29 February Saturday
2020 Wk-9 • Day (060-306)

machine learning

Why → To make new predictions.
Not for finding old ones.

How →

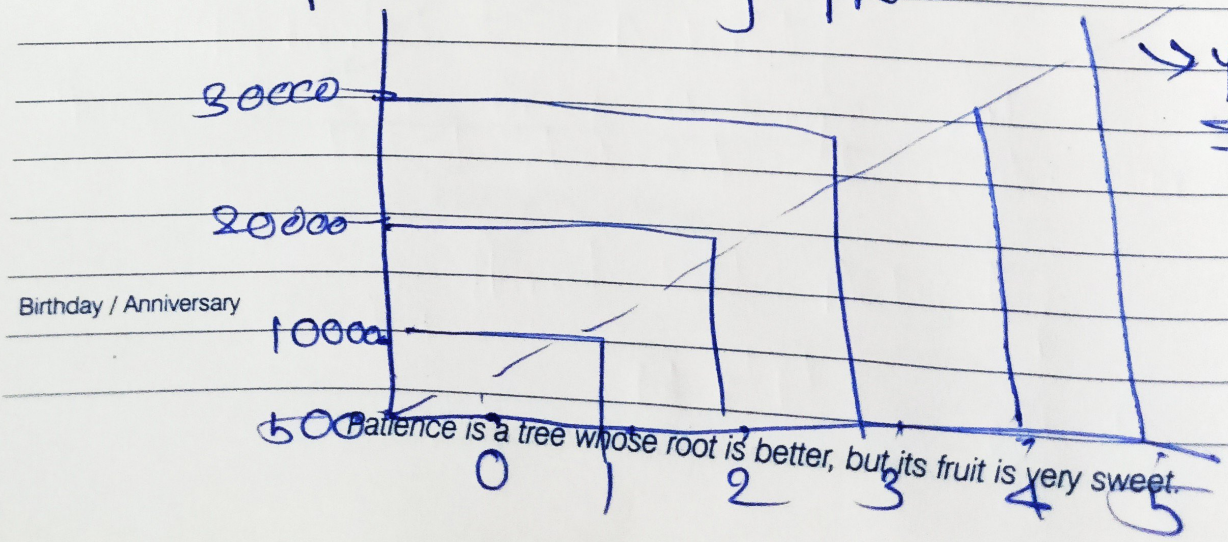
Years of experience	Salary
0	500
1	10000
2	20000
3	30000
4	40000
5	50000
10	100000

We see a pattern.
Rather a formula,

Years of experience * 10000 = salary
except for 0 y.o.e

01 Sunday

If we put this to graph



Birthday / Anniversary

Patience is a tree whose root is better, but its fruit is very sweet.

April 2020

Wk	Su	Mo	Tu	We	Th	Fr	Sa
14				1	2	3	4
15	5	6	7	8	9	10	11
16	12	13	14	15	16	17	18
17	19	20	21	22	23	24	25
18	26	27	28	29	30		

May 2020

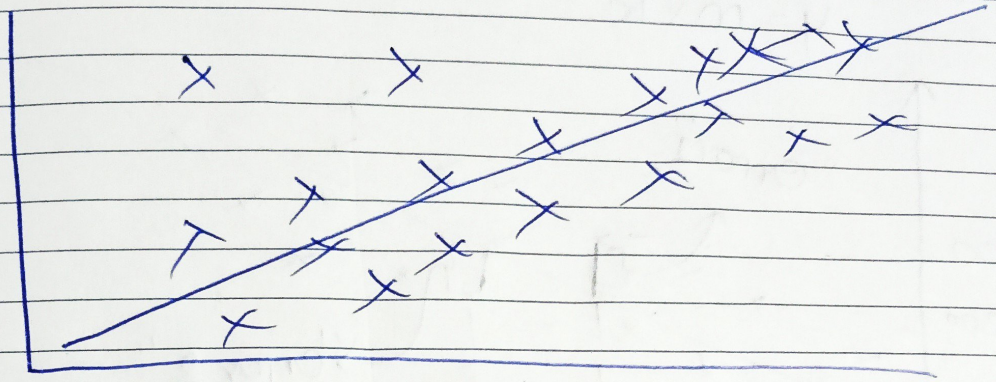
Wk	Su	Mo	Tu	We	Th	Fr	Sa
18	31					1	2
19	3	4	5	6	7	8	9
20	10	11	12	13	14	15	16
21	17	18	19	20	21	22	23
22	24	25	26	27	28	29	30

March
Monday
Wk-10•Day (062-304)
02
2020

MAR

7am Now in real world dataset are noisey
ie. not exactly linear

8



9

10

11

12 pm

But still plotting a ~~lint~~ line through that will give us a rough estimate.

