

## Introduction to Machine Learning

- Bookish Definition: Machine learning (ML) is the study of computer algorithms that improve automatically through experience.
- Pattern Recognition, Prediction
- Data
- Example: Salary, Age, Experience, Years of Education
- What pattern/prediction possible?
- Correlation? Can we go further?
- The magical formula for output and input
- Function - Math starts
- Slope Equation - High School Math
- Two points vs More points for Slope Equation
- Slope Equation becomes Approximation Equation
- Error / Loss
- Everything boils down to finding that magical  $m$  in the equation that reduces the loss
- And that magic is of learning is .....

Data:

Salary, Age, Experience, Years of Education Post High School

10k	20	1	4
20K	22	2	4
50K	25	4	4
100K	22	0	8

Salary, Age, Experience and Years of Education Post High School -> Variables / Features

1st row is 1 observation, similarly all the rows are observations

Prediction: I want to predict Salary using other Variables.

Output: Salary

Input: Age, Experience, Years of Education

Input and I want to predict the output.

All we need now is a function that MAPS Input to Output.

$f(\text{age, experience, years of education}) = \text{Salary}$

This is the example of typical Mathematical Function that we all know.

Correlation?

Can we predict Salary using Correlation only?

If A and B have a correlation of 0.9, can I now how A changes with changes in B? (task for you.)

$f(\text{age, exp, years of education}) = \text{Salary}$

Salary = 2000 \* Years of education -> Very bad equation

Years of education 1 -> 2000, 4 -> 8000

Salary = 5000 \* Experience

Exp 1 -> 5000, Exp 3 -> 15000

Exp 0 -> Salary -> 0

Salary = 5000 \* Experience + 2000

$Y = 2X + 10$  -> Function -> Magical Equation for our case

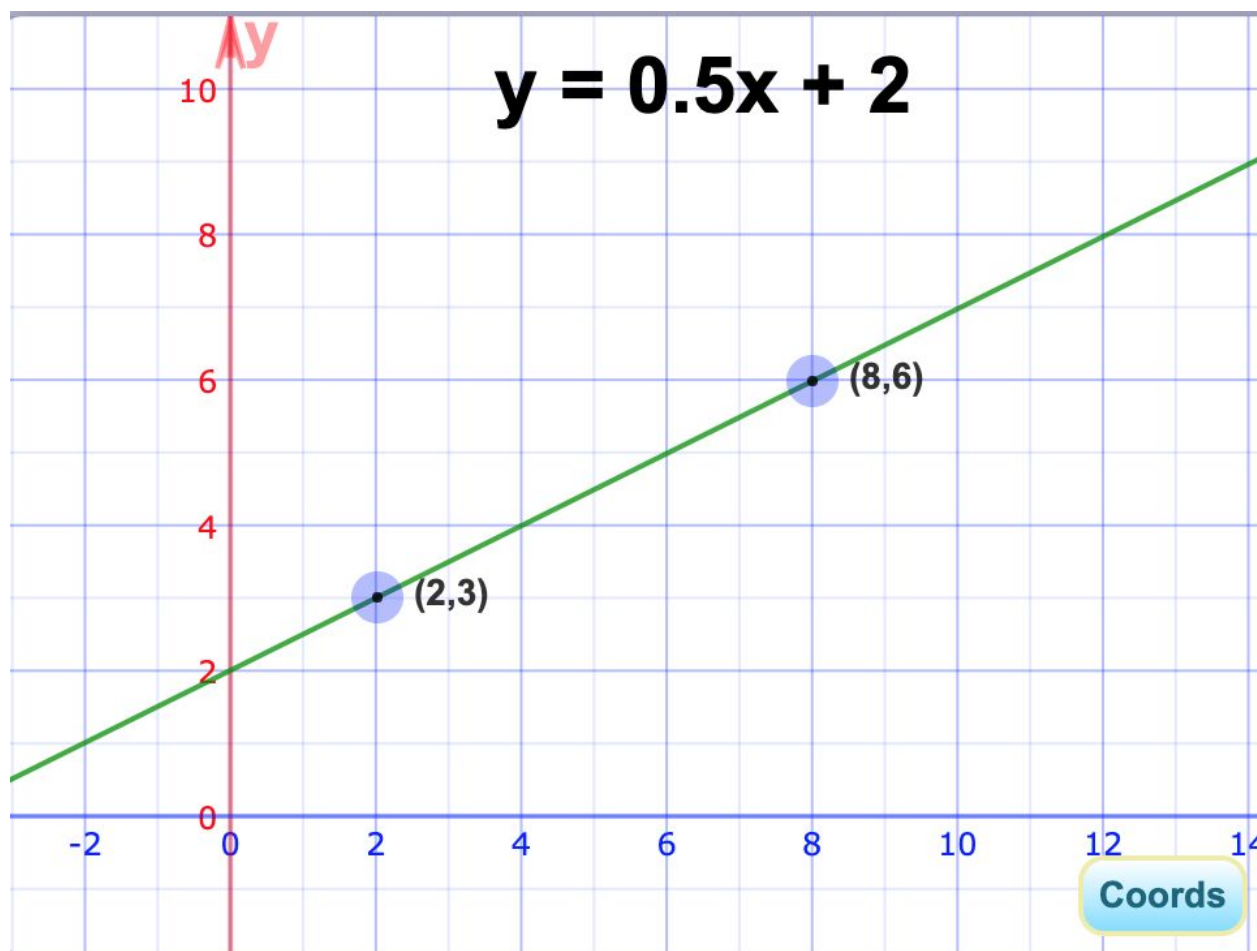
If I increase X by 1 unit, by what unit does Y increase/decrease?

