Internship 26

Machine Learning\_Worksheet\_1

Question 1. (A) Least Square Error

Question 2. (A) Linear Regression is sensitive to Outliers

Question 3. (B) Negative

Question 4. (B) Correlation

Question 5. (C ) Low Bias And High Variance

Question 6. (A) Descriptive Model

Question 7. (D) Regularization

Question 8. (D) SMOTE

Question 9. (A) TPR and FPR

Question 10. (B) False

Question 11. (B) Apply PCA to project High dimensional data

Question 12. (A) We don’t have to choose the learning rate., B) It becomes slow when number of features is very large.

Question 13. **Explain the term regularization?**

One of the major aspects of training our machine learning model is avoiding overfitting*.*The model will have a low accuracy if it is overfitting.

**Regularizations** are techniques used to reduce the error by fitting a function appropriately on the given training set and avoid overfitting. This is a form of regression, that regularizes or shrinks the coefficient estimates towards zero. In simple words, this technique converts a complex model into a simpler one, so as to avoid the risk of overfitting and shrinks the coefficients, for lesser computational cost.

The fitting procedure involves a loss function, known as residual sum of squares or RSS. The coefficients are chosen, such that they minimize this loss function. In Ridge regression, RSS is modified by adding the shrinkage quantity. The increase in flexibility of a model is represented by increase in its coefficients, and if we want to minimize the function, then these coefficients need to be small. This is how the Ridge regression technique prevents coefficients from rising too high. Lasso is another variation, in which the function is minimized. This variation differs from ridge regression only in penalizing the high coefficients*.*

A standard least squares model tends to have some variance in it, i.e., this model won’t generalize well for a data set different than its training data. Regularization, significantly reduces the variance of the model, without substantial increase in its bias*.* The tuning parameter, used in the regularization techniques, controls the impact on bias and variance. As the value of tuning parameter increases, it reduces the value of coefficients and thus reducing the variance. Till a point, this increase in parameter value is beneficial as it is only reducing the variance (hence avoiding overfitting), without losing any important properties in the data. But after certain value, the model starts losing important properties, giving rise to bias in the model and thus underfitting. Therefore, the value of tuning parameter should be carefully selected. It is a useful technique that can help in improving the accuracy of your regression models. A popular library for implementing these regularization algorithms is scikit learn.

Question 14. **Which particular algorithms are used for regularization?**

**Following Algorithms are used for Regularization :-**

1. **Ridge Regression :-** Ridge regression is a method for analysing data that suffer from multi-collinearity.Ridge regression adds a penalty (L2 penalty) to the loss function that is equivalent to the square of the magnitude of the coefficients. The regularization parameter (lambda) regularizes the coefficients such that if the coefficients take large values, the loss function is penalized. When lambda tends to 0, the penalty term has no eﬀect, and the estimates produced by ridge regression will be equal to least-squares i.e., the loss function resembles the loss function of the Linear Regression algorithm. Hence, a lower value of lambda will resemble a model close to the Linear regression model. When lambda tends to infinity, the impact of the shrinkage penalty grows, and the ridge regression coeﬃcient estimates will **approach zero** (coefficients are close to zero, but not zero). Ridge regression is also known as the L2 Regularization.Hence**Ridge regression shrinks the coefficients as it helps to reduce the model complexity and multi-collinearity.**
2. **Lasso (Least Absolute Shrinkage and Selection Operator) Regression :-** LASSO is a regression analysis method that performs both feature selection and regularization in order to enhance the prediction accuracy of the model. LASSO regression adds a penalty (L1 penalty) to the loss function that is equivalent to the magnitude of the coefficients. In LASSO regression, the penalty has the eﬀect of forcing some of the coeﬃcient estimates to be exactly equal to zero when the regularization parameter lambda is suﬃciently large. LASSO regression is also known as the L1 Regularization. Hence, **LASSO regression converts coefficients of less important features to zero, which indeed helps in feature selection, and it shrinks the coefficients of remaining features to reduce the model complexity, hence avoiding overfitting.**
3. **Elastic Net Regression :-** Elastic-Net is a regularized regression method that linearly combines the L1 and L2 penalties of the LASSO and Ridge methods respectively.

The working of all these algorithms is quite similar to that of Linear Regression, it’s just the loss function that keeps on changing.

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

* An error term is a residual variable produced by a statistical model, which is created when the model does not fully represent the actual relationship between the independent variables and the dependent variables. As a result of this incomplete relationship, the error term is the amount at which the equation may differ during analysis. The error term is also known as the residual, disturbance, or remainder term, and is variously represented in models by the letters e or u. An error term appears in a statistical model, like a regression model, to indicate the uncertainty in the model. An error term essentially means that the model is not completely accurate and tends to differing results during real-world applications. Although the error term and residual are often used synonymously, there is an important difference. An error term is generally unobservable and a residual is observable and calculable, making it much easier to quantify and visualize. While an error term represents the way observed data differs from the actual data for example Actual Population, a residual represents the way observed data differs from sample population data.