

## ★ Flow -

## ★ Content -

- ① Linear Algebra
- ② Calculus
- ③ Co-ordinate Geometry
- ④ Optimization

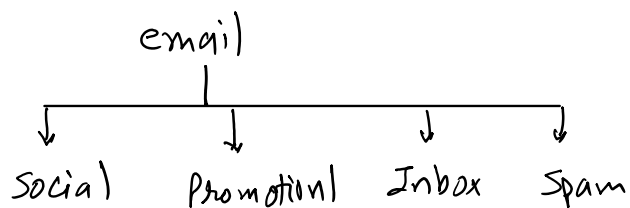
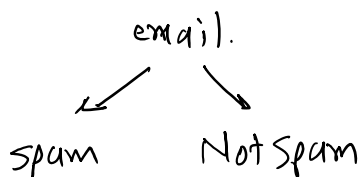
① Why are we studying this topic? How is this related to ML?

- ② Understand the concept
- ③ Visualization of the problem/sol<sup>n</sup>
- ④ Maths behind the concept
- ⑤ Code (wherever needed/possible)

## ★ Introduction

→ Two types of problems in ML

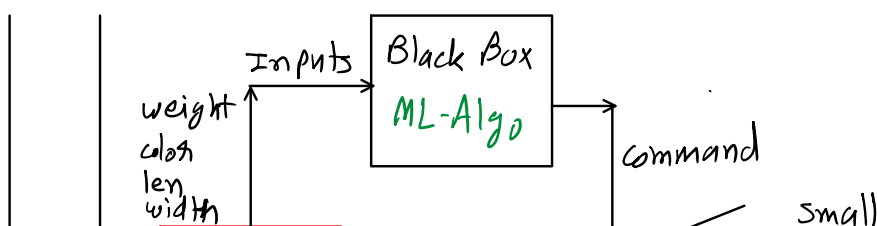
(i) Classification - categorizing an incoming mail

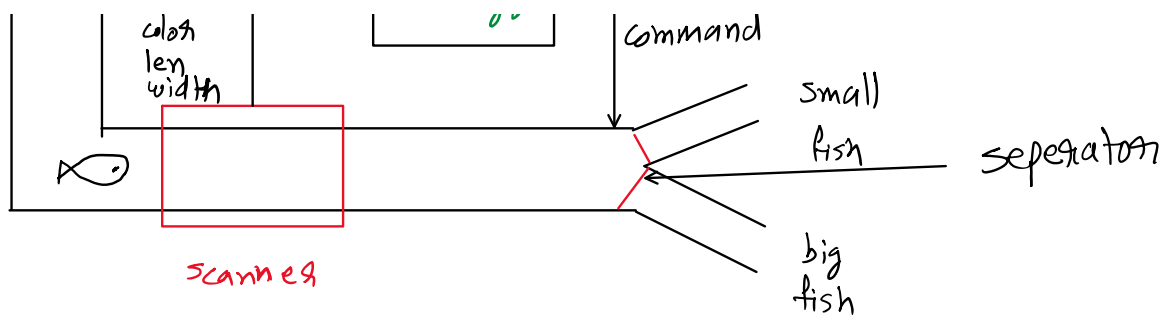


(ii) Regression - Predicting price of a house

BHK	Area	Locality	Parking	...	price
4	100	posh	2		3.6
<div></div>					<div></div>

## ★ Our Core Example - Fish Sorting Mechanism





★ What if a human (newly employed) is going to do this job? How will he/she learn which one is a big fish / small fish?

★ Process:

① Training - A supervisor will show him/her which are big vs. small fishes for a large no. of fishes.

Training Data

② Testing - The new employee will identify / label the big & small fishes while the supervisor watches.

Test Data

③ Deploy -

If they perform well then he/she will be deployed else he/she will be put back into the training.

