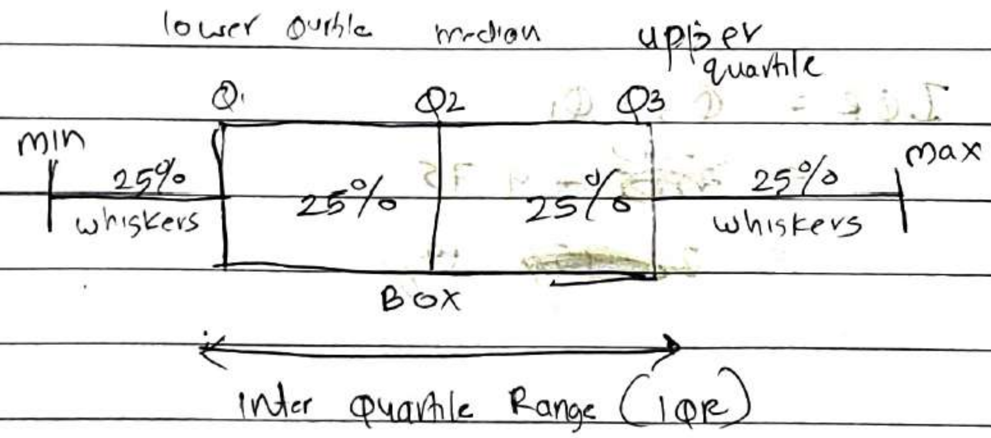
- minimum :- It is the minimum value in the dataset.
- First Quartile (Q_1) - 25 % of the data lies below the first (lower) quartile.
- median (Q_2) - It is the mid - point of data set.
- Third quartile (Q_3) - 75% of the data lies below the third (upper) quartile.
- maximum :- It is the maximum value in the dataset excluding the outliers.



Q heights of 40 students in a class having 5 number sumary as follows. construct a box plot for the dataset.

min = 59

max = 77

Q_1 = 64.5

Q_2 = 66

Q_3 = 70

3) Minimize $f(x) = x^2 + 2x$ interval $(-3, 4)$, $n = 5$

$a = -3,\ b = 4.$

~~L = (-3) 4.~~

$L = b - a.$

$4 - (-3)$

$L = 7.$

set $k = 2$

$n = 5.$

$$L_k^* = \left[\frac{F_{n-k+1}}{F_{n+1}}\right] L$$

7.

$$L_2^* = \left[\frac{F_{5-2+1}}{F_{5+1}}\right] \times 7 \Rightarrow \frac{F_4}{F_6} \Rightarrow \frac{5}{13} \times 7.$$

$L_k^* = 2.69.$

$x_1 = a + L_k^*$

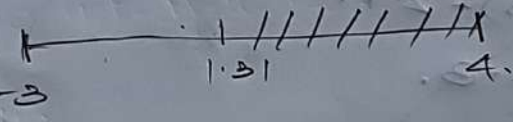
$= (-3) + 2.69.$

$= -0.31$ //

$x_2 = b - L_k^*$

$= 4 - 2.69.$

$= 1.31$ //

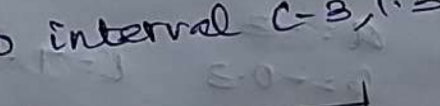


$f(x_1) = f(-0.31) = (-0.31)^2 + 2\times(-0.31) = -0.52.$

$f(x_2) = f(1.31) = (1.31)^2 + 2(1.31) = 4.33$

$f(x_2) > f(x_1)$ discard; New interval $(-3, 1.31)$.

$(1.31, 4)$.



