RandomForest Algorithm.

- Random Forest is type of machine learning algorithm based on ensemble learning.Ensemble learning is a type of learning where we joint deffrent type of algorithm or same algorithm mutiple times to form more powerfull prediction model.
- Random Forest combines multiple algorithm for the same time that means mutiple decison tree resulting in forest of tree hence the name of random forest.
- Random forest algorithm can be used for both regression and classification task.
- Random Forest has variety of applications,such as recommendation engine,image classification and feature selection.
- It can be used for classify the loyal loan applications,identify fraududent activity and predict diseases.It lies at the base of Borula algorithm,which select the important features in the dataset.
- In the random Forest,The tree will work on the same or diffrent algorthm simultaeously for same dataset on the same or dirrent classes and then finally combine their result and monetoring it's result by dirrent techniques:-
- In that we will be use the majority voting techinique and avraging based on that we will be make final class.
- Collection of the trees = Forest
- For the clasification whoever given output will give us more vote or popular that would be final result or final feature.
- For the regression we will consider the average of all the tree output.

How the Random Forest Algorithm work ?

- So,the techinically it is a ensemble learning based on the divide and conorr approach and it divides the decision tress in so many decison trees then it finally concorr those trees and then it's do the ensembling based on the outcome of the those trees.
- The collection of decision tree classfier is known as forest.
- Individual decison tree are generated based on attribute selection indicator which is also feature and based on the attribute selection measures (gini ratio,gini index, information gain) based on those it makes the forest of the trees and each tree depends on the independent random sample..
- In the classificcation problem each tree vote and most popular class is choosen as final result.
- In the case of the regression,average of all tree output is result consider as final result.
- Simple and powerfull as compare other non-linear classification.

Stpes Of RandomForestClassifier or Regressor

- Select the random sample from the given datasets.
- Construct the decision tree for each sample and get a prediction result from each decision tree.
- Perform a vote for each predicted result.
- Select the prediction result with most votes as final prediction.

Important Feature For Classification.

- Randomforest can identify even single important in our datasets.Randomforest offer these selection indicators.
- RandomForest uses the Gini Importance or mean decresase in impurity (MDI) to calculate the importance of each feature.Gini important is also known as the total decrease in node impurity.This is how the model is fit or accuracy decreases when you drop a variable.The larger decrease the more signoficant variable is.Here,the mean decrease is a significant parameter for variable selection.The Gini index can describe the overall explanatory power of the variables.
    - RandomForest is the the algorirthm calculate the contribution each feature in prediction by using the feature importances_.
    - This is helps to get the relevance of the datasets features.

Random Forests Vs Decision Trees

- Random Forest is set of mutiple decision tree.
- Decision tree is makes the low bised and high variance prediction
- Wherease the RandomForest create the low biased and low variance in the prediction.
- It uses the parellel RandomForest and each tree get the data in sample which include some sort of columns and rows and then make the prediction with replacement.
- Deep decison tree may suffer from overfitting,but random forest prevents overfitting by creating tress on random subsets.
- Decision trees are coputationaly faster.
- Random forest is difficult to interpret,while a decision tree is easily interpretable and can be converted to rules.
- Under the presency of the outliers the model of decision tree gor overfitted but RandomForestClassifier will not overfit the model under the presency of the outliers because it is robust to the outliers.

Advantages of the Decision Tree.

- Randomforest is slow in prediction and it has multiple decision trees because in the lost of decision trees are working together and that's it takes the lot of time for prediction or we can say it is time consuming.
- This model is difficult to interprete if we compare to the other model.