Ensemble learning and random forest

Random forest is the version of ensemble learning

Gradient boosting is another version of ensemble learning

Ensemble learning:

In this I took multiple algorithms or the same algorithm multiple times and put them together which makes much more powerful than the original

Step 1: pick random K data points form the training set

we use the Dataset we have and we only pick K data points from that training set

I build a decision tree associated to these K data points rather than building the decision tree based on everything I have done

It’s a subset of the dataset

I then choose the number of trees I want to build and keep repeating the steps

This builds a lot of regression decision tress

finally, use all of the trees to predict so far

datapoint makes each one of the entries, predict the value of y for the data point which is in question and assign the new data point

Instead of one prediction we get a lot of predictions by default

Dataset:

Here, I am using dataset of different positions and levels going from 1 to 10 and salary going from 45000 dollars to 1000000 dollars

**Dataset**

For this we use Position and salary dataset

We are using simple dataset

**Scenario:**

The Human Resources department wants to hire someone

A person is offered a position with tedious interviews

Now we have to determine the salary expectation

Now the person is asking 160000 $

Now the person says, that I have earned the same in my previous position

The company need to at least meet the previous salary

The HR department wants us to determine if the statement is true or bluff

This will be determined using polynomial regression model to build previous salary of the candidate

Dataset: salaries of different companies from business analyst to CEO

Compare the salaries of different positions to this person designation and company

**Assumptions:**

**According to the person Resume and Profile:**

On LinkedIn it shows the person has been a regional manager

this person has been manager for past 2 years

In this, now depending on person experience and designation the company considered the **level of person to be 6.5**