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

Ans:-Linear kernel :- Linear kernel is used when we have to separate the data linearly.

It is used for text classification. It is one of the most common kernels to be used.

Training a SVM with a Linear Kernel is Faster than with any other Kernel.

RBF Kernels:- In machine learning, the radial basis function kernel, or RBF kernel, is a popular kernel function used in various kernelized learning algorithms. In particular, it is commonly used in support vector machine classification.

Polynomial kernels:-  the polynomial kernel is a kernel function commonly used with support vector machines (SVMs) and other kernelized models, that represents the similarity of vectors (training samples) in a feature space over polynomials of the original variables, allowing learning of non-linear models.

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

Ans:- R-squared evaluates the scatter of the data points around the fitted regression line. It is also called the [coefficient](https://statisticsbyjim.com/glossary/regression-coefficient/) of determination, or the coefficient of multiple determination for multiple regression. For the same data set, higher R-squared values represent smaller differences between the observed data and the fitted values.

The [Residuals](https://statisticsbyjim.com/glossary/residuals/) versus Fits plot emphasizes this unwanted pattern. An unbiased model has residuals that are randomly scattered around zero. Non-random residual patterns indicate a bad fit despite a high R2. Always check your residual plots!

RSS is good is better measure of goodness of fit of model in regression .

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:- The coefficient of determination is used as a measure of how well a regression line explains the relationship between a dependent variable (Y) and an independent variable (X). The closer the coefficient of determination is to 1, the more closely the regression line fits the sample data.

**formula**: R2 = MSS/**TSS** = (**TSS** − RSS)/**TSS**

ESS:-  the explained sum of squares (ESS), alternatively known as the model sum of squares or sum of squares due to regression ("SSR" – not to be confused with the residual sum of squares RSS or sum of squares of errors), is a quantity used in describing how well a model, often a regression model.

RSS:- A residual sum of squares (RSS) is a statistical technique used to measure the amount of variance in a data set that is not explained by a regression model. ... The residual sum of squares measures the amount of error remaining between the regression function and the data set

What is Gini –impurity index?

Ans:- Gini index or Gini impurity measures the degree or probability of a particular variable being wrongly classified when it is randomly chosen. ... A Gini Index of 0.5 denotes equally distributed elements into some classes.

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

What is an ensemble technique in machine learning?

Ans:-Ensemble learning combines the predictions from multiple neural network models to reduce the variance of predictions and reduce generalization error. Techniques for ensemble learning can be grouped by the element that is varied, such as training data, the model, and how predictions are combined.

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

Ans:-Bagging is a way to decrease the variance in the prediction by generating additional data for training from dataset using combinations with repetitions to produce multi-sets of the original data. Boosting is an iterative technique which adjusts the weight of an observation based on the last classification.

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

Ans:-Out-of-bag (OOB) error, also called out-of-bag estimate, is a method of measuring the prediction error of random forests, boosted decision trees, and other machine learning models utilizing bootstrap aggregating (bagging) to sub-sample data samples used for training.

9. What is K-fold cross-validation?

Ans:- K-Fold CV is where a given data set is split into a *K* number of sections/folds where each fold is used as a testing set at some point. Lets take the scenario of 5-Fold cross validation(K=5). Here, the data set is split into 5 folds. In the first iteration, the first fold is used to test the model and the rest are used to train the model. In the second iteration, 2nd fold is used as the testing set while the rest serve as the training set. This process is repeated until each fold of the 5 folds have been used as the testing set.

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

Ans:- In machine learning, hyperparameter optimization or tuning is the problem of choosing a set of optimal hyperparameters for a learning algorithm. A hyperparameter is a parameter whose value is used to control the learning process. By contrast, the values of other parameters (typically node weights) are learned.

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

Ans:- In machine learning, hyperparameter optimization or tuning is the problem of choosing a set of optimal hyperparameters for a learning algorithm. A hyperparameter is a parameter whose value is used to control the learning process. By contrast, the values of other parameters (typically node weights) are learned

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

Ans:-Bias is the simplifying 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.

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

Ans:-Regularisation is a technique used to reduce the errors by fitting the function appropriately on the given training set and avoid overfitting. .

14. Differentiate between Adaboost and Gradient Boosting

Ans:- Adaboost is more about 'voting weights' and gradient boosting is more about 'adding gradient optimization'. Adaboost doesn't overfit because it is more about 'organizing people to vote' than 'voting'. In fact, if you have a gradient boosting model, you can use it in adaboost along with other models.

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

Ans:-Logistic regression has traditionally been used to come up with a hyperplane that separates the feature space into classes. But if we suspect that the decision boundary is nonlinear we may get better results by attempting some nonlinear functional forms for the logit function.