Assignment 39 Machine Learning Answers

1) A- Least Square Error
2) A- Linear regression is sensitive to outliers
3) B-Negative
4) B- Correlation
5) C- low bias and high variance
6) B- predictive model
7) D- regularization
8) D- SMOTE
9) D- Recall and Precision
10) A-true
11) B- apply PCA to project high dimensional data
12) B-it becomes slow when the number of features is very large
    C- we need to iterate.

13)Regularization is the process of adding information in order to solve an ill-posed problem or to prevent overfitting.

- When we use regression models to train some data, there is good chance that the model will overfit the given data set. Regularization helps sort this overfitting problem by restricting the degrees of freedom of a given equation that is simply reducing the number of degrees of a polynomial function by reducing their corresponding weights.
- In a linear equation, we do not want huge weights or coefficients as a small change in weight can make a large difference for the dependent variable(Y). So, regularization constraints the weights of such features to avoid overfitting.
- To regularize the model, a shrinkage penalty is added to the cost function.
- **WHY REGULARIZATION?**
- Regularization helps to reduce the variance of the model, without a substantial increase in the bias.
- If there is variance in the model that means that the model won't fit well for dataset different than training data.
- The tuning parameter ʎ controls this bias and variance trade off. When the value of ʎ is increased up to a certain limit, it reduces the variance without losing any important properties in data. But after a certain limit, the model will start losing some important properties which will increase the bias in the data. Thus, the selection of good value of ʎ is the key.
- The value of ʎ is selected using cross-validation method. A set of ʎ is selected and cross-validation error is calculated for each value of ʎ and that value of ʎ is selected for which the cross-validation error is minimum.

13) Different types of regularization in regression are:
    I)Ridge Regression
    II)LASSO (Least Absolute Shrinkage and Selection Operator) Regression
    III)Elastic-Net Regression(less popular)

    The above three algorithms are used for Regularization. But in practical field, only two particular algorithms Ridge Regression and LASSO are used to prevent overfitting.

LASSO Regression (L1 FORM):

LASSO regression penalizes the model based on sum of magnitude of the coefficients. The regularization term is given by;

Regularization= ʎ * $\sum$ | βj |

Where ʎ is the shrinkage factor(learning rate).

Ridge Regression(L2 form):

Ridge regression penalizes the model based on the sum of the squares of magnitudes of the coefficients. The regularization term is given by ;

Regularization= ʎ*$\sum$ | $\beta j2$ |

Difference between Ridge and LASSO regression:

Ridge regression shrinks the coefficients for those predictors which contribute very less in the model but have huge weights, very close to zero. But it never makes them exactly zero. Thus, the final model will still contain all those predictors, though with less weights. This does not help in interpreting the model very well. This is where LASSO regression differs with Ridge regression.

In LASSO regression, the L1 penalty does reduce some coefficients exactly to zero when we use a sufficiently large tuning parameter ʎ. So, in addition to regularizing**, LASSO also performs feature selection.**

14) Error term present in Linear Regression Equation:

Linear regression is one of the most fundamental and widely known Machine Learning Algorithms. Building blocks of LR model are;

- Discreat/ continuous independent variables
- A best-fit regression line
- Continuous dependent variables

The equation of linear regression is;

Y=a + b*X + e, where a is the intercept

b is the slope and e is the error or residual

the above equation is used to predict the value of the target variable based on the given predictor variables.

Error or residual is the distance between the actual Y. The best fit line is obtained by minimizing the residual.