RPubs *by RStudio*

Random Forest prediction model in R

Random Forest

Intuition Lecture 126 https://www.udemy.com/machinelearning/learn/lecture/5714412 (https://www.udemy.com/machinelearning/learn/lecture/5714412)

Lecutre 129 https://www.udemy.com/machinelearning/learn/lecture/5771094 (https://www.udemy.com/machinelearning/learn/lecture/5771094)

Decision tree algorithms are about splitting the data into classifications to then have an algorithm that will predict where new points of data will land. Those classifications are based on values of the independent and dependent variables.

Random forests are about having multiple trees, a forest of trees. Those trees can all be of the same type or algorithm or the forest can be made up of a mixture of tree types (algorithms). There are some very interesting further metaphorical thoughts that describe how the forest acts (decides).

Again as with Decision Trees the Random Forest is not based on euclidian distances but rather classifications.

Check Working directory getwd() to always know where you are working.

Importing the dataset

we are after the age and salary and the y/n purchased so in R that's columns 3-5

```
dataset = read.csv('Social_Network_Ads.csv')
dataset = dataset[3:5]
```

Have a look at data

```
summary(dataset)
```