If there a way we can know this directly without plugging any values?

X = 1, Y -> 12, X = 2, Y -> 14, for every unit change in X, you get 2 unit changes in Y.

$Y = 5X_1 + 10X_2 + 43X_3 + 500$

If I change X<sub>2</sub> by one unit, by what unit does Y change?



This concept of Unit Change in Output for Unit Change in Input comes with the concept of Slope in Mathematics.

$$Y = 2X + 10$$

$Y = mx + C$  (This comes directly 8th grade high school.)

$M$  = Slope

$C$  = intercept = value of  $Y$  when  $x = 0$

Slope = Change in  $Y$  Unit / Change in  $X$  Unit

What is one term in Mathematics that is same as Slope?

For small change in  $x$ , that is  $dx$ , what is the change in  $y$ , that is  $dy$ ?

$dy/dx$  = Derivatives in calculus = Slope

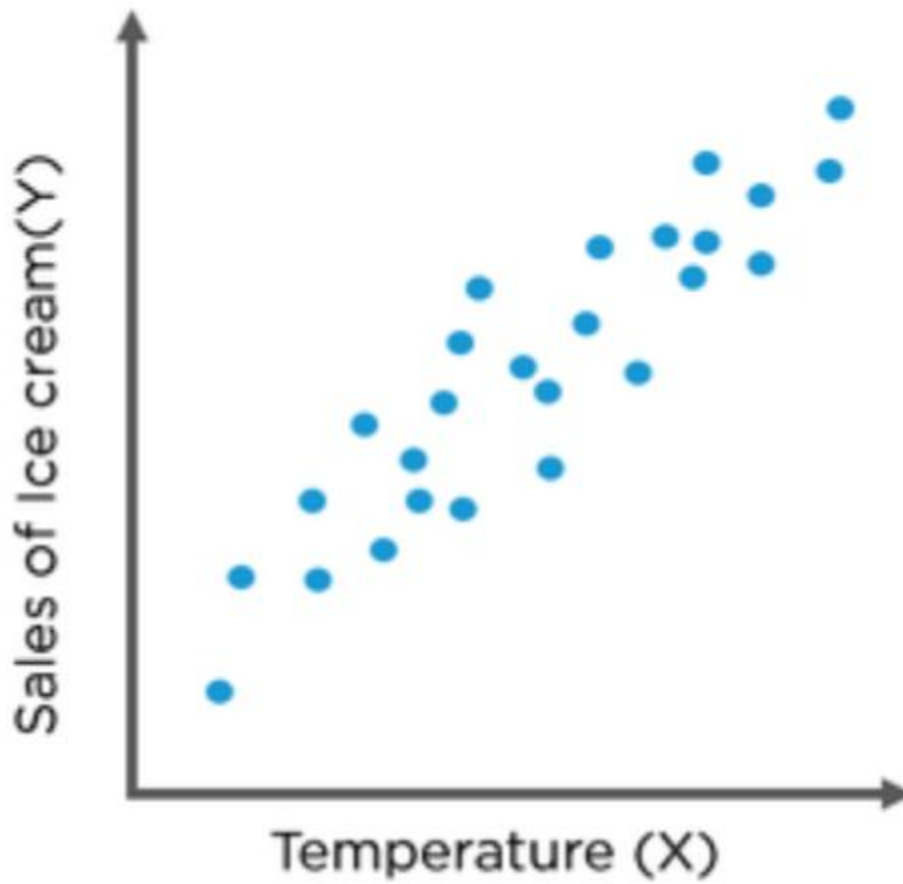
Infinitesimal  $dx \rightarrow$  Derivatives

