ASSIGNMENT-39

MACHINE LEARNING

ANSWERSHEET

Q1. (A) Least square Error

Q2. (A) Linear regression is Sensitive to outliers

Q3. (B) Negative

Q4. (B) Correlation

Q5. (D) None of these

Q6. (B) Predictive modal

Q7. (B) Regularization

Q8. (D) SMOTE

Q9. (A) TPR AND FPR

Q10. (False)

Q11. (A)Construction bag of words from a email

(B)Apply PCA to project high dimensional data

(C)Removing stop words

Q12. (A) We don't have to choose the learning rate.

(B) It becomes slow when number of features is very large.

(C) We need to iterate.

Q13. Regularization is a technique used in regression to reduce the complexity of the model and to shrink the coefficients of the independent features. This technique converts a complex model into a simpler one, so as to avoid the risk of overfitting and shrinks the coefficients, for lesser computational cost.

Q14. Ridge Regression

LASSO (Least Absolute Shrinkage and Selection Operator) Regression

Elastic-Net Regression

Q15. It is often said that the error term in a regression equation represents the effect of the variables that were omitted from the equation. This is unsatisfactory, even in simple contexts, as the following discussion should indicate. Suppose subjects are IID, and all variables are jointly normal with expectation 0. Suppose the explanatory variables have variance 1. The explanatory variables may be correlated amongst themselves, but anyp of them have a non-singularp-dimensional distribution