<u>MACHINE LEARNING ASSESMENT ANSWERS</u>

Q1> ANS: **(A) Least Square Error**

Q2> ANS**: (A) Linear regression is sensitive to outliers**

Q3> ANS: **(B) Negative**

Q4> ANS: **(B) Correlation**

Q5> ANS: **(C) Low bias and high variance**

Q6> ANS: **(B) Predictive model**

Q7> ANS:**(D) Regularization**

Q8> ANS: **(D) SMOTE**

Q9> ANS: **(A) TPR and FPR**

Q10> ANS: **(B) False**

Q11> ANS: **(B) Apply PCA to project high dimensional data**

Q12>ANS**: A) We don't have to choose the learning rate.**

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

Q13. **Explain the term regularization?**

- Regularization is a technique used to reduce the errors by fitting the function appropriately on the given training set and avoid overfitting. The commonly used regularization techniques are:

1. L1 regularization
2. L2 regularization
3. Dropout regularization

Q14. **Which particular algorithms are used for regularization?**

- **Regularization algorithms**
  - Ridge Regression
  - LASSO (Least Absolute Shrinkage and Selection Operator) Regression
  - Elastic-Net Regression

Ridge Regression

Ridge regression is a method for analyzing data that suffer from multi-collinearity.

$$Loss = \sum_{i=1}^{n} (y_i - (w_i x_i + c))^2 + \lambda \sum_{i=1}^{n} w_i^2$$

Loss Function for Ridge Regression

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 *(λ)* regularizes the coefficients such that if the coefficients take large values, the loss function is penalized.

Ridge regression is also known as the ***L2 Regularization.***

To sum up, **Ridge regression shrinks the coefficients as it helps to reduce the model complexity and multi-collinearity.**

LASSO Regression

LASSO is a regression analysis method that performs both feature selection and regularization in order to enhance the prediction accuracy of the model.

$$Loss = \sum_{i=1}^{n} (y_i - (w_i x_i + c))^2 + \lambda \sum_{i=1}^{n} |w_i|$$

Loss Function for LASSO Regression

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 λ is sufficiently large.

*LASSO regression is also known as the **L1 Regularization (L1 penalty).***
LASSO Regression: Coefficient values if λ = 0.05, and 0.5 respectively ||

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.

$$Loss = \sum_{i=0}^{n} (y_i - (w_i x_i + c))^2 + \lambda_1 \sum_{i=0}^{n} |w_i| + \lambda_2 \sum_{i=0}^{n} w_i^2$$

Loss Function for Elastic-Net Regression

Q15.Explain the term error present in linear regression equation?

The standard error of the regression (S), also known as the standard error of the estimate, represents the average distance that the observed values fall from
the regression line. Conveniently, it tells you how wrong the regression model is on average using the units of the response variable. Smaller values are better because it indicates that the observations are closer to the fitted line.

Unlike R-squared, you can use the standard error of the regression to assess the precision of the predictions. Approximately 95% of the observations should fall within plus/minus 2*standard error of the regression from the regression line, which is also a quick approximation of a 95% prediction interval. If want to use a regression model to make predictions, assessing the standard error of the regression might be more important than assessing R-squared.

Linear regression most often uses mean-square error (MSE) to calculate the error of the model. MSE is calculated by:

1. measuring the distance of the observed y-values from the predicted y-values at each value of x;
2. squaring each of these distances;
3. calculating the mean of each of the squared distances.

Linear regression fits a line to the data by finding the regression coefficient that results in the smallest MSE.

mean-square error (MSE)

The Mean Squared Error (MSE) is a measure of how close a fitted line is to data points. For every data point, you take the distance vertically from the point to the corresponding y value on the curve fit (the error), and square the value. Then you add up all those values for all data points, and, in the case of a fit with two parameters such as a linear fit, divide by the number of points minus two.** The squaring is done so negative values do not cancel positive values. The smaller the Mean Squared Error, the closer the fit is to the data. The

MSE has the units squared of whatever is plotted on the vertical axis.

Another quantity that we calculate is the Root Mean Squared Error (RMSE). It is just the square root of the mean square error. That is probably the most easily interpreted statistic, since it has the same units as the quantity plotted on the vertical axis.