$k = n$

$2 \neq 5$ Proceed further

2nd Iteration.

set $k = k+1.$

$k = 3.$

$L = 7.$

$a = -3$

$b = 1.31.$

$$L_k^* = \left[\frac{F_{n-k+1}}{F_{n+1}}\right] L \Rightarrow L_3^* = \left[\frac{F_{5-3+1}}{F_{5+1}}\right] \Rightarrow \cancel{\frac{F_3}{F_4}} L$$

$\frac{F_3}{F_6} \times L \Rightarrow \frac{3}{13} \times 7.$ $\cancel{\frac{F_3}{F_6}} \times L \Rightarrow$ ~~$\frac{3}{5} \times 7 \Rightarrow L_k^* = 4.2$ //~~

$= 1.61$ //

$x_1 = a + L_k^*$
$= -3 + 4.2$
$= 1.2$

$x_2 = b - L_k^*$
$= 1.31 - 4.2$
$= -2.89$

$x_1 = a + L_k^*$
$= -3 + 1.61$
$= -1.39$

$x_2 = b - L_k^*$
$= 1.31 - 1.61$
$= -0.3$

$f(x_1) = (-1.39)^2 + 2(-1.39) = -0.84$
$f(x_2) = (-0.3)^2 + 2(-0.3) = -0.51$
$f(x_2) > f(x_1)$.

$-3 \quad -1.39 \quad -0.3 \quad 1.31$

New interval $(-3, -0.3)$

$a = -3$
$b = -0.3$

$L = b - a$
$= (-0.3) - (-3) = 2.7$

Check $k = n$:
$3 \neq 5$.
Proceed Further
So,

$x_1 = a + L_k^*$
$= -3 + 2.7$
$= -0.3$

$x_2 = b - L_k^*$
$= (-0.3) - 2.7$
$x_2 = -3$

3rd Iteration:-

So, $k = k+1$
$k = 3+1$
$= 4$

$a = -3$
$b = -0.3$
$n = 5$
$L = 7$

$$L_k^* = \left[\frac{F_{n-k+1}}{F_{n+1}}\right] \times L \Rightarrow \left[\frac{F_{5-4+1}}{F_{5+1}}\right] \times L$$

$$\Rightarrow \frac{F_2}{F_6} \times L \Rightarrow \frac{2}{13} \times 7 \Rightarrow L_k^* = 1.07$$

$x_1 = a + L_k^*$
$= -3 + 1.07$
$= -1.93$

$x_2 = b - L_k^*$
$= -0.3 - (1.07)$
$= -1.37$

$f(x_1) = (-1.93)^2 + 2(-1.93) = -0.13$

$f(x_2) = (-1.37)^2 + 2(-1.37) = -0.86$.

$f(x_1) > f(x_2)$, New interval $(-1.93, 0.3)$

Check $K = n$.

$4 \neq 5$.

Proceed Further

4th Iteration :-

$K = K+1$

$= 4+1$

$K = 5$

$a = -1.93$

$b = -0.3$

$L = 7$, $n = 5$.

$$L_K^* = \left[\frac{F_{n-K+1}}{F_{n+1}}\right] \times L \Rightarrow \left[\frac{F_{5-5+1}}{F_{5+1}}\right] \times L$$

$$= \frac{F_1}{F_6} \times L \Rightarrow \frac{1}{13} \times 7.$$

$$L_K^* = 0.53$$

$x_1 = a + 0.53$

$= (-1.93) + 0.53$

$= -1.4$

$x_2 = b - 0.53$

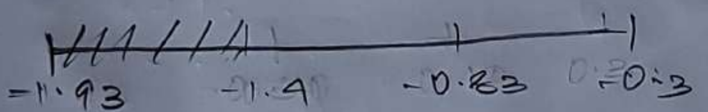
$= -0.3 - 0.53$.

$= -0.83$

$f(x_1) = (-1.4)^2 + 2(-1.4) = -0.84$.

$f(x_2) = (-0.83)^2 + 2(-0.83) = \cancel{-0.49}\ -0.98$

~~$f(x_2) > f(x_1)$~~, $f(x_1) > f(x_2)$



$K = n$.

$5 = 5$

Interval $(-1.4, -0.3)$

4) $f(x) = x^2 + 2x$ interval $(-3, 4)$, $n=6$.

$a=-3,\ b=4.$ $\quad L = 4-(-3)$

$L = 7$

Set $K=2$

$n=6$

$$L_K^* = \left[\frac{F_{n-K+1}}{F_{n+1}}\right] L$$

$$L_K^* = \left[\frac{F_{6-2+1}}{F_{6+1}}\right] = \frac{F_5}{F_7} \times L = \frac{8}{21} \times 7.$$

$L =$

$f(w_1) > f(w_2)$

Discard $(0.763, 1)$

New interval ~~(0.382, 0.618)~~. $(0.382, 0.763)$

3rd Iteration:-

$a_w = 0.382$, ~~$b_w = 0.618$~~.
$b_w = 0.763$.

