MACHINE LEARNING _WORKSHEET

1. A.) Least square error
2. A.)Sensitive to outliers.
3. A.) Positive
4. A.) Regression
5. C.) low bias and high Variance.
6. B.) Predictive model
7. D.) Regularization
8. D) Smote
9. C) Sensitivity and Specificity

10.) B.) False

11.) B.) Apply PCA to project high dimensional data

12 .) A,B,C

13.) So if we explain the term regularization to the area of Machine learning , then there is one ex like - if we worked on some machine learning model then we might have came to the problem of very high training accuracy and the testing accuracy is low .This happens with two cases where once the variance is high and the bias is low which is named as Overfitting, and when the variance is low and the bias is high that is termed as underfitting .Thus regularization refers to modification that can be made to a learning algorithm to reduce the training error it reduces by ignoring the less important features .So it also prevents fitting issues making the model more robust and decreasing the complexity of a model .

So atlast as a summary we can say that regualaization are used to make model accurate and correct , so it performs better than earlier .

14.) As now we know that regularization is technique to solve the fitting problem in a ML model and to get a better accuracy in the result .

So as this technique is used so we also have, some algorithms behind them to do it in a way better & effective manner

I find these algorithms particularly useful. Most data scientists will find some of their models to overfit at some point during their careers. The general idea behind these algorithms is that they try to minimize and even prevent overfitting.

A. Ridge regression / (L2) Algorithm:

This algo's target is to solve the the overfitting in a model when a data suffers from multicollinearity (means a high degree of correlation between the 2 varaibles so in other words it can also be said as the degree of tolerance). Ridge regressions adds a small squared bias factors to the variables. Such a bias factor pick ups the feature variable , introducing a small amt of bias in the model but greatly reducing the variance .

B. Lasso regression /L1 Algorithm:

The word lasso stands for , least absolute shrinkage and selection operator . Lasso regression is a regualrization technique which is used for a more accurate prediction . so L1 regularization adds penalties results in the coeffiecients values that are closer to zero . This is ideal for making a simple model .

C. Elastic net Algorithm:-

So this is the algo which is a combination of both ridge and lasso , elastic net reduces the impact of different features while not eliminating all of the features. Lasso will eliminate many features , and reduce overfitting in the model, ridge will reduce the impact of the features that are not important for the prediction. But the the elastic net combines the feature

elimination like lasso and feature coefficient reduction like ridge to improve the accuracy of the model and generally we use it when we are confused to select between the ridge and lasso.

15) The term error or e in the linear regression equation is a variable produced by a statistical linear regression equation or model , which was formed when the model does not fully represents the actual relationship between the independent variable and the dependent variables. As a result of this incomplete relationship the error term is the amount by which a linear regression model may differ from the , the actual values during a analysis performance .

For ex_ let y as be the output of the linear regression equation If the actual y differs from the expected y in the equation during a test then the error term does not equal to 0, which means there are others factors which is influencing y.

Another real world example of use of linear regression model and the error term :
In a linear regression model tracking a stock's price over time, the error term is the difference between the expected price at particular time and the price that was actually observed.

In terms of equation if the change seen in the coefficient value in the independent variable for the change in actual dependent variable or the output is not equal to the predicted dependent variable or predicted output then the term error is added in the linear regression equation.