Machine Learning

Answer from 1 to 11

Ans1 - Least Square Error
Ans2 - Linear regression is sensitive to outliers
Ans3 - Negative
Ans4 - Correlation
Ans5 - Low bias and high variance
Ans6 - Descriptive model
Ans7 - Regularization
Ans8 - Regularization
Ans9 - TPR and FPR
Ans10- False
Ans11- Apply PCA to project high dimensional data
Ans12- A- We don't have to choose the learning rate
B- It becomes slow when the number of features is very large

Answer from 13 to 15

Ans13 - Regularization is one of the most important concepts of machine learning. It is a technique to prevent the model from overfitting by adding extra information to it.

Sometimes the machine learning model performs well with the training data but does not perform well with the training data but does not perform well with the test data.

Regularization is a technique used to reduce errors by fitting the function appropriately on the given training set and avoiding overfitting.

Ans14 - Algorithms used for regularization

- **Ridge Regression** - Its purpose is to overcome problems such as data overfitting and multicollinearity in data. Ridge regression adjusts the variables by a modest squared bias factor. Ridge regression decreases the two coefficients towards each other when the variables are highly linked.

- **LASSO** - It simply penalizes Large coefficients, in contrast to Ridge Regression. When the hyperparameter is big enough, Lasso has the effect of driving certain coefficient estimations to be absolutely zero. If we have a large number of features, LASSO works effectively for feature selection.

Ans15 - In a regression analysis, the error term also known as the residual represents the difference between the observed value of the dependent variable and the value predicted by the regression model.

It captures the variability in the dependent variable that is not explained by the independent variables included in the model. The error term is an important component of regression analysis as it reflects the presence of unobserved factors or random variation that affect the dependent variable.

It is typically assumed to follow certain statistical properties, such as being normally distributed with a mean of zero.