

ASSIGNMENT – 39

MACHINE LEARNING

S.NO	Questions	Answers
1	Which of the following methods do we use to find the best fit line for data in Linear Regression?	A) Least Square Error
2	Which of the following statement is true about outliers in linear regression?	A) Linear regression is sensitive to outliers
3	A-line falls from left to right if a slope is _____?	A) Positive
4	Which of the following will have symmetric relation between dependent variable and independent variable?	B) Correlation
5	Which of the following is the reason for over fitting condition?	C) Low bias and high variance
6	If output involves label then that model is called as:	B) Predictive modal
7	7. Lasso and Ridge regression techniques belong to _____?	D) Regularization
8	To overcome with imbalance dataset which technique can be used?	D) SMOTE
9	The AUC Receiver Operator Characteristic (AUCROC) curve is an evaluation metric for binary classification problems. It uses _____ to make graph?	A) TPR and FPR
10	In AUC Receiver Operator Characteristic (AUCROC) curve for the better model area under the curve should be less.	B) False
11	Pick the feature extraction from below:	A) Construction bag of words from a email
12	Which of the following is true about Normal Equation used to compute the coefficient of the Linear Regression?	A) We don't have to choose the learning rate. B) It becomes slow when number of features is very large.

Q13 and Q15 are subjective answer type questions, Answer them briefly.

13. Explain the term regularization?

Regularization is one of the most important concepts in ML. It is a technique to prevent the model from overfitting by adding extra information to it.

Sometimes the machine Learning model performs well with Training Data-Set. But not good with the Testing Data-Set. It means that it is not performing well with unseen data. To overcome this problem and to increase the ML model preformation we want to go for regularization of Dataset where we do the following.

- L1 Method/Lasso Method
- L2 Method/Ridge Method.

Lasso/L1 Method:

Here in Lasso Method will internally tune the Coefficient of Variables such that certain variables will become zero so that the machine will come to know which data variables to be used exactly for learning the ML model.

Ridge/L2 Method:

Here this method will reduce the value of the Coefficient between the variables to the lowest. So that the predicted result of the ML model will be much better.

14. Which particular algorithms are used for regularization?

- L1 Method/Lasso Regression
- L2 Method/Ridge Regression

Lasso /L1 Regression:

Here in Lasso Method will internally tune the Coefficient of Variables such that certain variables will become zero so that the machine will come to know which data variables to be used exactly for learning the ML model. (Least Absolute and selection Operator). Here the Penalty term contains the absolute weight and because of which the slope of the variable can be reduced to zero.

Ridge /L2 Regression:

Here this method will reduce the value of the Coefficient between the variables to the lowest. So that the predicted result of the ML model will be much better. Here the Penalty term contains the square of weights because of which the slope(coefficient) of the variable can be reduced but not to zero.

15. Explain the term error present in the linear regression equation?

The term error in Linear Regression Equation is nothing but when a Dataset is provided to a Model for training it will learn the actual output and during the testing, the model will provide an output that may be different from the actual output. The difference in the outputs is called an error in the Linear Regression equation. Those residual errors are acceptable and if the errors are less then the prediction will be good.

$$Y = a + bX + e$$

e = error.