Machine Learning

1. D) Both A and B
2. A) Linear regression is sensitive to outliers
3. B) Negative
4. C) Both of them
5. C) Low bias and high variance
6. B) Predictive modal
7. D) Regularization
8. C) Kernel
9. A) TPR and FPR
10. B) False
11. A) Construction bag of words from a email

    B) Apply PCA to project high dimensional data
12. A) We don't have to choose the learning rate.

    D) It does not make use of the dependent variable.
13. Regularization techniques that are used to calibrate machine learning models in order to minimize the adjusted loss function and prevent overfitting or under fitting. Using Regularization, we can fit our machine learning model appropriately on a given test set and hence reduce the errors in it. There are two main types of regularization techniques:
    - Ridge Regularization : with the L2 norm
    - Lasso Regularization : having the L1 norm

    - Ridge Regression, modifies the over-fitted or under fitted models by adding the penalty equivalent to the sum of the squares of the magnitude of coefficients. It means, the mathematical function representing our machine learning model is minimized and coefficients are calculated. The magnitude of coefficients is squared and added. Ridge Regression performs regularization by shrinking the coefficients present.

    - Lasso regression modifies the over-fitted or under-fitted models by adding the penalty equivalent to the sum of the absolute values of coefficients. It also performs coefficient minimization, but instead of squaring the magnitudes of the coefficients, it takes the true values of coefficients. This means that the coefficient sum can also be 0, because of the presence of negative coefficients.

Regularization is widely used in linear regression, logistic regression, and other machine learning algorithms to improve model robustness and enhance generalization to new, unseen data.

14. Regularization can be applied to various machine learning algorithms. That's are:

i. Linear Regression with L1/L2 Regularization (Lasso/Ridge)
ii. Logistic Regression with L1/L2 Regularization
iii. Support Vector Machines (SVM)
iv. Neural Networks
v. Decision Trees and Random Forests
vi. Elastic Net

These algorithms could be implemented with regularization to control the complexity of the models and prevent overfitting.

15. In linear regression, the term "error" refers to the difference between the actual observed values (dependent variable) and the values predicted by the linear regression model. This difference was denoted as the residual or error term.

The linear regression equation is literally expressed as:

$$Yi = \beta 0 + \beta 1Xi + €i$$

- Y = is the observed or dependent variable for the ith observation
- Xi = is the independent variable for the ith observation
- β0 = is the intercept (y-intercept)
- β1 = is the slope of the regression line
- €i = is the error term for the ith observation

The error term (€i) represent the part of the observed values that is not explained by the linear relationship between the independent variable and the dependent variable. In other words, it captures the discrepancy between the actual data points and the values predicted by the regression model.