DS_2303	Machine Learning	Assignment 39

1.	A) Least Square Error
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.	A) Cross validation
8.	A) Cross validation
9.	A) TPR and FPR
10.	A) True
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 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: 1) Ridge Regularization 2) Lasso Regularization. ▪ A regression model which uses L1 Regularization technique is called **LASSO** (Least Absolute Shrinkage and Selection Operator) regression. ▪ A regression model that uses L2 regularization technique is called Ridge regression. ▪ A regression model which uses L1 Regularization technique is called LASSO(Least Absolute Shrinkage and Selection Operator) regression. ▪ A regression model that uses L2 regularization technique is called Ridge regression. ▪ Lasso Regression adds absolute value of magnitude of coefficient as penalty term to the loss function (L).
14.	1) Regularization is a technique used to reduce the errors by fitting the function appropriately on the given training set and avoid overfitting. The commonly used regularization techniques are : ▪ L1 regularization ▪ L2 regularization ▪ Dropout regularization 2) A regression model which uses L1 Regularization technique is called **LASSO** (Least Absolute Shrinkage and Selection Operator) regression. 3) A regression model that uses L2 regularization technique is called Ridge regression.

	4) A regression model which uses L1 Regularization technique is called LASSO (Least Absolute Shrinkage and Selection Operator) regression. 5) A regression model that uses L2 regularization technique is called Ridge regression. 6) Lasso Regression adds absolute value of magnitude of coefficient as penalty term to the loss function (L). 7) Lasso is an acronym for least absolute shrinkage and selection operator, and lasso regression adds the "absolute value of magnitude" of the coefficient as a penalty term to the loss function. $$\sum_{i=1}^{n}(Y_i - \sum_{j=1}^{p} X_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{p} \lvert\beta_j\rvert$$ 8) Again, if lambda is zero, then we'll get back OLS (ordinary least squares) whereas a very large value will make coefficients zero, which means it will become under fit. 9) Ridge regression adds the "squared magnitude" of the coefficient as the penalty term to the loss function. The highlighted part below represents the L2 regularization element. $$\sum_{i=1}^{n}(y_i - \sum_{j=1}^{p} x_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{p} \beta_j^2$$ 10) If lambda is 0, we can assume we get back OLS. 11) However, if lambda is really large, it will add too much weight and result in under fitting. However, how we chose lambda is critical. This method works quite well for avoiding overfitting concerns. 12) The main distinction between both strategies is that lasso reduces the coefficient of the less significant characteristic to zero, eliminating certain features entirely. 13) In other words, when we have a large number of features, L1 regularization is effective for feature selection.
15.	▪ The supervised machine learning approach known as "linear regression" identifies the linear relationship between the dependent and independent variables by determining the best fit linear line between them. ▪ A statistical technique called regression analysis can be used to test the claim that a variable is reliant on one or more other variables. ▪ Regression analysis can also give a rough idea of how much a change in one variable would affect another. ▪ This final characteristic is crucial for predicting future values, of course. ▪ For assessing and summarizing the effectiveness of a regression model, three error metrics are frequently employed. ▪ These include: MSE, or mean square error. RMSE (Root Mean Squared Error). MAE, or Mean Absolute Error

	▪ Because independent variables are never ideal predictors of the dependent variables in reality, a regression line will always contain an error term. ▪ Standard error of the regression = **(SQRT (1 minus adjusted-R-squared)) x STDEV. S(Y).** ▪ So, for models fitted to the same sample of the same dependent variable, adjusted R-squared always goes up when the standard error of the regression goes down ▪ Instead, the line represents an estimation based on the data at hand. ▪ Therefore, the error term informs you of your level of confidence in the formula. ▪ A Linear Regression model's main aim is to find the best fit linear line and the optimal values of intercept and coefficients such that the error is minimized. ▪ Error is the difference between the actual value and Predicted value and the goal is to reduce this difference.

****