$L_w = 0.763 - 0.382$

$L_w = 0.381$.

$w_1 = a_w + (0.618) L_w$

$= 0.382 + (0.618) \times 0.381$.

$w_1 = 0.617$ //

$w_2 = b_w - (0.618) L_w$

$= \cancel{1}\ 0.763 - (0.618) \times 0.381$

~~$w_2 = 0.764$~~ // $w_2 = 0.527$

$f(w_1) = 25w^2 + \frac{54}{5w}$

$= 25(0.617)^2 + \frac{54}{5(0.617)}$

$= 27.02$

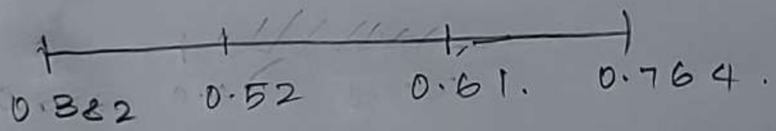
$f(w_2) = 25w^2 + \frac{54}{5w}$

$= 25(\cancel{0.764}\ 0.527)^2 + \frac{54}{5(\cancel{0.764}\ 0.527)}$

$=$ ~~28.72~~. ~~33.66~~.

$= 27.43$

~~$f(w_2) > f(w_1)$~~ $f(w_2) > f(w_1)$



Hold of Search Method:

⇒ It is an iterative process to find the minimum of a function in a certain domain.

Alg:

Step 1: Choose a lower bound 'a' & an upper bound 'b' & let e be a small number.

⇒ Normalise the variable 'x' by using the eqⁿ.

$$w = \frac{x-a}{b-a}. \quad a_w = 0, b_w = 1.$$

$L_w = b_w - a_w$, Set $a = 1$. ⇒ Golden ratio.

$= 1$

$w_1 = a_w + (0.618) L_w.$

$w_2 = b_w - (0.618) L_w.$

compute $f(w_1)$ & $f(w_2)$ use region elimination rule & set new a_w & b_w.

stop when $|L_w| < \varepsilon$.

~~Find the~~

1) Find minimum of $f(x) = x^2 + \frac{54}{x}$ interval $[0, 5]$.

1st Iteration:

$a = 0, b = 5.$

Normalise, $w = \frac{x-a}{b-a} \Rightarrow \frac{x-0}{5-0} \Rightarrow \frac{x}{5} \Rightarrow w$

$w = x/5$ $\quad x = 5w$

$f(x) = (5w)^2 + \frac{54}{5w}$ $\quad L_w = b_w - a_w = 1.$

$= 25w^2 + \frac{54}{5w}$ $\quad a_w = 0$, $b_w = 1.$

Compute w_1 & w_2.

$w_1 = a_w + (0.618) L_w.$

$= 0 + (0.618) \times 1 = 0.618$

$w_2 = b_w - (0.618) L_w$

$= 1 - (0.618) \times 1 = 0.382$

$$f(\omega_1) = 25\omega^2 + \frac{54}{5\omega}$$

$$= 25(0.618)^2 + \frac{54}{5(0.618)} = 27.02$$

$$f(\omega_2) = 25\omega^2 + \frac{54}{5\omega}$$

$$= 25(0.382)^2 + \frac{54}{5(0.382)} = 31.92$$

$f(\omega_1) < f(\omega_2)$

Discard $(0, 0.382)$

New interval $(0.382, 1)$ 2nd Iteration:-

$a_\omega = 0.382$, $b_\omega = 1$. ~~$a_\omega = 0.382$~~ ~~$b_\omega = 1$.~~

$L_\omega = 1 - 0.382$

$L_\omega = 0.618$.

