

MACHINE LEARNING

1. (a)

2.(b)

3.(b)

4.(c)

5.(d)

6.(b)

7.(a)

8.(a)

9.(a)

10.(a)

11.Ans:-One-hot encoding creates d-dimensional vectors for each instance where d is the unique number of feature values in the dataset. For a feature having a large number of unique feature values or categories, one-hot encoding is not a great choice.

12.Ans:-When we are using an imbalanced dataset, we can oversample the minority class using replacement.

This technique is called oversampling. Similarly, we can randomly delete rows from the majority class to match them with the minority class which is called undersampling.

13.Ans:-The key difference between ADASYN and SMOTE is that the former uses a density distribution, as a criterion to automatically decide the number of synthetic samples that must be generated for each minority sample by adaptively changing the weights of the different minority samples to compensate for the skewed.

14.Ans:-GridSearchCV is a library function that is a member of sklearn's model_selection package. It helps to loop through predefined hyperparameters and fit your estimator (model) on your training set. So, in the end, you can select the best parameters from the listed hyperparameters.

15.Ans:- There are three error metrics that are commonly used for evaluating and reporting the performance of a regression model; they are: Mean Squared Error (MSE). Root Mean Squared Error (RMSE). Mean Absolute Error (MAE).

