**Machine Learning**

Q1) A – Least square error

Q2) A – Linear Regression is sensitive to outliers

Q3) B – Negative

Q4) A – Regression

Q5) C – Low Bias and High Variance

Q6) B – Predictive modal

Q7) D – Regularization

Q8) D – SMOTE

Q9) C – Sensitivity and Specificity

Q10) B – False

Q11) B – Apply PCA to Project high dimensional data

Multiple Option Correct

Q12) A , B

Q13)

Regularization is a way to avoid  overfitting by penalizing high-valued regression coefficients. In simple terms, it**reduces parameters and shrinks (simplifies) the model.** This more streamlined, more parsimonious model will likely perform better at predictions. Regularization adds penalties to more complex models and then sorts potential models from least overfit to greatest; The model with the lowest “overfitting” score is usually the best choice for predictive power.

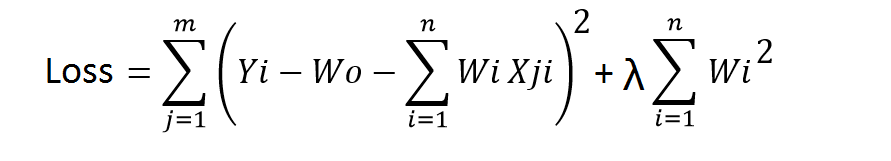
Q14)

There are three main regularization techniques, namely:

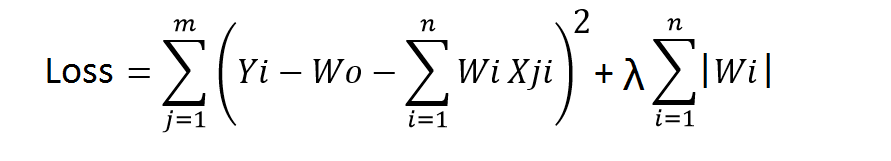
1. Ridge Regression (L2 Norm)
2. Lasso (L1 Norm)
3. Dropout

1.Ridge regression is also called L2 norm or regularization.

When using this technique, we add the sum of weight’s square to a loss function and thus create a new loss function which is denoted thus:



2. Lasso regression is L1 norm and denoted as below:



This technique is different from ridge regression as it uses absolute weight values for normalization. λ is again a tuning parameter and behaves in the same as it does when using ridge regression.

3. Dropout is a regularization technique used in neural networks. It prevents complex co-adaptations from other neurons.

In neural nets, fully connected layers are more prone to overfit on training data. Using dropout, you can drop connections with *1-p* probability for each of the specified layers. Where *p* is called **keep probability** **parameter** and which needs to be tuned.

Q15)

An error term is a residual variable produced by a statistical or mathematical 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 empirical analysis.

The error term is also known as the residual, disturbance, or remainder term, and is variously represented in models by the letters e .

* An error term appears in a statistical model, like a regression model, to indicate the uncertainty in the model.
* The error term is a residual variable that accounts for a lack of perfect goodness of fit.
* Heteroskedastic refers to a condition in which the variance of the residual term, or error term, in a regression model varies widely.

An error term essentially means that the model is not completely accurate and results in differing results during real-world applications. For example, assume there is a multiple linear regression function that takes the following form:

***Y*=*αX*+*βρ*+*ϵ***

**Where :** *α*,*β*=Constant parameters *X* , *ρ*=Independent variables ,

*ϵ* = Error term​﻿

When the actual Y differs from the expected or predicted Y in the model during an empirical test, then the error term does not equal 0, which means there are other factors that influence Y.