MACHINE LEARNING

1. D) Both A and B
2. A) Linear regression is sensitive to outliers
3. B) Negative
4. B) Correlation
5. C) Low bias and high variance
6. B) Predictive model
7. D) Regularization
8. D) SMOTE
9. A) TPR and FPR
10. B) False
11. B) Apply PCA to project high dimensional data
12. A) We don't have to choose the learning rate.
    B) It becomes slow when number of features is very large.

13. Regularization refers to techniques used in machine learning to prevent overfitting and improve the generalization performance of models. The goal of regularization is to find a balance between fitting the training data well i.e. reducing bias and avoiding overfitting i.e. reducing variance Overfitting occurs when a model learns the noise or random fluctuations in the training data and as a result performs poorly on unseen or new data. Regularization helps to prevent overfitting by adding a penalty term to the loss function which term discourages the model from fitting too closely to the training data and encourages it to learn more generalizable patterns. There are two popular types of regularization techniques – L1 regularization or Lasso and L2 regularization or Ridge.
    L1 regularization adds the sum of absolute values of the coefficients to the loss function as a penalty term, , promoting sparsity and feature selection. L2 regularization adds the sum of squared values of the coefficients, as a penalty term, promoting smaller and more evenly distributed coefficients.
    Regularization helps to improve model performance by controlling model complexity, especially when dealing with limited training data or high-dimensional feature spaces.

14. There are two types of algorithms primarily used for regularization – L1 and L2.
    L1 or Lasso regularization adds the absolute values of the coefficients as a penalty term. It promotes sparsity and feature selection making some coefficients become exactly zero.

L2 or Ridge regularization adds the squared values of the coefficients as penalty term. This promotes smaller and more evenly distributed coefficients
There is a third type of regularization known as Elastic Net regularization. It combines both L1 and L2 penalties to create a balance between sparsity and also include correlated features.
Regularization can be applied to various machine learning algorithms like Linear Regression, Logistic Regression, SVMs, Neural Networks, Decision Trees, Random Forests and Gradient Boosting to prevent overfitting and improve the performance of the models.

15. In the context of linear regression, the term “error” is the difference between the predicted values and the actual observed values of the dependent variable or target variable. It represents the variability or noise in the data that cannot be captured by the linear relationship between the independent variables and the dependent variable.
The goal of linear regression is to find the best-fitting line that minimises the sum of squared errors. The error term is denoted by ε.
$\varepsilon = y_{obs} - y_{pred}$
where,
ε = error
y_{obs} = actual observed value of the dependent variable
y_{pred} = predicted value of the dependent variable based on the linear regression model
In general, the linear regression model assumes that the errors are normally distributed with a mean of zero and constant variance. The model is considered to be a good fit if the errors are normally distributed.