

Name – Raja Kumar

Batch - DS2312

Assignment- Machine Learning

1. (A) Least Square Error
2. (A) Linear regression is sensitive to outliers
3. (B) Negative
4. (A) Regression
5. (C) Low bias and high variance
6. (B) Predictive model
7. (D) Regularization
8. (D) SMOTE
9. (A) TPR and FPR
10. (B) False
11. (B) Apply PCA to project high dimensional data
12. (A) We don't have to choose the learning rate. (B) It becomes slow when the number of features is very large. (D) It does not make use of the dependent variable.
13. REGULARIZATION: Machine learning model many times the models performs well on training data but performs low on test data. because the model was able to understand the noises and the trend in the training data but was not able to do the same on the test data. Suppose in a exam the teacher gave the same question which was taught in the class. In that case students were able to get good marks but they got low marks when the questions were different in the exam as compared to what was taught in the class. This is the same case with models. They performs good on training data but not that much well on the test data. This is called the overfitting of the model. And to avoid this overfitting we do the regularization of the model. And this is called regularization. The regularization models discourage the use of large coefficients in the model that means they discourage the use of those features which have great impact on the label. By this they are able to build a good and simple model and read the data in a better precise way. Lasso and ridge are two models widely used in regularization.
14. Algorithms used for regularization: 1. lasso (least absolute shrinkage and selection operator)- let us assume we have mobile no., experience, skill, email as features and we have to predict salary which is the label. so in this case lasso will give zero importance to phone no., skill and email because they are not important for the label. If we have add more or more features than chances of overfitting will be more so it minimizes the features as much as possible. Lasso acts as a feature selector. It is able to identify the features in best possible way. In lasso we can also control the learning rate which means we can control at what speed our model is understanding the data. If we give more time than changes are there that the model will be good. And we can also control the iteration rate which means we can control how many times our model is reading the data. It can be 1, 2, 3 or any other number of times. This depends up to us only 2. Ridge- It stands for ridge regression. It is also almost same as lasso. But the difference is that lasso gives zero importance to the features which are not important for the model but ridge gives the importance but very very low. Suppose it gives 80% importance to skill and 90% to experience than it will give 1% or 2% to phone no. and email. It works in this way. In ridge also we can control the speed at which our model is getting trained which is

called as alpha in the machine learning world and we can also control the iteration of our model.

15. Explain the term error present in linear regression equation: The error is the difference between the actual value of the label and the value of the label as predicted by the model. $\text{error} = \text{actual} - \text{the value predicted by the label}$. The model predicts the label based on the various algorithms and by understanding the relation between various features with the label. It then predicts the label based on its understanding. If the model is perfect then the error will be zero which is very very rare. Every model in real time will have something as error. Mathematically it can be explained as: $\text{error} = y_i - y_i^{\wedge}$. y_i is the actual value of the label and y_i^{\wedge} is the predicted value by the linear regression model.