MACHINE LEARNING

1. D) All of the above

2. D) All of the above

3. D) kernel will be changed to linear

4. C) both A & B

5. C) In case of classification problem, the prediction is made by taking mode of the class labels predicted by the component trees.

6. B) Gradient Descent algorithm can keep oscillating around the optimal solution and may not settle.

7. B) Bias will decrease, Variance increase

8. B) model is overfitting

9. Class of A=40%=40/100=2/5

Class of B=60%=60/100=3/5

Gini Index= 1- ((2/5))^2+((3/5))^2=12/25=0.48

Gini index= (3/5)0.48+(2/5)0.48=0.288+0.192=0.48

Entropy= -[(2/5log(2/5)+3/5log(3/5)]=0.971

10. Random Forest has a higher training time than a single decision tree. We should take this into consideration because as we increase the number of trees in a random forest, the time taken to train each of them also increases. Also, Random forest aggregate the tree to limit overfitting as well as error due to bias and therefore produce a better accuracy.

11. We mainly require scaling to normalize the dataset while building the machine learning model. Also, another importance of scaling is to make the dataset standardized which helps us understand the input parameters in a better way.

Two scaling techniques are nominal scale, ordinal scale.

12. Advantages:

* We can use fixed learning rate during training without worrying about learning rate decay.
* It has straight trajectory towards the minimum and it is guaranteed to converge in theory to the global minimum if the loss function is convex and to a local minimum if the loss function is not convex.
* It has unbiased estimate of gradients. The more the examples, the lower the standard error.

13. No accuracy will not be good measure for model performance as the dataset is imbalanced.

We need to use different evaluation metrics for check the accuracy. Also if the dataset is imbalanced we need to talk to concerned stakeholders to find the important parameters only used for model building and use feature selection and feature engineering to rebuild the dataset again.

14.Formula for F-Measure = (2 \* Precision \* Recall) / (Precision + Recall).

The F-score is a measure of a model’s accuracy on a dataset. It is used to evaluate binary classification systems, which classify examples into ‘positive’ or ‘negative’.

The F-score is a way of combining the precision and recall of the model, and it is defined as the harmonic mean of the model’s precision and recall.

The F-score is commonly used for evaluating information retrieval systems such as search engines, and also for many kinds of machine learning models, in particular in natural language processing.

It is possible to adjust the F-score to give more importance to precision over recall, or vice-versa. Common adjusted F-scores are the F0.5-score and the F2-score, as well as the standard F1-score.

15. Fit performs the training, transform changes the data in the pipeline in order to pass it on to the next stage in the pipeline, and fit\_transform does both the fitting and the transforming in one possibly optimized step. fit computes the mean and std to be used for later scaling technique.