

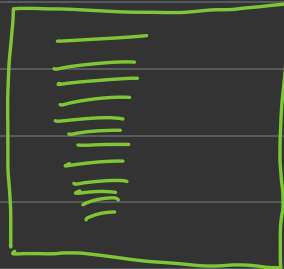
03 Sept 2022

Term 4 - ML

April 2022

- Data is the new oil.
  - ML is a powerful concept & framework for making the best out of it.
  - Age of Information or Age of data.
- eg: Google knows everything about my location.

MPT  
(Mumbai & N. Mumbai)



Data

Information

[Darshan + 49 more ppl]



Bus Service	Car
X	✓ convenient & easy
AC	AC
✓	Guaranteed Seat
P2P	Home - Destn
15min	Time conv.
₹ 2950/- pm	10 000 - 120000 rs / month
✗ No Driving	Driving - Tired & fatigue
Sleep, Meetings, Netflix	Home - Fresh

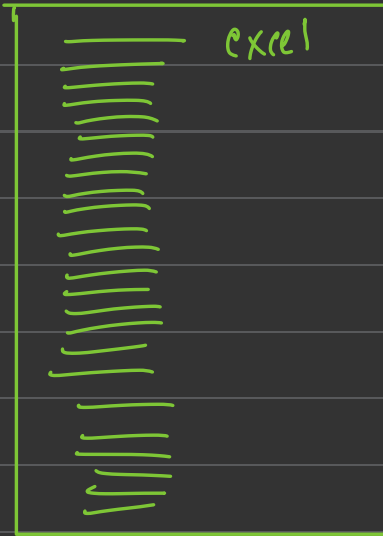
Who is most intelligent lifeform on the planet at the moment?  
Human.

- m/c: To mk data driven decisions at a scale.

your friend (10yrs)

Will you lend a 10000 vs?  
10yrs (100 interaction)

past  
experience



M/C

80 yes ✓ ∴ Money Lended

20 No

10 interaction Human

7  
yes  
✓

3  
No

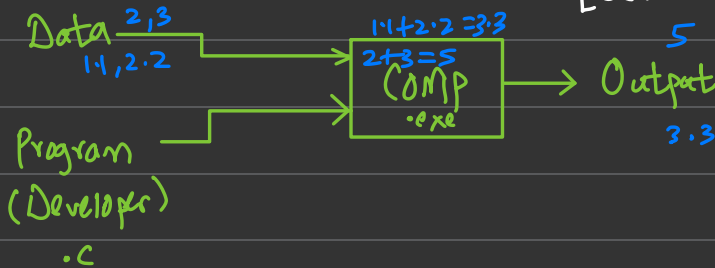
∴ money lended.

# Traditional Programming Paradigm

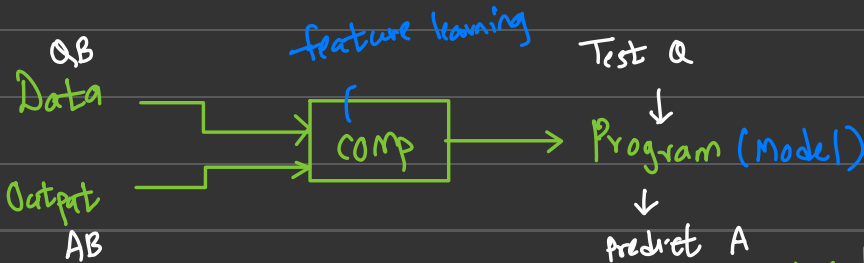
HDFC  $\xrightarrow{\text{website}}$  Infosys

Task: Add two nos

[Deterministic]



# ML Programming Paradigm



QB	AB
1	1
2	2
3	3
...	...
...	...
100	100

Study  
80

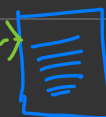
80 QB  
80 AB



QB & AB

20 Q

20 A



Predicted

Actual (True)

1  
Compare Actual & Predicted  
answers & rate the performance  
of a student.

———— X ————

04 Sept 2022

<sup>y</sup>  
Dependent  
Variable (1)

<sup>X</sup>  
Independent Variables (1 or more)

① Price of car ~ Brand + Engine Power + Mileage + Transmission Type + Color + Variant + Fuel Type + Body Type + ABS + EBD + ADAS + HP + Torque

Regression

② Price of Mobile ~ Brand + RAM + Storage + Battery + Screen Size + Camera + OS + 5G/4G + OLED/AMOLED + Processor

③ Apple vs Orange ~ Color + Taste + Shape + Texture + Nutrition Value + Calorie Value + Price

Classification

④ Iris (Setosa vs Versicolour vs Virginica) ~ Sepal Length + Sepal Width + Petal Length + Petal Width

[https://miro.medium.com/max/1000/1\\*IFC\\_U5j\\_Y8IXF4Ga87KNVg.png](https://miro.medium.com/max/1000/1*IFC_U5j_Y8IXF4Ga87KNVg.png)

[https://miro.medium.com/max/638/0\\*2c7voFri9clXGrc4](https://miro.medium.com/max/638/0*2c7voFri9clXGrc4)