→ We do exactly the same process to deploy an ML Program (ML Model)

★ Let's have a look at a sample data:

features / Independent Variables					label / Dependent Variable
length	width	color	weight	.....	Type
25	10	Red	150		5

OR target

→ Data point /

25	10	Red	150	5
300	100	Black	2000	3

— target

→ Data point /  
Record

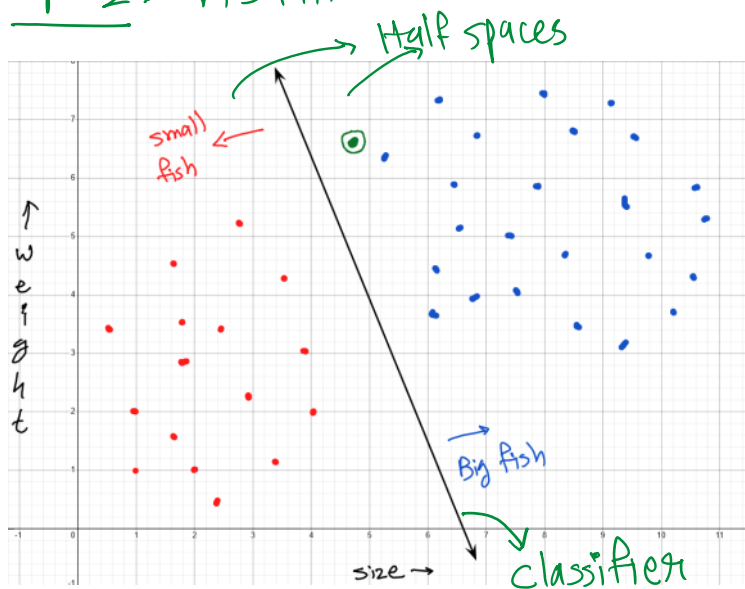
## Step-1: Splitting Train-Test Data

→ We will first divide our data into two parts - Training Data & Testing Data.

This split may be 60:40 | 70:30 | 80:20 or any other combination

## Step-2: Visualize the data

— For simplicity, we will take just two features - size & weight



→ After visualizing the data points in form of a scatter plot, we can say that if our ML program can draw / find a line that separates these two types of dots

then its job will be done. Therefore 'Learning' of machine is nothing but finding this line / boundary.

→ This line is also known as 'separator' / 'boundary' / 'classifier'

∴ We need to study the maths of a straight line.

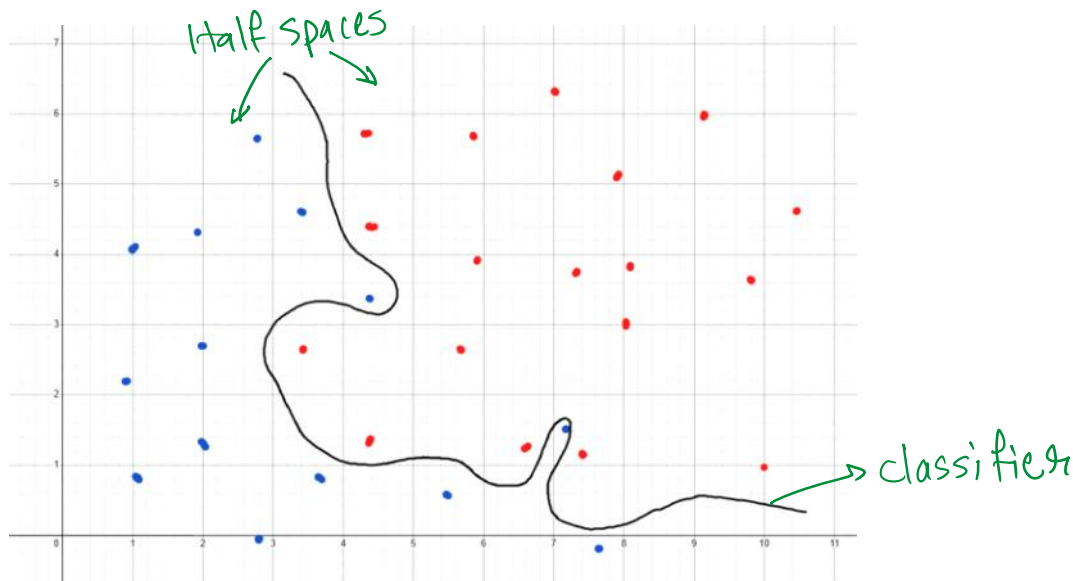
→ Some questions:

→ Some questions:

① What if it is not possible to draw a straight-line that perfectly separate the dots?

