Machine learning

1. A) Least square error
2. A) Linear regression is sensitive to outliers
3. B) Negative
4. A) Regression
5. C) Low bias and high variance
6. B) Predictive model
7. D) Regularization
8. D) SMOTE
9. A) TPR and FPR
10. A) True
11. B) applying PCA to project high dimensional data
12. A) We don't have to choose the learning rate. B) It becomes slow when the number of features is very large.
13. Regularization is nothing but adding extra information (penalty) to the data set to prevent over fitting.

    Regularization try to keep the model simple and prevent it from learning to much from the noise in the data.
14. Algorithms used for Regularization are:
    a. Lazzo (L1) - eliminate the unimportant features
    b. Ridge (L2) - reduces their impact without completely eliminating them
    c. Elastic net - it's a combination of both L1 and L2

    Need to determine the regularization parameter - the value that control the impact of the regularization term on the loss function, a higher value increase the influence of the regularization term thus increasing the simplicity of the model and reduces the risk of over fitting.
15. The term error indicates uncertainty in the model, it is the difference between what the model is predicting and the actual value. This can range from being relatively small to huge, even within one model, across the observed data points.