Week 01

1] What is Machine learning ?

" Field of study that gives computers the ability to learn without being explicitly programmed"
- Arthur Samuel

Types of machine learning

1] Supervised learning (Mostly used)
2] Unsupervised learning
3] Reinforcement learning
4] Semi supervised learning

2] Supervised learning

*] Supervised learning is a type of machine learning where the model is trained on a dataset that has inputs (features) and known outputs (labels)

*] The algorithm learns a mapping function from inputs → outputs, so that when it sees new data, it can predict the output

Main 2 types of supervised learning:

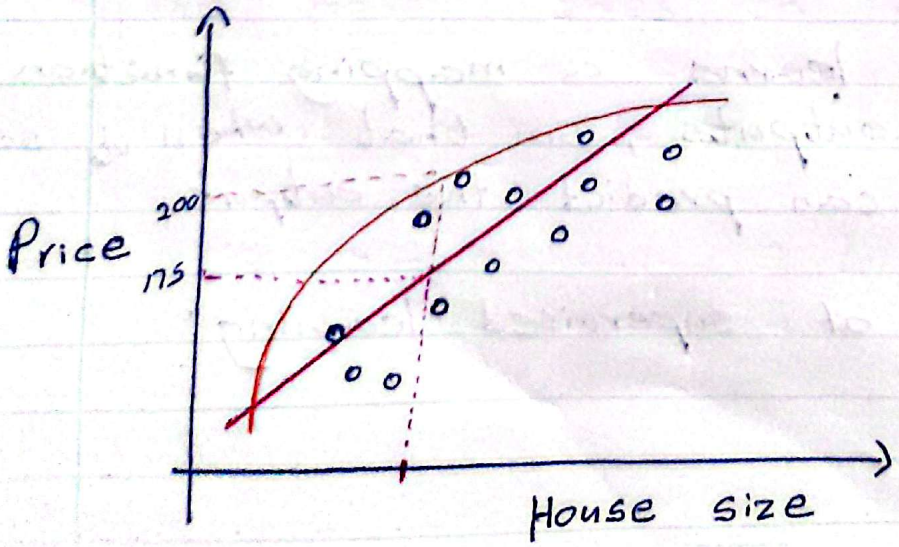
1] Regression
2] Classification

Applications of supervised learning

Input (X)	Output (Y)	Application
email	spam (0/1)	Spam filtering
audio	text transcript	Speech Recognition
English	Spanish	machine translation
ad, user info	click (0/1)	Online advertising
image, radar in.	position of cars	Self-driving car
image of phone	defect (0/1)	Visual inspection