1

Goal: Goal is to find best possible value for "m" & "c"

2

3

This can be achieved only by failing again and again.-
ie. by looking through large datasets.

4

5

6

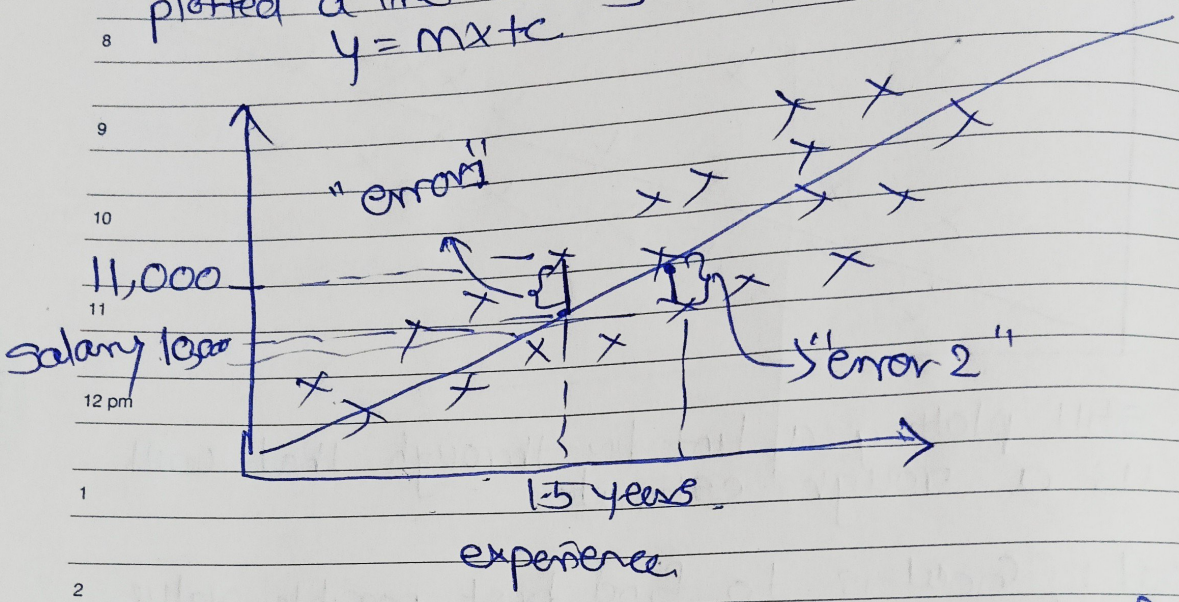
7

Birthday / Anniversary

Today is your day! Your mountain is waiting. So... get on your way.

Now, lets say for given data you have plotted a line having equation

$y = mx + c$



You our line is predicting a salary of ₹10,000 for 1.5 yrs of experience. but actual value is ₹11,000

We can say that our line is off by 1000.

But So our formula for error can be

$\sum$ [Predicted − Actual]

This is for only a single value. so for total dataset formula will be

$$\sum_{j=1}^{n} [\text{Predicted} - \text{Actual}]$$

$\frac{1}{n}$ ← (mean error)

Birthday / Anniversary

7am We can seen in graph for error 2 our [Predicted - Actual] will be -ve

8

So, in general for given graph our error can

9

come out to be zero due to set of +ve & -ve

10 values.

To overcome this problem i.e we take

11 square of [Predicted - Actual]

12 pm $$error = \frac{\sum_{i=1}^{n} [Predicted - Actual]^2}{N}$$

1 due to squaring, even for smallest values error become excessively large.

2

To avoid this we take the square root of

3 the whole term.

4 Thus

$$error = \sqrt{\frac{\sum_{i=1}^{n} [Predicted - Actual]^2}{N}}$$

5

6

This is a typical loss/cost function

7

called RMSE.

Birthday / Anniversary

However long the moon disappears, somebody it must shine again.

March
Thursday
20 | Wk-10 • Day (065-301)

7am Linear regression →