**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?

**Ans-**RSS measures the level of variance in the error term or residuals of a regression model.Rss is just the absolute amount of explained variation.Rsqured is the absolute amount of variation as a proportion of total variation

1. 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.

**Ans-**TSS tells us how much variation there is in dependent variable.ESS is a statistical quantity used in modelling of a process.ESS gives an estimate of how well a model explains the observed data for the process.RSS is the measure of discrepancy between the data and an estimation model such as linear regression.

R^2 =MSS/TSS=(TSS-RSS)/TSS

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

**Ans-**Regularization is a technique used to reduce the errors by fitting the functions appropriately on the giving training set and avoid overfitting.The commonly used regularization techniques are L1 and L2 regularization.

1. What is Gini–impurity index?

**Ans-**Gini impurity measures the degree or probability of a particular variable being wrongly classified when it is randomly chosen.It is one of the methods used in decision tree algorithms to decide the optimal split from a root node,and subsequent splits

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

**Ans-**Decision trees are prone to overfitting ,especially when a tree is particularly deep.This is due to the amount of specificity we look at leading to smaller sample of events that meet the previous assumptions.This small sample could lead to unsound conclusions.

1. What is an ensemble technique in machine learning?

**Ans-**Ensemble methods are techniques that create multiple models and then combine them to produce improved results.Ensemble methods usually produce more accurate solutions than a single model would .

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

**Ans-**Bagging is a method of merging the same type of predictions.Boosting is a method of merging different types of predictions.Bagging decreases variance,not bias and solves overfitting issues in a model.Boosting decreases bias,not variance.

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

**Ans-**Out of bag error is the average error for each calculated using predictions from the trees that donot contain in their respective bootstrap sample.

9. What is K-fold cross-validation?

**Ans-**Cross-validation is a resampling procedure used to evaluate machine learning models on a limited data sample.Tht is to use a limited sample inorder to estimate how the model is expected to perform in general when used to make predictions on data not used during the training of the model.

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

**Ans-**Hyperparameter tuning is the problem of choosing a set of optimal hyperparameters for a learning algorithm.Hyperparameter is a parameter whose value is used to control the learning process.It is crucial as they control the overall behaviour of a machine learning model.Every machine learning models have different hyperparameters that can be set.

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

**Ans-**When the learning rate is too large gradient descent can inadvertently increase rather than decrease the training error.When the learning rate is too small ,training is not only slower but may become permanently stuck with a high training error.

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

**Ans-**Logistic regression is known and used as a linear classifier.It is used to comeup with a hyperplane in feature space to separate observations that belongs to a class from all the other observations that donot belong to that class.The decision boundary is thus linear.

13. Differentiate between Adaboost and Gradient Boosting.

**Ans-**Adaboost is the first designed boosting algorithm with a particular loss function.Gradient boosting is a generic algorithm that assist in searching the approximate solution to the additive modelling problem.This makes gradient boosting more flexible than adaboost

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

**Ans-**Bias is the simplified assumptions made by the model to make the target function easier to approximate.variance is the amount that the estimate of the target function will change given different training data.Trade-off is tension between the error introduced by the bias and the variance.

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

**Ans-**Linear kernel is used when the data is linearly separable.It can be separated using single line.It is used when there are large number of features in a particular dataset.Radial bias function kernel is used in various kernelized learning algorithms.It is similar to k-nearest neighbourhood algorithm.It has the advantage of KNN and overcomes the space complexity problem as it needs to store support vectors during training and not the entire dataset.polynomial kernel function represents the similarity of vectors in a feature space over polynomials of the original variables,allowing learning of non linear models.