1) Regression

Predict a number (continuos values) from infinetly many possible outputs

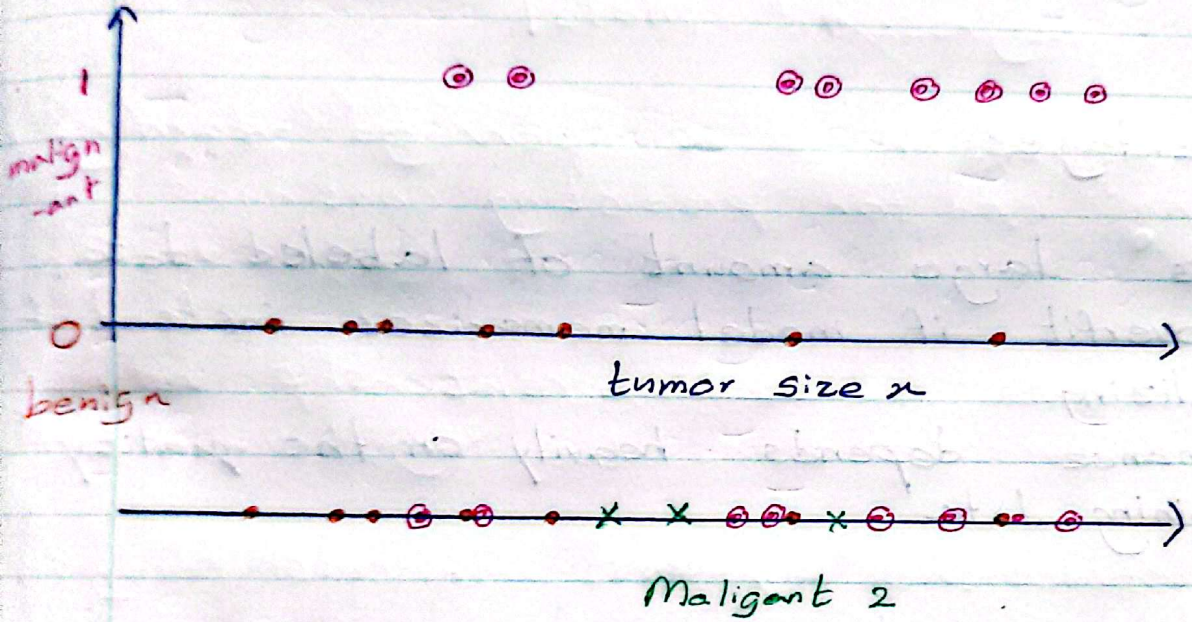


Output prediction Can be a straight line, curve or even more complex function

2] Classification

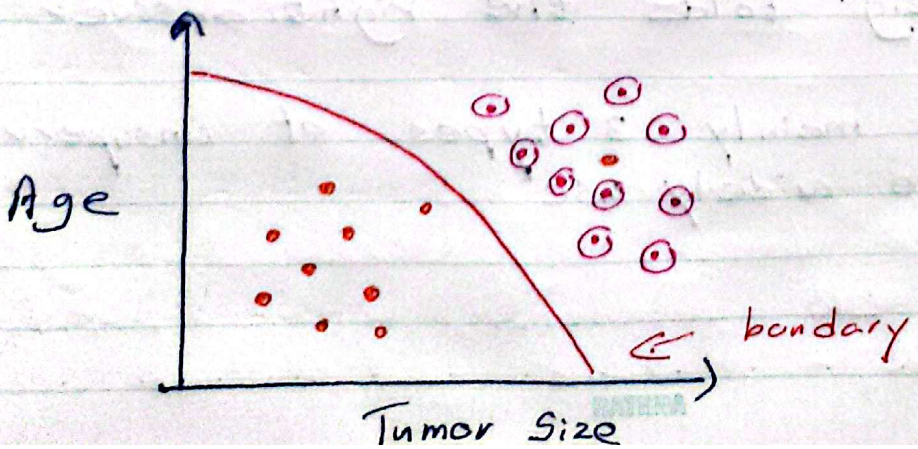
Predict categories / classes (Non-numerical or numerical) from small number of possible outputs

Breast cancer prediction :



*) Can be more than 2 outputs but ^ has only finite no. of outputs

*) There also can be more than two inputs (most cases). In that case algorithms job is to predict a perfect boundary to seperate outputs (categories)



Advantages

*) Easy to understand
*) Widely used.
*) produce high accuracy if we have enough quality labled data
*) Works for both simple and complex tasks

Disadvantages

*) Requires large amount of labeled data.
*) Can overfit if model memorizes instead of generalizing
*) Performance depends heavily on the quality of training data

3) Unsupervised learning

*) Unsupervised learning is a type of machine learning where the dataset has only inputs (features) but no labels (output)

*) The model tries to find hidden patterns, structures, or relationships in the data without beig told the right answer

There are mainly 3 types of unsupervised learning we aidely use :

1] Clustering:

Group the similar data points together

2] Anomaly detection:

Find unusual data points

3] Dimensionality Reduction:

Reduces features while keeping important information

Applications

*) Marketing: Grouping customers into segments for targeted ads (c)
*) Healthcare: Find disease subtypes from patient data (c)
*) Finance: Detecting unusual spending patterns (a)
*) Image processing: Compressing images with fewer features

Advantages

*) Unlabled data are cheap and widely available
*) Helps discover hidden patterns humans may miss.
*) Useful for exploratory data analysis (EDA)

