Project - 3

Machine learning

Answer sheet

Ans.1 – (a) least square error

Ans.2 – (b) linear regression is not sensitive to outliers

Ans.3 – (b) negative

Ans.4 – (a) regression

Ans.5 – (c) low bias and high variance

Ans.6 – (b) predictive model

Ans.7 – (d) regularization

Ans.8 – (a) cross validation

Ans.9 – (a) TPR AND FPR

Ans.10-(b) false

Ans.11–(b) apply PCA to project high dimensional data

Ans.12-(a, b, and c)

Subjective type questions –

Ans.13 – regularization is a set of methods for reducing overfitting in machine learning models. Typically , regularization trades a marginal decrease in training accuracy for an increase in generalizability.

Ans.14 – elastic net (lesso regression + ridge regression) algorithm are used for regularization.

Ans.15 – in a linear regression equation , an error term represents the difference between the observed value of the dependent variable (y) and the predicted value of the dependent variable based on the regression line.

The error term is denoted by $\grave{\varepsilon}$ (epsilon) and is added to the linear regression equation as follows.

$Y = \beta_0 + \beta_1 x + \grave{\varepsilon}$

Where :-

- Y is the dependent variable
- X is the independent variable
- β_0 is the intercept or constant term
- β_1 is the slope coefficient
- $\grave{\varepsilon}$ is the error term

the error term accounts for the variability in the data that is not explained by the linear relationship between x and y. it presents the random fluctuations or noise in the data.