Compute ω_1 & ω_2

$\omega_1 = a_\omega + (0.618) L_\omega$

$= 0.382 + (0.618) \times 0.618$

$\omega_1 = 0.763$

$\omega_2 = b_\omega - (0.618) L_\omega$

$= 1 - (0.618) \times 0.618$

$\omega_2 = 0.618$

$$f(\omega_1) = 25\omega^2 + \frac{54}{5\omega}$$

~~$= 25(0.618)$~~

$$= 25(0.763)^2 + \frac{54}{5(0.763)}$$

$f(\omega_1) = 28.70$

$$f(\omega_2) = 25\omega^2 + \frac{54}{5\omega}$$

$$= 25(0.618)^2 + \frac{54}{5(0.618)}$$

$f(\omega_2) = 27.02$

Arvin Jayakumar

Supervised learning:

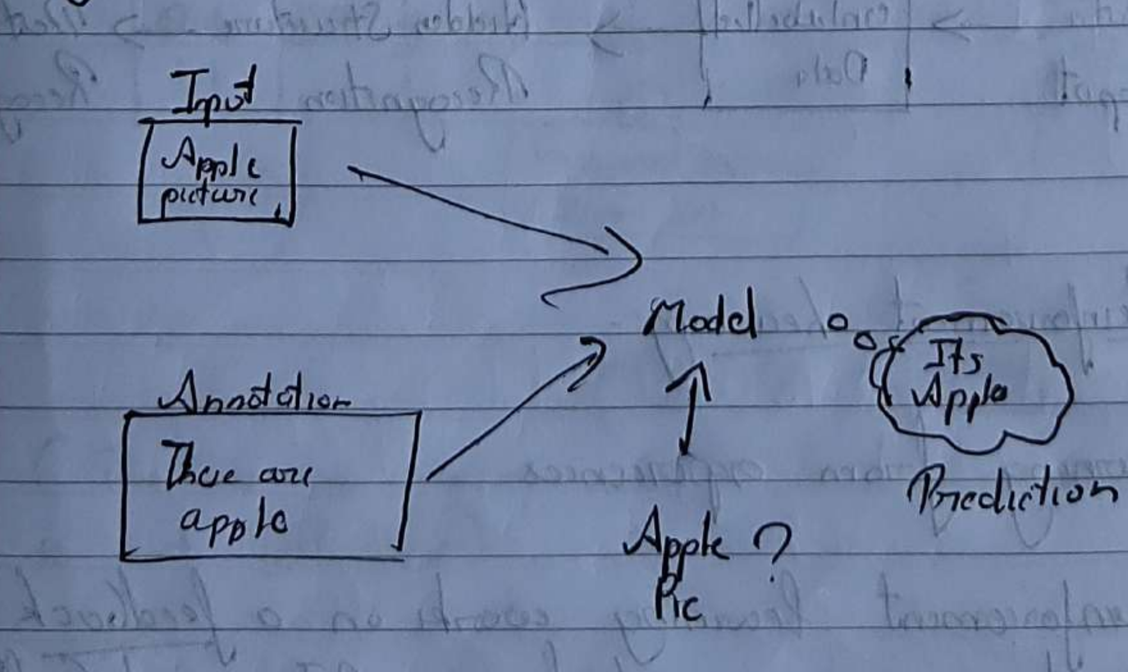
Supervised machine learning is based on supervision.

In this technique, we train the machine using the "labelled" dataset and based on the training the machine predicts the outcome.

The main goal of the supervised learning technique is to map the input variable (x) with output variable y.

Supervised learning Type:

i Classification
ii Regression



Module 3
Machine Learning

Machine Learning is a growing technology that enable computer to learn automatically from past data.

It uses various algorithms for building mathematical models and making prediction using historical data or information.

The term machine learning was introduced by Arthur Samuel in 1959.

Def -> Machine learning enables a machine to automatically learn from data improve performance from experiences and predict things.

Types of ML
i Supervised Learning
ii Unsupervised Learning
iii Reinforcement learning

Arvin Jayakumar

Supervised learning :-

Supervised machine learning is based on supervision.