Disadvantages

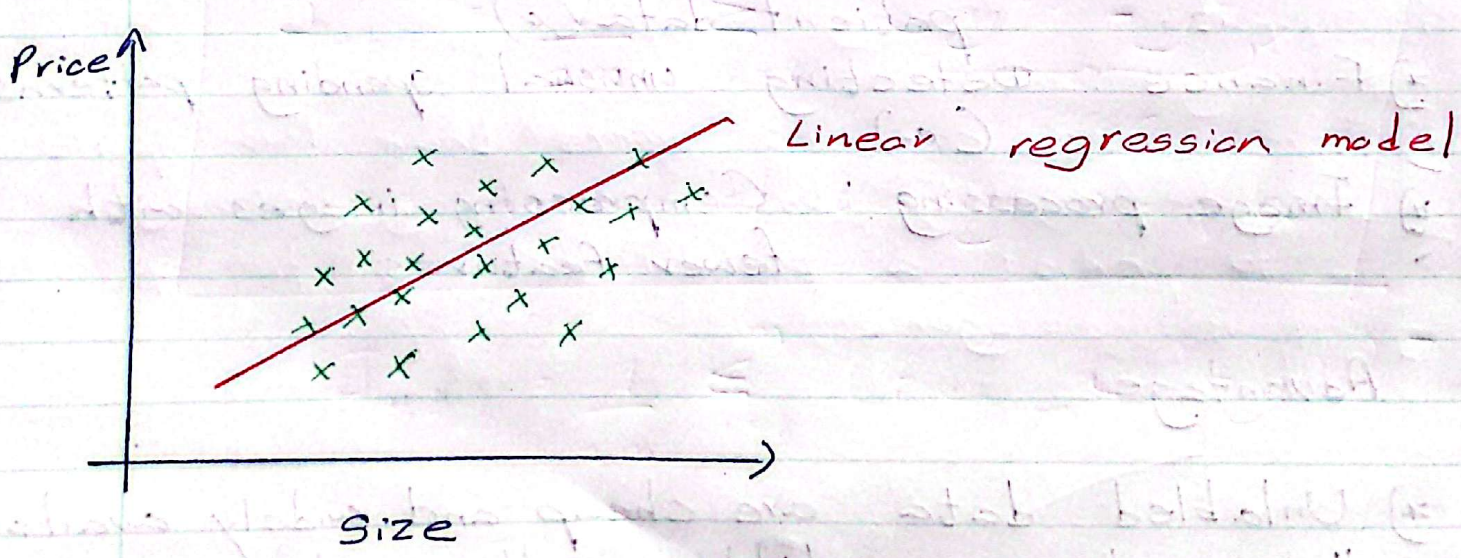
*) No clear "Correct answer"
*) Results may be hard to interpret
*) May require tuning

4] Linear Regression Model

*] Linear regression is a supervised learning algorithm used for predicting continuos values

*) It assumes a linear relationship between input features (x) and the target variable(Y)

*] The model tries to fit the best straight line through the data points, so we can predict outcomes



Two types of linear regression

1] Univariate LR (simple LR) - one input feature
2) Mutliple LR - multiple input features

Terminology

	Size (ft^2)	price ($)
(1)	2104	400
(2)	1416	232
(3)	1534	315
⋮	⋮	⋮
(47)	3210	870

*) Training set = Data used to train the model

*) x = input variable feature

*) y = output variable / target variable

*) m = number of training examples

*) (x, y) = Single training example

*) $(x^{(i)}, y^{(i)})$ = i^{th} training example

Process of Supervised learning

training set (features, targets)

↓

learning algorithm

↓

$x \longrightarrow \boxed{f} \longrightarrow \hat{y}$ (prediction)

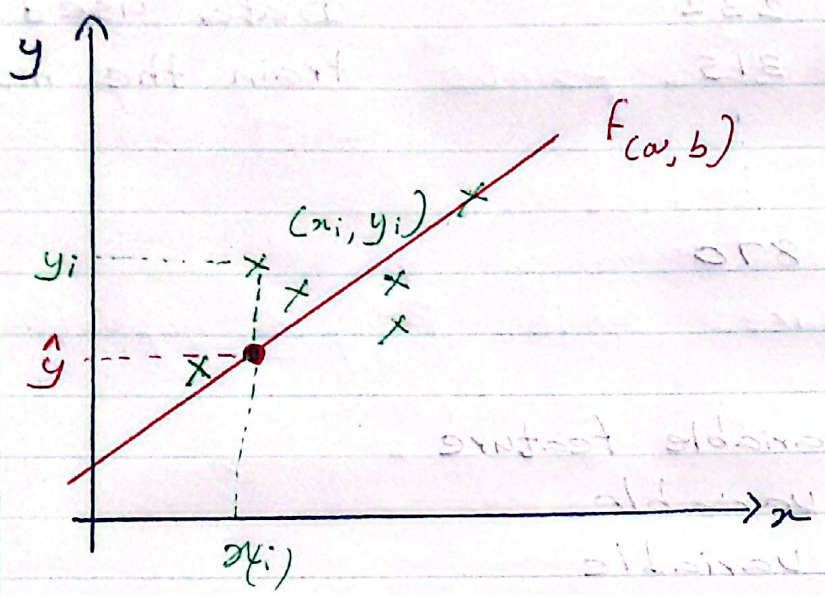
($\hat{y}$ = Estimated y)

