

# **MACHINE LEARNING**

**Q1 to Q11, only one option is correct, choose the correct option:**

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 modal
7. (D) Regularization
8. (D) SMOTE
9. (C) Sensitivity and Specificity
10. (A) True
11. (A) Construction bag of words from a email  
(B) Apply PCA to project high dimensional data  
(C) Removing stop words
12. (A) We don't have to choose learning rate  
(B) It becomes slow when number of features is very large.  
(C) We need to iterate

**Q13 and Q15 are subjective answer type questions, Answer them briefly.**

**13.)** Regularization is a technique used to reduce the errors by fitting the function appropriately on the given training set and avoid overfitting. Regularization refers to the modifications that can be made to a learning algorithm that helps to reduce this generalization error and not the training error. It reduces by ignoring the less important features. It also helps prevent overfitting, making the model more robust and decreasing the complexity of a model.

**14.)** The particular algorithm are used for regularization are:-

- Ridge Regression

The Ridge regression technique is used to analyze the model where the variables may be having multicollinearity. It reduces the insignificant independent variables though it does not remove them completely. This type of regularization uses the L2 norm for regularization.

- LASSO (Least Absolute Shrinkage and Selection Operator) Regression

Least Absolute Shrinkage and Selection Operator (or LASSO) Regression penalizes the coefficients to the extent that it becomes zero. It eliminates the insignificant independent variables. This regularization technique uses the L1 norm for regularization.

- Elastic-Net Regression

The Elastic Net Regression technique is a combination of the Ridge and Lasso regression technique. It is the linear combination of penalties for both the L1-norm and L2-norm regularization. The model using elastic net regression allows the learning of the sparse model where some of the points are zero, similar to Lasso regularization, and yet maintains the Ridge regression properties. Therefore, the model is trained on both the L1 and L2 norms.

**15.)** The error term in the linear regression model is the difference between the expected value at a particular time and the value that was actually observed. In instances where the value is exactly what was anticipated at a particular time, the value will fall on the trend line and the error term will be zero. The error in linear regression equation are what your model estimates may not match what is actually observed. If your model specification, your decision to use linear regression, is a good one then the difference between your model estimate and

what is actually observed is small. This difference is the error. It is used to account for the difference between what is observed and what your model estimates.