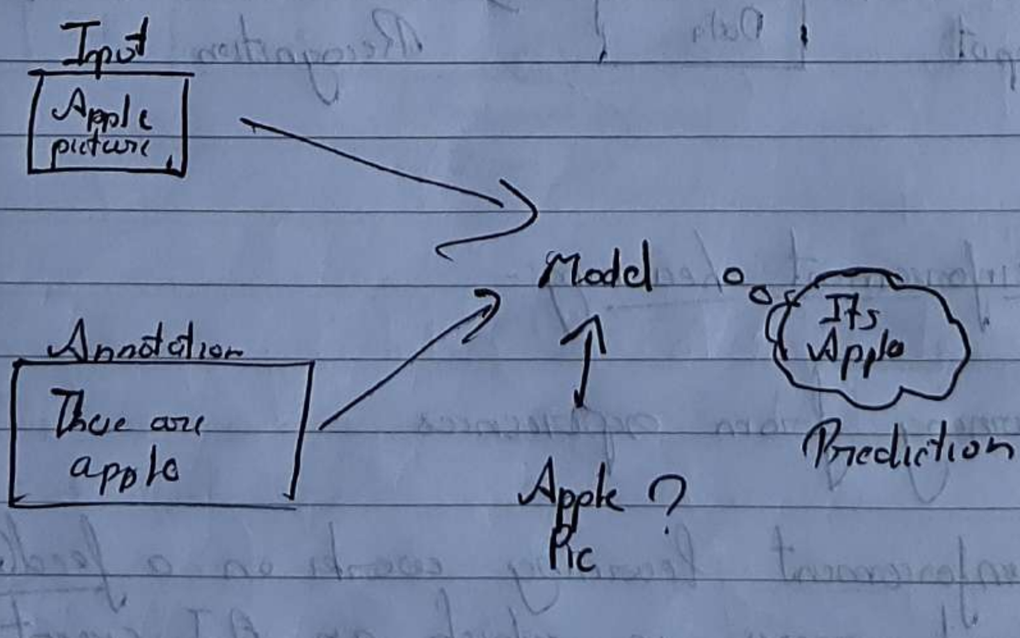
In this technique, we train the machine using the "labelled" dataset and based on the training the machine predicts the outcome.

The main goal of the supervised learning technique is to map the input variable (x) with output variable y.

Supervised learning Types:

i Classification

ii Regression

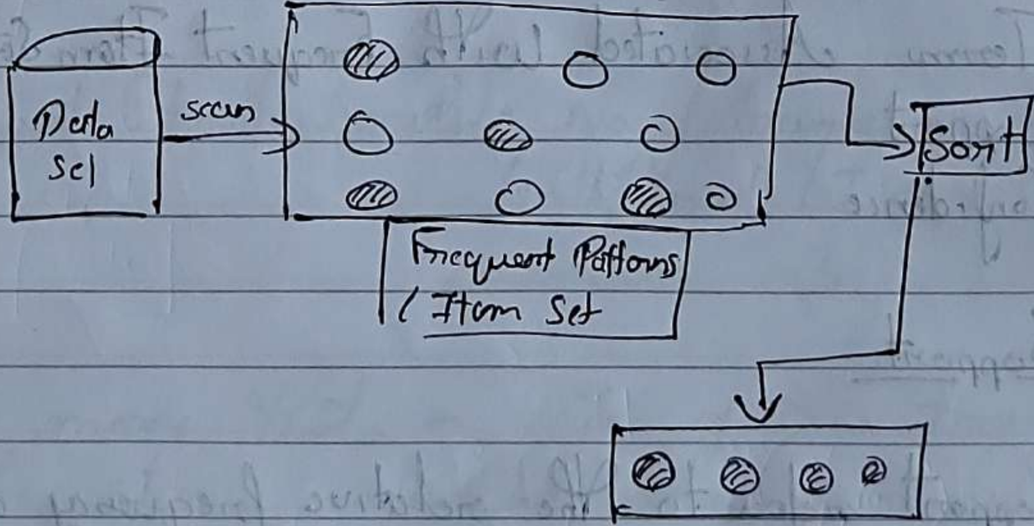


component) automatically explore it's surroundings, learning from experiences and improving it's performance.

→ The agent gets rewarded for each good action and get punished for each bad action, hence the goal of reinforcement learning agent is to maximize the rewards.

Frequent Pattern Mining :-

Frequent pattern refers to itemset, subsequences or substructures that appear frequently in a dataset.



Sorting (This is an unsupervised dataset because the data is not labelled)

Unsupervised Learning:-

→ There is no supervision

→ The machine is trained using unlabelled dataset and the machine predicts the output.

