

1. B
2. D
3. A
4. D
5. C
6. C
7. B
8. B
- 9.
10. Random forest algorithm avoids and prevents overfitting by using multiple trees. The results are not accurate. This gives accurate and precise results. Decision trees require low computation, thus reducing time to implement and carrying low accuracy.
11. Scaling is required to rescale the data and it's used when we want features to be compared on the same scale for our algorithm. And, when all features are in the same scale, it also helps algorithms to understand the relative relationship better. The two techniques used in scaling are Absolute Maximum Scaling, Min-Max Scaling.
12. Having features on a similar scale will help the gradient descent converge more quickly towards the minima. Specifically, in the case of Neural Networks Algorithms, feature scaling benefits optimization by: It makes the training faster. It prevents the optimization from getting stuck in local optima.
13. Accuracy is not a good metric for imbalanced datasets.
This model would receive a very good accuracy score as it predicted correctly for the majority of observations, but this hides the true performance of the model which is objectively not good as it only predicts for one class.
14. F1 score is a machine learning evaluation metric that measures a model's accuracy. It combines the precision and recall scores of a model. The accuracy metric computes how many times a model made a correct prediction across the entire dataset. This can be a reliable metric only if the dataset is class-balanced; that is, each class of the dataset has the same number of samples.

$$F1 \text{ Score} = 2 \times \text{Precision} \times \text{Recall} / (\text{Precision} + \text{Recall})$$

15. the `fit()` method will allow us to get the parameters of the scaling function. The `transform()` method will transform the dataset to proceed with further data analysis steps. The `fit_transform()` method will determine the parameters and transform the dataset.