feature → model (function / hypothesis) → target

Represent (f) => $f_{w,b}(x) = wx + b$

$f(x) = wx + b$

5] Cost function formula



$\hat{y}^{(i)} = f_{w,b}(x^{(i)})$

$f_{w,b}(x^{(i)}) = wx^{(i)} + b$

*] Cost function : Squared error cost function

$$J_{(w,b)} = \frac{1}{2m} \sum_{i=1}^{m} \underbrace{(\hat{y}^{(i)} - y^{(i)})}_{\text{error}}^2$$

m = number of training examples

$$J_{(w,b)} = \frac{1}{2m} \sum_{i=1}^{m} (f_{w,b}(x^{(i)}) - y^{(i)})^2$$

Goal : Find w, b
where $\hat{y}^{(i)}$ is closed to $y^{(i)}$ for all $(x^{(i)}, y^{(i)})$

model → $f_{w,b}(x) = wx + b$

parameters → w, b

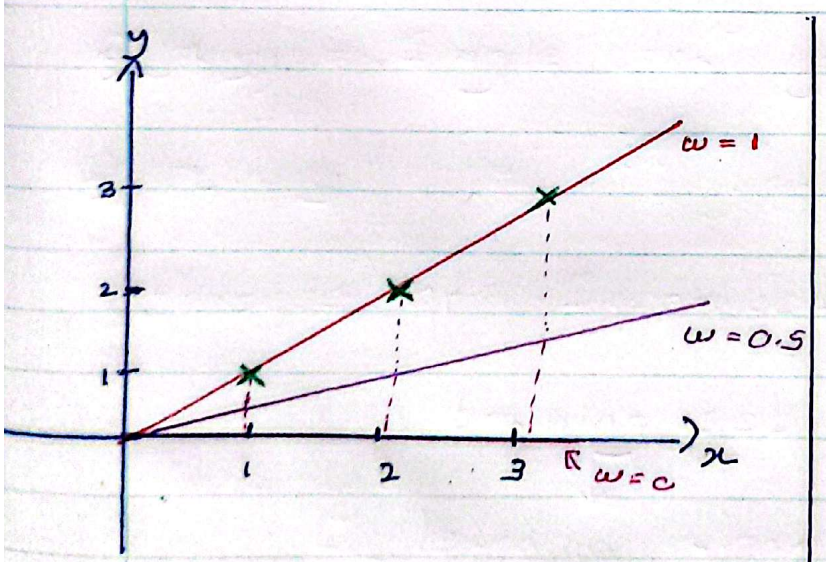
Cost function → $J(w,b) = \frac{1}{2m} \sum_{i=1}^{m} (f_{w,b}(x^{(i)}) - y^{(i)})^2$

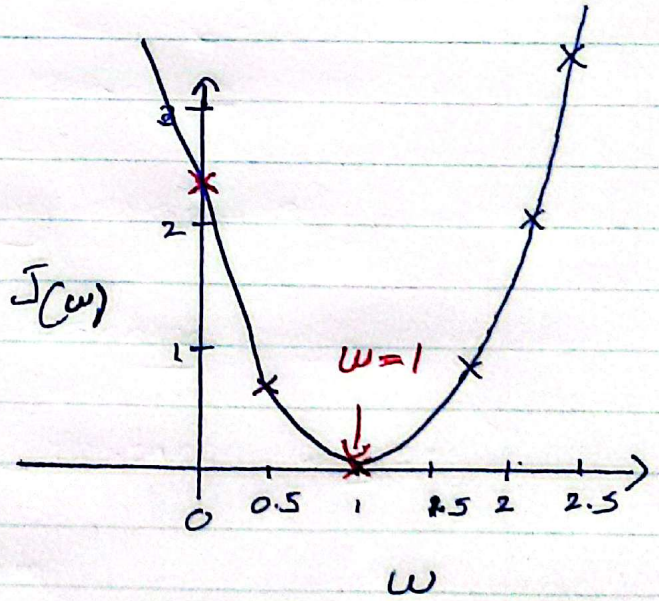
goal → minimize $J(w,b)$
w, b

Simplified → $f_w(x) = wx$

$$J(w) = \frac{1}{2m} \sum_{i=1}^{m} (f_w(x^{(i)}) - y^{(i)})^2$$

Goal → minimize $J(w)$
w





$J(w^1) = \frac{1}{2m}(0^2 + 0^2 + 0^2) = 0$

$J(w)^{0.5} = \frac{1}{2m}((0.5-1)^2 + (1-2)^2 + (1.5-3)^2) = 0.58$

$J(w)^0 = \frac{1}{2m}(1^2 + 2^2 + 3^2) = 2.3$

Chose w to minimize $J(w)$