MACHINE LEARNING ASSIGNMENT - 5

Q1 to Q15 are subjective answer type questions, Answer them briefly.

1. R-squared or Residual Sum of Squares (RSS) which one of these two is a better measure of goodness of fit model in regression and why?

Answer: R-squared and Residual Sum of Squares (RSS) both are commonly used measures of goodness of fit in regression analysis. R-squared provides a measure of overall goodness of fit and explanatory power, RSS offers insight into the residual errors of the model and its predictive accuracy. Both measures are important and can be used together to gain a comprehensive understanding of the performance of a regression model.

2. What are TSS (Total Sum of Squares), ESS (Explained Sum of Squares) and RSS (Residual Sum of Squares) in regression. Also mention the equation relating these three metrics with each other.

3. What is the need of regularization in machine learning?

Answer: Regularization is a technique used in machine learning to prevent overfitting and improve the generalization ability of a model.

Need of regularization: - Regularization helps to the issue of overfitting occurs when a model learns to fit the training data too closely

4. What is Gini–impurity index?

5. Are unregularized decision-trees prone to overfitting? If yes, why?

6. What is an ensemble technique in machine learning?

7. What is the difference between Bagging and Boosting techniques?

8. What is out-of-bag error in random forests?

9. What is K-fold cross-validation?

10. What is hyper parameter tuning in machine learning and why it is done?

11. What issues can occur if we have a large learning rate in Gradient Descent?

12. Can we use Logistic Regression for classification of Non-Linear Data? If not, why?

13. Differentiate between Adaboost and Gradient Boosting.

14. What is bias-variance trade off in machine learning?

15. Give short description each of Linear, RBF, Polynomial kernels used in SVM.