Machine Learning worksheet 1

1. A) Least Square Error

2. A) Linear regression is sensitive to outliers (The slope of the regression line will change due to outliers in most of the cases)

3. B) Negative

4. B) Correlation (the relationship is symmetric between x and y in case of correlation but in case of regression it is not symmetric)

5. C) Low bias and high variance

6. B) Predictive Model

7. D) Regularization

8. D) SMOTE

note: Approach to deal with the imbalanced dataset problem

a) Choose a Proper Evaluation Metric.

b)Classifier divided by the total number of predictions

c)Resampling (Oversampling and Undersampling)

d) SMOTE

e) BalancedBaggingClassifier

9. A) TPR and FPR

10. B) False

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

12. A, B and C

A) We don't have to choose the learning rate.

B) It becomes slow when number of features is very large.

C) We need to iterate

Machine Learning worksheet 1

13. **Answer**

Prevent overfitting is Regularization

There is a principle called Occam's Razor, which states: "When faced with two equally good hypotheses, always choose the simpler."

Regularization is an application of Occam's Razor. It is one of the key concepts in Machine learning as it helps choose a simple model rather than a complex one.

Regularization refers to the modifications that can be made to a learning algorithm that helps to reduce this generalization error and not the training error. It reduces by ignoring the less important features. It also helps prevent overfitting, making the model more robust and decreasing its complexity of a model.

The regularization techniques in machine learning are:

Lasso regression: having the L1 norm

Ridge regression: with the L2 norm

Elastic net regression: It is a combination of Ridge and Lasso regression.

Regularization Techniques

a**. Ridge Regression**

The Ridge regression technique is used to analyze the model where the variables may be having multicollinearity. It reduces the insignificant independent variables though it does not remove them completely. This type of regularization uses the L2 norm for regularization.

It uses the L2-norm as the penalty.

L2 penalty is the square of the magnitudes of beta coefficients.

It is also known as L2-regularization.

L2 shrinks the coefficients, however never make them to zero.

The output of L2 regularization is non-sparse.

Machine Learning worksheet 1

b. **Lasso Regression**

Least Absolute Shrinkage and Selection Operator (or LASSO) Regression penalizes the coefficients to the extent that it becomes zero. It eliminates the insignificant independent variables. This regularization technique uses the L1 norm for regularization.

It adds L1-norm as the penalty.

L1 is the absolute value of the beta coefficients.

It is also known as the L-1 regularization.

The output of L1 regularization is sparse.

c. **Elastic Net Regression**

The Elastic Net Regression technique is a combination of the Ridge and Lasso regression technique. It is the linear combination of penalties for both the L1-norm and L2-norm regularization.

The model using elastic net regression allows the learning of the sparse model where some of the points are zero, similar to Lasso regularization, and yet maintains the Ridge regression properties. Therefore, the model is trained on both the L1 and L2 norms.

14. **Answer**

There are two main types of regularization techniques: Ridge Regularization and Lasso Regularization.

Ridge Regularization :

Also known as Ridge Regression, it modifies the over-fitted or under-fitted models by adding the penalty equivalent to the sum of the squares of the magnitude of coefficients.

This means that the mathematical function representing our machine learning model is minimized and coefficients are calculated. The magnitude of coefficients is squared and added. Ridge Regression performs regularization by shrinking the coefficients present.

Lasso Regression

It modifies the over-fitted or under-fitted models by adding the penalty equivalent to the sum of the absolute values of coefficients.

Machine Learning worksheet 1

Lasso regression also performs coefficient minimization, but instead of squaring the magnitudes of the coefficients, it takes the true values of coefficients. This means that the coefficient sum can also be 0, because of the presence of negative coefficients

15. **Answer**

The error term of a regression equation represents all of the variations in the dependent variable not explained by the weighted independent variables.

A regression equation is a formula for a straight line — in this case, the best-fit line through a scatterplot of data. If there were no errors, all the data points would be located on the regression line; to the extent, that they do not represent error; this is what the error term summarizes.