

Random forest prediction pseudocode:

To perform prediction using the trained random forest algorithm uses the below pseudocode.

Takes the test features and use the rules of each randomly created decision tree to predict the outcome and stores the predicted outcome (target)

Calculate the votes for each predicted target.

Consider the high voted predicted target as the final prediction from the random forest algorithm.

To perform the prediction using the trained random forest algorithm we need to pass the test features through the rules of each randomly created trees. Suppose lets say we formed 100 random decision trees to from the random forest.

Each random forest will predict different target (outcome) for the same test feature. Then by considering each predicted target votes will be calculated. Suppose the 100 random decision trees are prediction some 3 unique targets x, y, z then the votes of x is nothing but out of 100 random decision tree how many trees prediction is x.

Likewise for other 2 targets (y, z). If x is getting high votes. Lets say out of 100 random decision tree 60 trees are predicting the target will be x. Then the final random forest returns the x as the predicted target.

This concept of voting is known as majority voting.

Now lets look into few applications of random forest algorithm.

Random Forest Applications

Random Forest Applications

The random algorithm used in wide varieties applications. In this article, we are going address few of them.

Below are some the application where random forest algorithm is widely used.

Banking Medicine Stock Market E-commerce Lets begin with the banking sector.

1.Banking:

In the banking sector, random forest algorithm widely used in two main application. These are for finding the loyal customer and finding the fraud customers.

The loyal customer means not the customer who pays well, but also the customer whom can take the huge amount as loan and pays the loan interest properly to the bank. As the growth of the bank purely depends on the loyal customers. The bank customers data highly analyzed to find the pattern for the loyal customer based the customer details.

In the same way, there is need to identify the customer who are not profitable for the bank, like taking the loan and paying the loan interest properly or find the outlier customers. If the bank can identify theses kind of customer before giving the loan the customer. Bank will get a chance to not approve the loan to these kinds of customers. In this case, also random forest algorithm is used to identify the customers who are not profitable for the bank.

2.Medicine

In medicine field, random forest algorithm is used identify the correct combination of the components to validate the medicine. Random forest algorithm also helpful for identifying the disease by analyzing the patients medical records.

3. Stock Market

In the stock market, random forest algorithm used to identify the stock behavior as well as the expected loss or profit by purchasing the particular stock.

4. E-commerce

In e-commerce, the random forest used only in the small segment of the recommendation engine for identifying the likely hood of customer liking the recommend products base on the similar kinds of customers.

Running random forest algorithm on very large dataset requires high-end GPU systems. If you are not having any GPU system. You can always run the machine learning models in cloud hosted desktop. You can use clouddesktoponline platform to run high-end machine learning models from sitting any corner of the world.

Advantages of Random Forest algorithm

Below are the advantages of random forest algorithm compared with other classification algorithms.

The overfitting problem will never come when we use the random forest algorithm in any classification problem. The same random forest algorithm can be used for both classification and regression task. The random forest algorithm can be used for feature engineering. Which means identifying the most important features out of the available features from the training dataset.

Compared with other classification techniques, there are three advantages as the author mentioned.

For applications in classification problems, Random Forest algorithm will avoid the overfitting problem For both classification and regression task, the same random forest algorithm can be used The Random Forest algorithm can be used for identifying the most important features from the training dataset, in other words, feature engineering.