→ We might need to draw a curve. OK

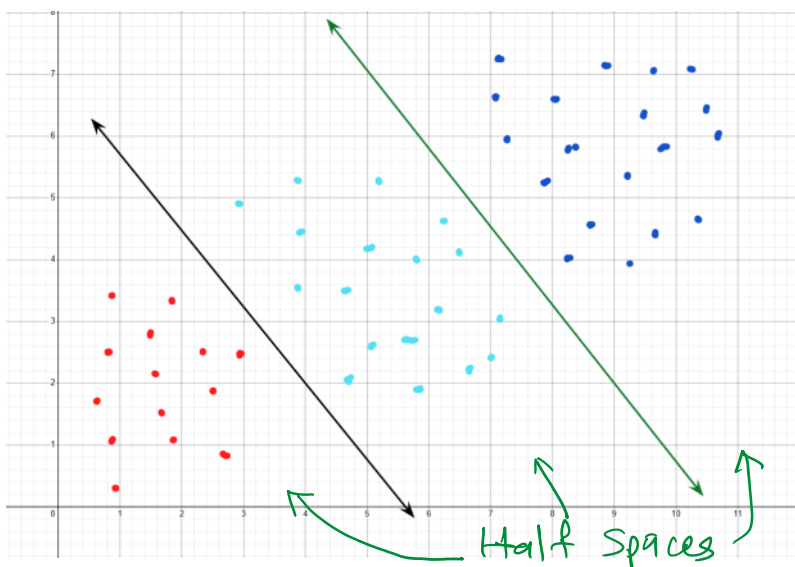
→ We need to compromise with the accuracy.



② What if there are more than two types?

eg. small / medium / large fishes

→ We will have two classifiers



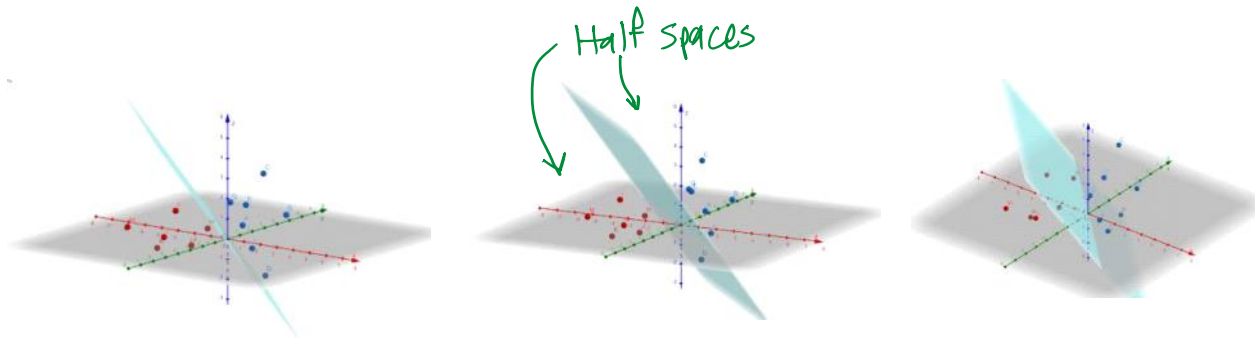
✳ What if an extra feature is added?

★ What if an extra feature is added?

eg, if we took length, width & weight, how would the graph look?  $\rightarrow$  3D

Why 3D? Because there are 3 features.

$\therefore$  Features are also called 'Dimensions'



★ What if we add one more feature? Now the space becomes 4D which is beyond visualization, but what will happen with our separator (boundary)?

space/no. of features	classifier
2D	$\rightarrow$ 1D $\rightarrow$ straight line
3D	$\rightarrow$ 2D $\rightarrow$ plane (= straight line in 2D)
4D	$\rightarrow$ 3D
$\vdots$	$\vdots$
$n$ D	$\rightarrow$ $(n-1)$ D

} straight lines in higher dimensions. } Hyper planes

At the end, we must learn maths behind the straight line because all the hyperplanes are nothing but straight lines in higher dimensions.