→ Aim of this algorithm is ~~used~~ to group or categorize the dataset according to similarities, pattern and difference.

Two Types

i. Clustering

ii. Association

Data Input → Unlabelled Data → Hidden Structure Recognition → Data Recognition

Reinforcement Learning:-

→ Learning from experiences

→ Reinforcement learning works on a feedback based process in which an AI agent (Software

by the total number of transactions

Support is calculated as follows
-> Support (X) = (Number of transactions containing X) / Total no of transactions.
where X is the item set for which we are calculating support.

For example, suppose we have a dataset of 1000 transactions and the itemset {milk, bread} appears in 100 of those transaction.

The support of the itemset {milk, bread} would be calculated as follows

Support ({milk, bread}) = No of transactions containing {milk & bread} / Total no of transaction
= 100/1000 = 10%

This means that in 10% of the transaction, the items milk and bread were both purchased

Algorithms that can be used for classification using frequent pattern mining

i Apriori Algorithm
ii FP Growth Algorithm

a Consider the dataset

Transaction id	items
T1	hot dogs, buns, ketchup
T2	hot dogs, buns
T3	hot dogs, coke, chips
T4	chips, coke
T5	chips, ketchup
T6	hotdogs, coke, chips

Find the frequent item set and generate association rules using apriori algorithm. Assume the minimum support threshold is 33.33 % and minimum confidence is 60 %.

Minimum Support Count

$$= \text{No of transactions} \times \frac{\text{support}}{100}$$

$$= \cancel{6}^{3} \times \frac{33.33}{\cancel{100}_{50}} = 1.99 = 2$$

Step 1

Item	Support Count
Hot Dogs	4
Buns	2
Ketchup	2
Coke	3
Chips	4

If any item having the count less than minimum support count, then we need to remove that item (Data pruning)

Step 2

Item	Support Count	
Hot Dogs, Buns	2	
Hot Dogs, Ketchup	1	x
Hot Dogs, Coke	2	
Hot Dogs, Chips	2	
Buns, Ketchup	1	x
Buns, Coke	0	x
Buns, Chips	0	x
Ketchup, Coke	0	x
Ketchup, Chips	1	x
Coke, Chips	3	

Item	Min Support
Hotdogs, chips, coke	2

Frequent itemset: {hot dogs, chips, coke}

Association Rule

{hotdogs, coke} ⟹ {chips}

$$\text{Confidence} = \frac{\text{Support count}(\{\text{hot dogs, chips}\})}{\text{Support count}(\{\text{hotdogs, coke}\})} \times 100$$

$$= \frac{2}{2} \times 100 \Rightarrow 100\%$$

{hotdogs, chips} ⟹ {coke}

$$\text{Confidence} = \frac{\text{Support count}(\{\text{hot dogs, } \cancel{\text{chips}}\ \text{coke}\})}{\text{Support count}(\{\text{hot dogs, } \cancel{\text{coke}}\ \text{chips}\})} \times 100$$

$$= \frac{2}{2} \times 100 = 100\%$$

{coke, chips} ⟹ {hotdogs}

$$\text{Confidence} = \frac{\text{Support count}(\{\text{coke, } \cancel{\text{chips}}\ \text{hotdog}\})}{\text{Support count}(\{\text{coke, } \cancel{\text{hotdogs}}\ \text{chips}\})} \times 100$$

$$= \frac{2}{3} \times 100 = 66.66\%$$

Step 3 Data Pruning

Item	Support Count
Hot Dog, bun	2
Hot Dogs, coke	2
Hot Dogs, chips	2
Coke, Chips	3

Step 4 Combination of 3

Item	Support Count
Hot Dogs, Buns, Ketup	1
Hot Dogs, Buns, Coke	0
Hot Dogs, Bun, Chips	0
Hot Dogs, Ketup, Coke	0
Hot Dogs, Ketchup, Chips	0
Hot Dogs, Coke, Chips	2
Buns, Ketup, Coke	0
Bun, Ketchup, Chips	0
Ketchup, Coke, Chips	0
Bun, Coke, Chips	

Step 5 Data Pruning