-p(B)) = 40%(1-40%) + 60%(1-60%) = 0.24 Entropy: -p(A)*log2(p(A)) - p(B)*log2(p(B)) = -40%*log2(40%) - 60%*log2(60%) = 0.97

Ans. 10) The advantages of Random Forests over Decision Tree are:

• Random Forests are less prone to overfitting as compared to decision tree.

• Random Forests are more robust to noise in the dataset.

• Random Forests provide better accuracy compared to decision tree.

Ans. 11) Scaling is the process of standardizing the range of features of a dataset. The need of scaling is to ensure that each feature contributes approximately proportionately to the final distance. Two techniques used for scaling are:

- Min-Max Scaling
- Standardization

Ans. 12) Scaling provides following advantages in optimization using gradient descent algorithm: • It helps to converge faster

• It helps to find global minima

Ans. 13) In case of a highly imbalanced dataset for a classification problem, accuracy is not a good metric to measure the performance of the model because accuracy is computed by dividing the number of correct predictions to

total predictions. As the majority class is over-represented, the classifier may predict the majority class most of the time and still have a high accuracy.

Ans. 14) F-score is a metric that combines precision and recall to provide a single measure of the performance of a classification model. The mathematical formula for f-score is:- F-score = (2 * Precision * Recall) / (Precision + Recall).

Ans. 15) In machine learning, fit(), transform() and fit_transform() are methods of the scikit-learn library used for preprocessing data:

• fit() method is used to fit the data to the model, it is used to calculate the internal parameters of the model.

• transform() method is used to transform the data according to the internal parameters calculated during the fit() method.

• fit_transform() method is used to fit the data to the model and then transform it in one step.