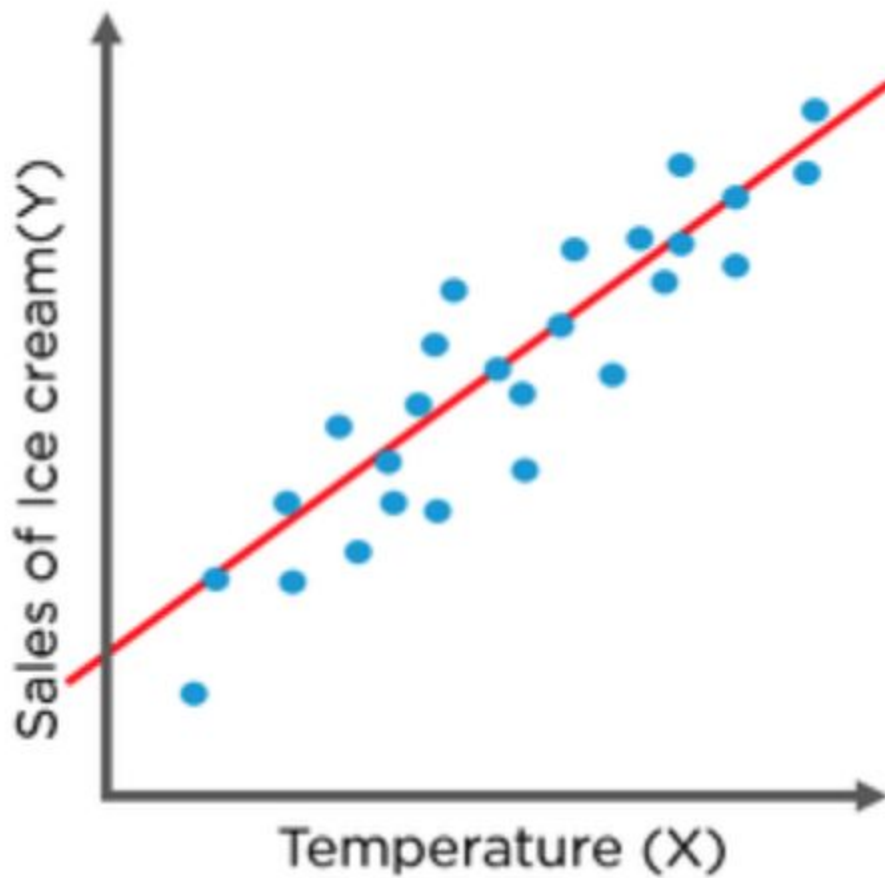
$$Y = 2X^2 + 1$$

For unit change in X, what is the change in Y? (This is for you to think.)

Salary, Exp = (10k, 4), and (50k, 4)

What if we have more points that 2??





Output =  $m \cdot \text{Input} + \text{Intercept} \Rightarrow (Y = mx + c)$

We can predict Sales of Icecream using Temperature input and our equation above.

Output  $\sim m \cdot \text{Input} + \text{Intercept}$  (for most of the points above)  $\Rightarrow$  Approximation

Output =  $m \cdot \text{Input} + \text{Intercept} + \text{Some Error Term}$ , can I write this???

One Output, Multiple Inputs  $\rightarrow$  You can not draw a 2-D image always

One Output, 10 Inputs  $\rightarrow$  You can not draw 10 D input digram

For example, in  $Y = 5X_1 + 10X_2 + 43X_3 + 500$ , how to find 5, 10, 43 and 500

All you will have at hand is set of data points like:

Y X1 X2 X3

20 3 5 8 = (20, 3, 5, 8) => Y = 20, X = (X1, X2, X3) = (3,5,8)

18 2 3 9

19 4 5 12

20 3 4 10

16 2 5 12

Function(X1, X2, X3) = Y

Some sort of function like:  $Y = 5X_1 + 10X_2 + 43X_3 + 500$

So again, how to find 5, 10, 43 and 500?

For more general case:

$Y = m_1X_1 + m_2X_2 + m_3X_3 + \text{Intercept}$

How to find  $m_1$ ,  $m_2$ ,  $m_3$  and Intercept from our data points? (Which are also known as Coefficients)

Let's say using some magic and the data given, you do manage to calculate  $m_1$ ,  $m_2$ ,  $m_3$  and Intercept.

After that, can you at least try to predict Y for any input set X1, X2 and X3?

This magic of calculating  $m_1$ ,  $m_2$ ,  $m_3$  and Intercept using Data is known as Machine Learning.

So, ML ultimately boils down to learning  $m_1$ ,  $m_2$ ,  $m_3$  and intercept from Data.

How to perform this magic / how to do Machine Learning, we will come to learn in our coming lectures.

Next up, we will do a session of Mathematical Brush Up!

First query: Machine Learning vs Modeling? We perform ML to get a final magical equation which is a Model of Input and Output.

Second query: Is Error term covered/compensated by the Intercept? No. Since whatever best fit line you come up with, it won't contain all the points despite having intercepts (which also changes depending on the best fit line calculated)