# ML Problems

Regression

y is Continuous

Classification

y is Categorical

Binary (2)

eg: Apple vs  
Orange

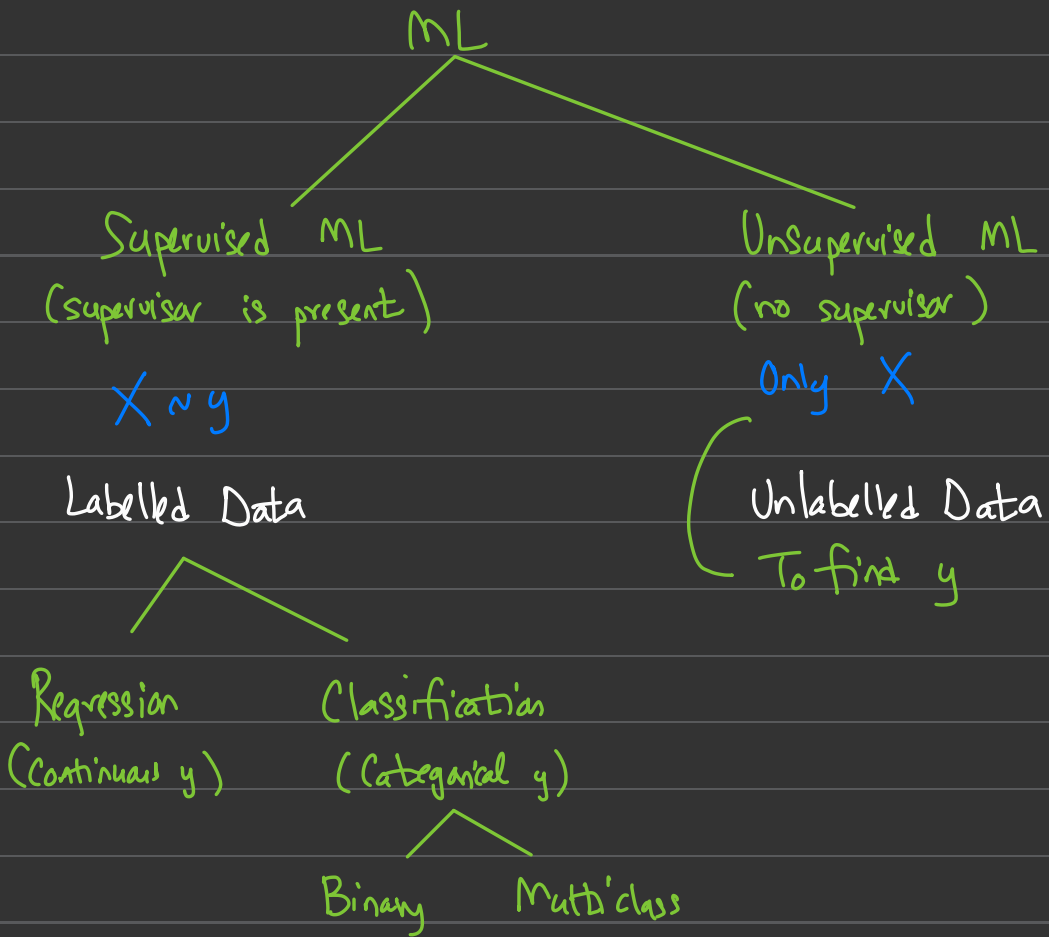
eg: Male vs  
Female

eg: Spam vs  
Ham

Multi-class (more <sup>than</sup> 2)

eg: Iris Dataset

eg: Apple vs Orange  
vs Mango vs Grapes



<https://www.aplustopper.com/wp-content/uploads/2017/05/Correlation-1.jpg>

Take Home Task: Difference b/w Pearson's,  
Spearman Rank & Kendall's  
Correlation Coefficient + Applications

<https://www.mathsisfun.com/data/correlation.html>



Graph Index

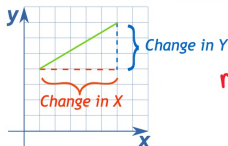
To calculate the Slope:

Equation of a  
Straight LineDivide the **change in height** by the **change in horizontal distance**

Slope  
 $m = 0.5$

If we move 1 unit on X-axis,  
we climb 0.5 units on Y-axis

Slope =  $\frac{\text{Change in Y}}{\text{Change in X}}$   
 $= \frac{y_2 - y_1}{x_2 - x_1}$



$$m = \frac{6-3}{8-2} = \frac{3}{6} = 0.5$$

$$m = \frac{3-6}{2-8} = \frac{-3}{-6} = 0.5$$

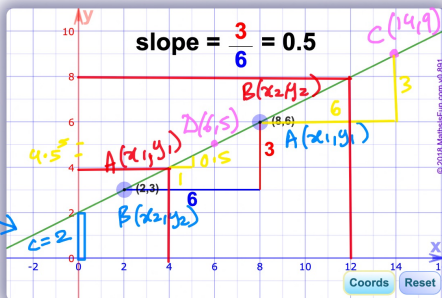
Have a play (drag the points):

y-intercept  
 $c = 2$

Eqn. of Line

$$y = mx + c$$

$$y = 0.5x + 2$$



$$m = \frac{9-5}{14-6} = \frac{4}{8} = 0.5$$

$$m = \frac{5-9}{6-14} = \frac{-4}{-8} = 0.5$$

Given  $x$ , we can find  $y$ .

If  $x = 4$ ,  $y = 0.5(4) + 2 = 4$ , If  $x = 12$ ,  $y = 0.5(12) + 2 = 8$

Examples:

