Ⓠ find LDA of given data :-

Infection	CRP (mg/L)	Temp (°c)
Viral	40.0	36.0
Viral	11.1	37.2
Viral	30.0	36.5
Viral	21.4	39.4
Viral	10.7	39.6
Viral	3.4	40.7
Bacterial	42.0	37.6
Bacterial	31.1	42.2
Bacterial	50.0	38.5
Bacterial	60.4	39.4
Bacterial	45.7	38.6
Bacterial	17.3	42.7

Ⓐ

Step-1 :-

$$\text{cov}_{\text{Viral Infection}} = \begin{bmatrix} 188 & -21 \\ -21 & 4 \end{bmatrix}$$

$$\text{cov}_{\text{Bacterial Infection}} = \begin{bmatrix} 228 & -24 \\ -24 & 4 \end{bmatrix}$$

Step-2 :-

$$W = \begin{bmatrix} \frac{188+228}{2} & \frac{-21-24}{2} \\ \frac{-21-24}{2} & \frac{4+4}{2} \end{bmatrix}$$

$$W = \begin{bmatrix} 208 & -22.5 \\ -22.5 & 4 \end{bmatrix}$$

Step-3 :-

$$\text{Total} = \text{Covar}_{\text{Total}} = \begin{bmatrix} 317.1 & -11.0 \\ -11.0 & 4.4 \end{bmatrix}$$

Step-4 :-

$$B = T - W = \begin{bmatrix} 317.1 - 208 & -11.0 + 22.5 \\ -11.0 + 22.5 & 4.4 - 4 \end{bmatrix}$$

$$= \begin{bmatrix} 109.1 & 11.5 \\ 11.5 & 0.4 \end{bmatrix}$$

Step-5 :-

$$W^{-1} = \begin{bmatrix} 0.012 & 0.066 \\ 0.066 & 0.609 \end{bmatrix}$$

$$S = W^{-1} B$$

$$= \begin{bmatrix} 0.012 & 0.066 \\ 0.066 & 0.609 \end{bmatrix} \begin{bmatrix} 109.1 & 11.5 \\ 11.5 & 0.4 \end{bmatrix}$$

$$\therefore S = \begin{bmatrix} 2.05 & 0.16 \\ 14.15 & 0.96 \end{bmatrix}$$

Step-6 :-

Eigen values of S are $\lambda_1 = 3.1053$

$\lambda_2 = -0.095$

Eigen vector of $\lambda_1 = 3.1053$ is

$$\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 0.150 \\ 0.989 \end{bmatrix}$$

Step 7)

$LD_1 = 0.150 \times CRP + 0.989\ Temp$

Scores:-

$\boxed{41.6, 38.4, 40.6, 42.2, 40.8, 40.8,}$ $\boxed{43.5, 46.4,}$

$\boxed{45.5, 48, 45, 44.8}$

$Var(Scores)_{Viral} = 1.60$

$Var(Scores)_{Bacterial} = 2.37$

$Var(Scores)_{pooled} = 1.985$

$$LD_1 = \frac{0.150}{\sqrt{1.985}} \times CRP + \frac{0.989}{\sqrt{1.985}} \times Temp.$$

$$LD_1 = 0.11(CRP) + 0.70(Temp).$$

Updated Scores:-

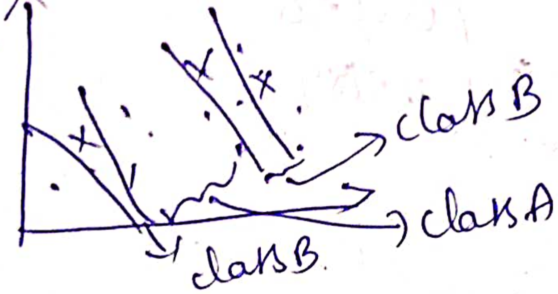
29.5, 27.3, 28.8, 29.9, 28.9, 28.9, 30.8, 32.9,
32.3, 34.0, 31.9, 31.8.

Cent Scores:-

−1.1, −3.3, −1.8, −0.7, −1.7, −1.7,
0.2, 2.3, 1.7, 3.5, 1.3, 1.2.

Decision Tree ⇒ Non linearly independ
seperable data.

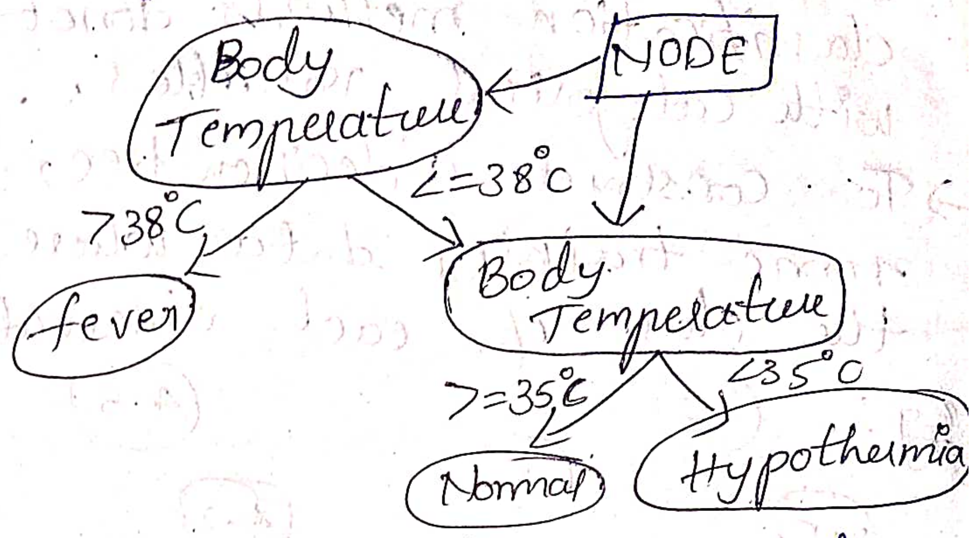
Impurity level ⇒ Gini level



* Decisio

→ consider the following decision tree to determine if someone has fever or hypothermea based on a person's body temperature in celsius.

→

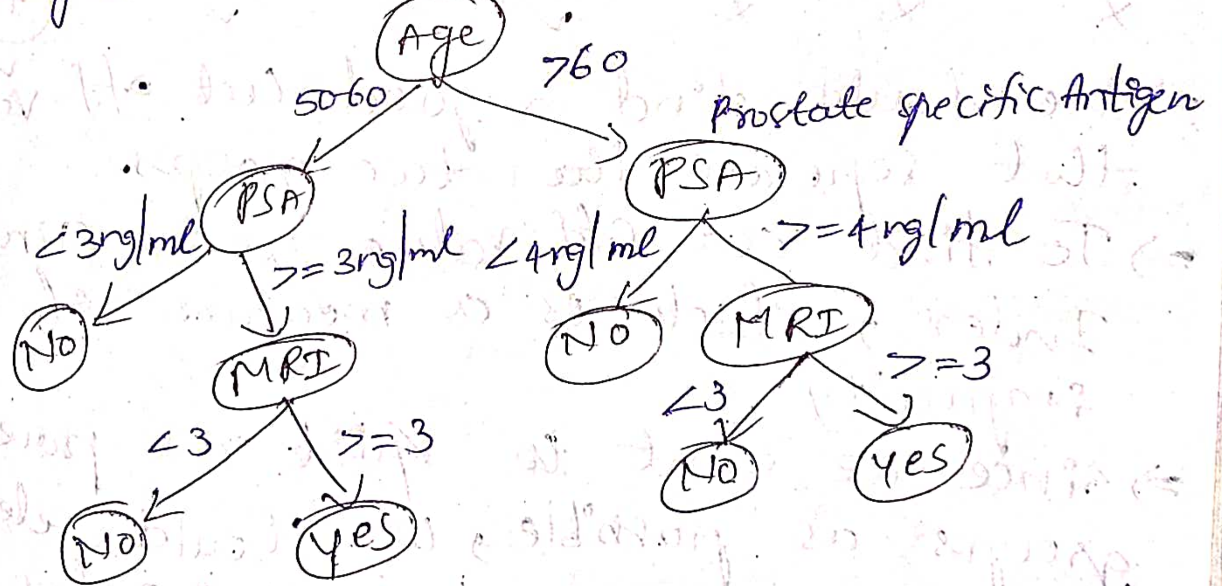


A classification tree consists of nodes, which represent a choice between alternatives based on the value of the measured variable.

A node at the top of the tree is called the "root node".

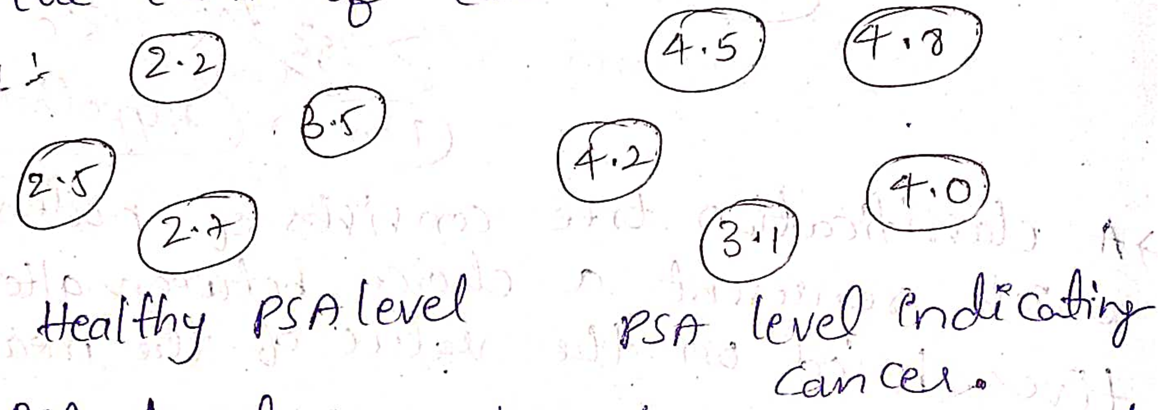
→ Leaf node represents the decision.

The arrows are called the branches (or) edges.

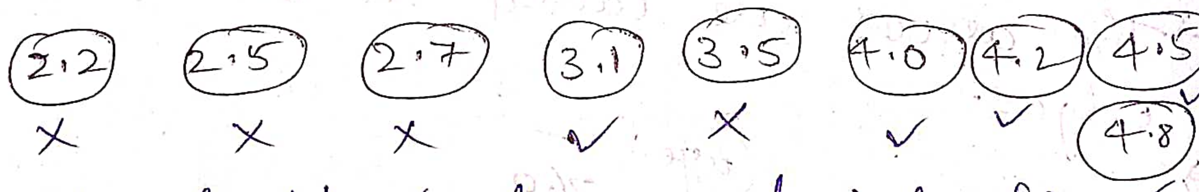


- → Decision Trees work fine with variables on both scales. We can even have a variable ~~on~~ or a nominal scale (non-numeric). This is one advantage with decision trees because most other classification methods donot work well with categorical variables.
- → To construct a decision tree, we need some training data where we know the class of each observation.

Eg :-



- ⇒ PSA level is a bit higher in affected persons compared to the healthy ones.
- ⇒ We begin ~~to~~ ~~sort~~ individuals based on their PSA concentrations.



- ⇒ We should find a good cut off value that seperate the two groups.
- ⇒ To find cut off value we use Gini Index, which is a measure of impurity.
- ⇒ since we want ~~to~~ ~~split~~ as pure groups as possible, we should select that results in a low Gini index as possible.

⇒ We [illegible] using the mean value of any two adjacent datapoints

$$Gini(T) = 1 - \sum_{j=1}^{n} p_j^2$$ (where p is the proportion in each class)

⇒ A cutoff value at 2.6 will split the data into 2 groups.

⇒ .2 2.5 | 2.7 3.1 3.5 4.0 4.2 4.5 4.8

× × × ✓ × ✓ ✓ ✓ ✓

⇒ lets calculate Gini index of second subgroup.

$$\Rightarrow Gini(T_2) = 1 - \sum_{j=1}^{n} p_j^2$$

$$= 1 - \left(\left(\frac{2}{7}\right)^2 + \left(\frac{5}{7}\right)^2\right)$$

$$= 0.41$$

Two out of 7 samples in the subgroup are healthy individuals, whereas five others are patients with prostate cancer.

$$\Rightarrow Gini(T_1) = 1 - \sum_{j=1}^{n} p_j^2$$

$$= 1 - \left(\left(\frac{2}{2}\right)^2 + \left(\frac{0}{2}\right)^2\right)$$

$$= 1 - 1 = 0$$

The Gini index is equal to zero which means that we have 100% purity or 0% impurity.

⇒ Then we calculate the weighted average of these two Gini indexes.

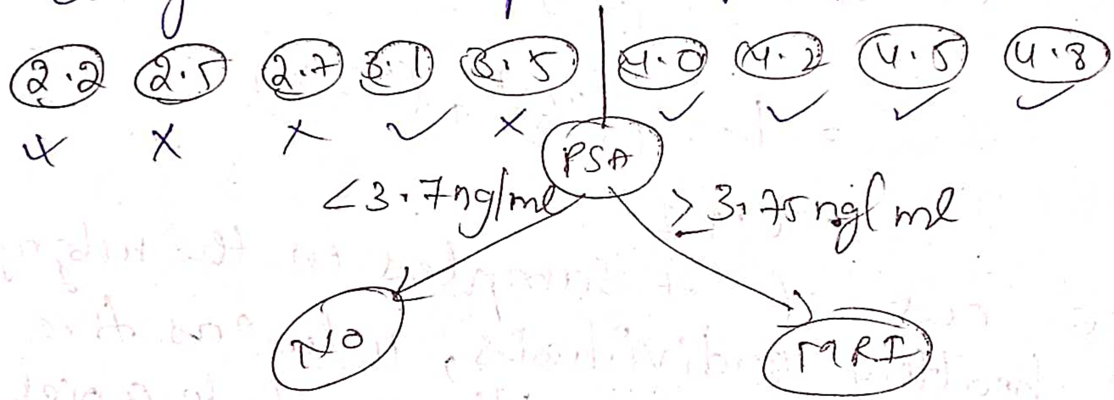
$$Gini(T) = \frac{n_1}{N} Gini(T_1) + \frac{n_2}{N} Gini(T_2)$$

$$= \frac{2}{9} \times 0 + \frac{7}{9} \times 0.41$$

$$= 0.32$$

⇒ By using a cutoff value of 2.6, the Gini Index is equal to 0.32

If cutoff value is 2.9 then Gini Index is 0.19. Similarly, for the cut off value of 3.75 then gini index is 0.18. We should select low gini index value because it generates purest groups.



⇒ We can therefore calculate the accuracy of this tree based on training data by dividing the true negative & positive by the total no. of observations.

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP}$$

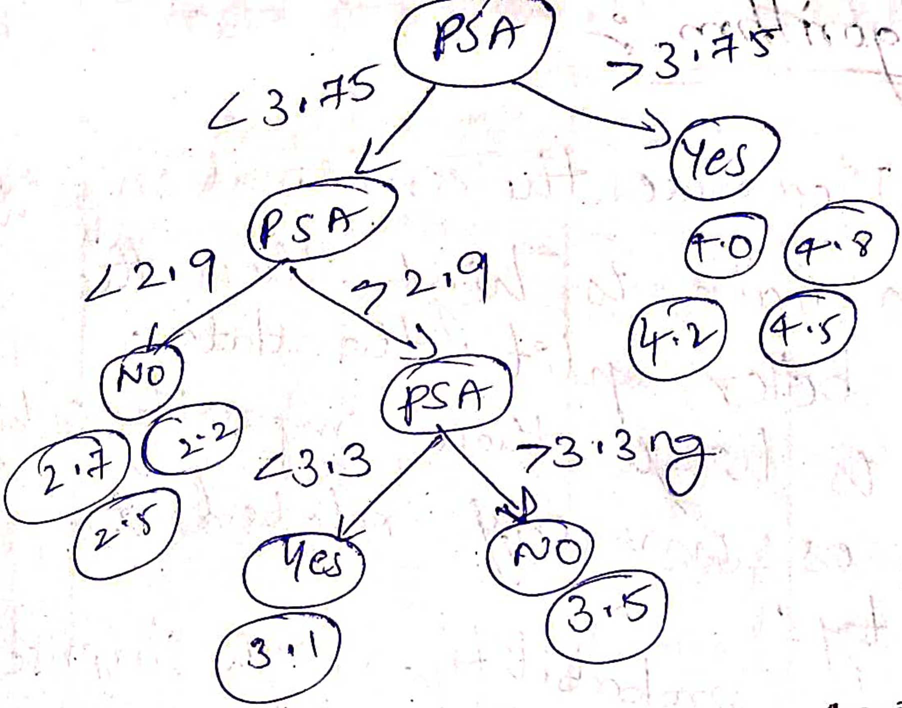
TN ⇒ ~~Posi~~ (Negative cancer) True Negative.

TP ⇒ (Positive Health) True positive.

FN ⇒ false ~~posi~~ Negative

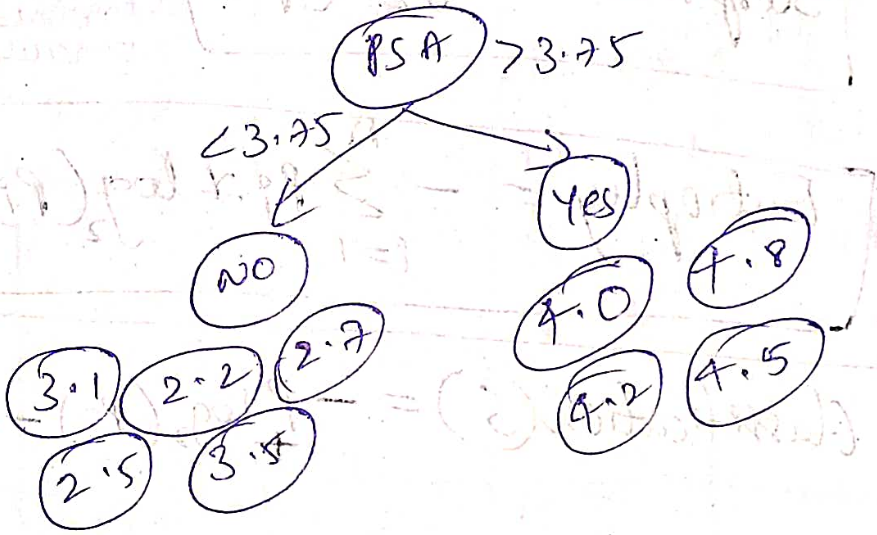
FP ⇒ false positive

$$Accuracy = \frac{4+4}{4+4+1+0} = 0.89 = \boxed{89\%}$$



It is difficult to apply a different set of data points due to overfitting

$$\text{Accuracy} = \frac{4+5}{4+5+0+0} = 1 = \underline{100}\%$$



$$\text{Accuracy} = \frac{8}{9} = 89\%$$

Decision Trees

Classification (Categorical)

Regression Tree (Numerical non-linear)

ID3 Algorithm :

Entrophy :- Measure the information gain Information Gain is high then the data is better split the data.

Entrophy is high then impurity is high

* Surprise is inversely related to probability.

① → Low probability → high surprise

② → High probability → low surprice.

$$\boxed{\text{Surprise} = \log_2\left(\frac{1}{P}\right)}$$

0 for probability one
undefined for unoccured event

$$\boxed{\text{Entrophy} = -\sum_{i=1}^{n} P_i * \log_2(P_i)}$$

$$\boxed{\text{Binomial Classification}(S) = -P_+\log_2(P_i) - P_-\log_2(}$$

Entrophy([a+, 0-]) => 0

Entrophy([0+, a-]) => 0

Entrophy([a+, a-]) => 1

→ Highly Information gain attribute is selected as "Rood Node" in decision tree.

Find ID3 Algorithms & draw decision tree.

Day	Outlook	Temp	Humidity	Wind	Play Tennis
D1	Sunny	Hot	High	Weak	NO
D2	Sunny	Hot	High	Strong	NO
D3	overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	weak	Yes
D5	Rain	Cool	Normal	weak	Yes
D6	Rain	cool	Normal	Strong	NO
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	weak	NO
D9	Sunny	Cool	Normal	weak	Yes
D10	Rain	Mild	Normal	weak	Yes
D11	Sunny	Mild	Normal	~~weak~~ strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	weak	Yes
D14	Rain	Mild	High	Strong	NO

$$E(S) = -\frac{9}{14}\log_2\left(\frac{9}{14}\right) - \frac{5}{14}\log_2\left(\frac{5}{14}\right) = 0.94$$

$$\text{Gain}(S, A) \equiv \text{Entrophy}(S) - \sum \frac{|S_v|}{|S|} \text{Entrophy}(S_v)$$

ⓐ

$S_{strong} \rightarrow [3+, 3-]$

$E(S_s) = 1$

$S_{weak} \rightarrow [6+, 2-]$

$$E(S_W) = -\frac{6}{8}\log_2\left(\frac{6}{8}\right) - \frac{2}{8}\log_2\left(\frac{2}{8}\right)$$

$= ~~0.093~~\ 0.3112 + 0.5$

$= 0.8112$

$\mathrm{Gain}(S, \text{wind}) = E(S) - \frac{6}{14} E(S_S) - \frac{8}{14} (S_W)$

$\mathrm{Gain}(S, \text{wind}) = 0.94 - \frac{6}{14} \times 1 - \frac{8}{14} \times 0.8112$

$= 0.94 - 0.4285 - 0.4635$

$\boxed{\therefore \mathrm{Gain}(S, \text{wind}) = 0.048}$

ⓑ $S_{\text{sunny}} \rightarrow [2+, 3-]$

$E(S_S) = -\frac{2}{5} \log_2\left(\frac{2}{5}\right) - \frac{3}{5} \log_2\left(\frac{3}{5}\right)$

$= 0.5287 + 0.4421$

$= 0.9708$

$S_{\text{overcast}} \rightarrow [4^+, 0-]$

$E(S_O) = 0$

$S_{\text{Rain}} \rightarrow [3+, 2-]$

$E(S_R) = 0.9708$

$\mathrm{Gain}(S, \text{Outlook}) = E(S) - \frac{5}{14} E(S_S) - \frac{4}{14} E(S_O) - \frac{5}{14} E(S_R)$

$= 0.94 - \frac{5}{14} \times 0.9708 - \frac{5}{14} \times 0.9708$

$= 0.94 - 0.3467 - 0.3467$

$= 0.2465$

$\boxed{\therefore \mathrm{Gain}(S, \text{overcast}) = 0.2465}$

ⓒ

$S_{\text{Hot}} \rightarrow [2+, 2-] \quad E(S_H) = 1$

$S_{\text{Mild}} \rightarrow [4+, 2-] \quad E(S_M) = 0.91383$

$S_{\text{Cool}} \rightarrow [3+, 1-] \quad E(S_C) = 0.8112$

$\mathrm{Gain}(S, \text{Temp}) = E(S) - \frac{4}{14} E(S_H) - \frac{6}{14} E(S_M) - \frac{4}{14} E(S)$

$= 0.94 - \frac{4}{14}(1) - \frac{6}{14}(0.9183) - \frac{4}{14}(0.8112)$

$= 0.94 - 0.2857 - 0.4587 - 0.2317$

$=$ ~~0·2719~~ ~~0·27079~~ 0.0441

$\boxed{\therefore \text{Gain}(S, \text{Temp}) = 0.0441}$

ⓓ ~~$S_{weak} \rightarrow [6+, 2-] \Rightarrow E(S_W) = 0.8113$~~

~~$S_{strong} \rightarrow [3+, 3-] \rightarrow E(S_S) = 1$~~

~~Gain(S, wind)~~

$S_{High} \rightarrow [3+, 4-] \Rightarrow E(S_H) = 0.9852$

$S_{Normal} \rightarrow [6+, 1-] \Rightarrow E(S_N) = 0.5916$

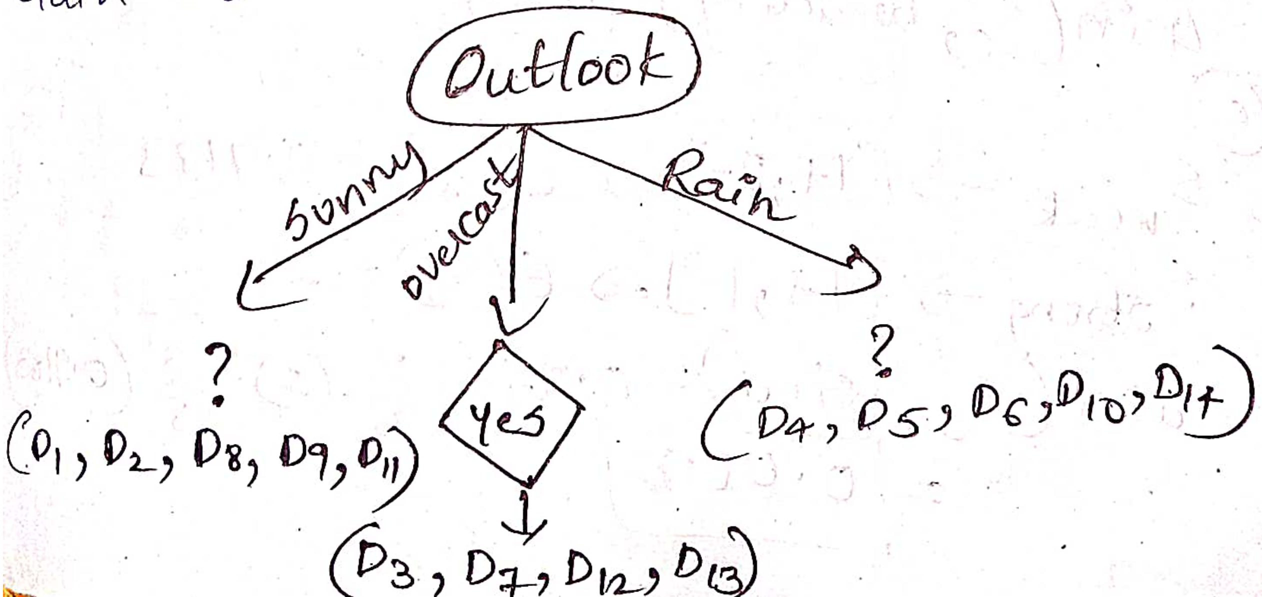
$\text{Gain}(S, \text{Humidity}) = E(S) - \frac{7}{14}(E S_H) - \frac{7}{14} E(S_H)$

$= 0.94 - \frac{7}{14}(0.9852) - \frac{7}{14}(0.5916)$

$= 0.1516$

$\boxed{\therefore \text{Gain}(S, \text{Humidity}) = 0.1516}$

Here outlook has high Information Gain. so outlook is root node.



Left ~~Right~~ Subtree Node:

Day	Temp	Humidity	wind	Play Tennis
D_1	Hot	High	weak	No
D_2	Hot	High	strong	NO
D_8	Mild	High	weak	No
D_9	Cool	Normal	weak	Yes
D_{11}	Mild	Normal	strong	Yes

ⓐ

$S_{Hot} \rightarrow [0+, 2-] = 0$

$S_{Mild} \rightarrow [1+, 1-] = 1$

$S_{cool} \rightarrow [1+, 0-] = 0$

$E(S) = -\frac{3}{5}\log_2\left(\frac{3}{5}\right) - \frac{2}{5}\log_2\left(\frac{2}{5}\right)$

$= 0.97$

$Gain(S_5, Temp) = 0.97 - \frac{2}{5} \times 1 = \boxed{0.57}$

ⓑ

$S_{High} \rightarrow [0+, 3-] = 0$

$S_{Normal} \rightarrow [2+, 0-] = 0$

$Gain(S_5, Humidity) = \boxed{0.97}$

ⓒ

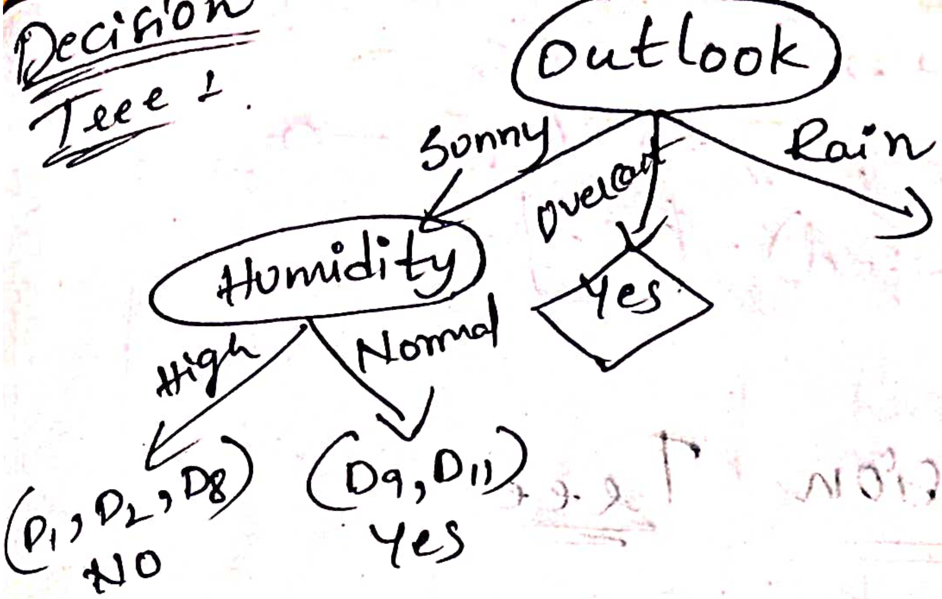
$S_{weak} \rightarrow [1+, 2-] \Rightarrow E(S_w) = 0.9183$

$S_{strong} \rightarrow [1+, 1-] \Rightarrow E(S_s) = 1$

$Gain(S_5, wind) = 0.97 - \frac{2}{5}(1) - \frac{3}{5}(0.9183)$

$= \boxed{0.0192}$

Decision Tree :



Right Subtree Node :

Day	Temp	Humidity	wind	Play Tennis
D4	Mild	High	weak	Yes
D5	cool	Normal	weak	Yes
D6	cool	Normal	strong	No
D10	Mild	Normal	weak	Yes
D14	Mild	High	strong	No

ⓐ $S_{Hot} \rightarrow [0+, 0-] \Rightarrow E(S_H) \Rightarrow 0$

$S_{Mild} \rightarrow [2+, 1-] \Rightarrow E(S_M) \Rightarrow 0.9183$

$S_{cool} \rightarrow [1+, 1-] \Rightarrow E(S_C) \Rightarrow 1$

$E(S) = -\frac{3}{5}\log\left(\frac{3}{5}\right) - \frac{2}{5}\log\left(\frac{2}{5}\right) = 0.97$

$Gain(S_R, T) = 0.97 - \frac{3}{5}(0.9183) - \frac{2}{5}(1)$

$= \cancel{0.01}\ \boxed{0.0192}$

ⓑ

$S_{High} \rightarrow [1+, 1-] \Rightarrow E(S_H) \Rightarrow 1$

$S_{Normal} \rightarrow [2+, 1-] \Rightarrow E(S_N) \Rightarrow 0.9183$

$Gain(S_R, Humidity) = 0.97 - \frac{2}{5}(1) - \frac{3}{5}(0.9183)$

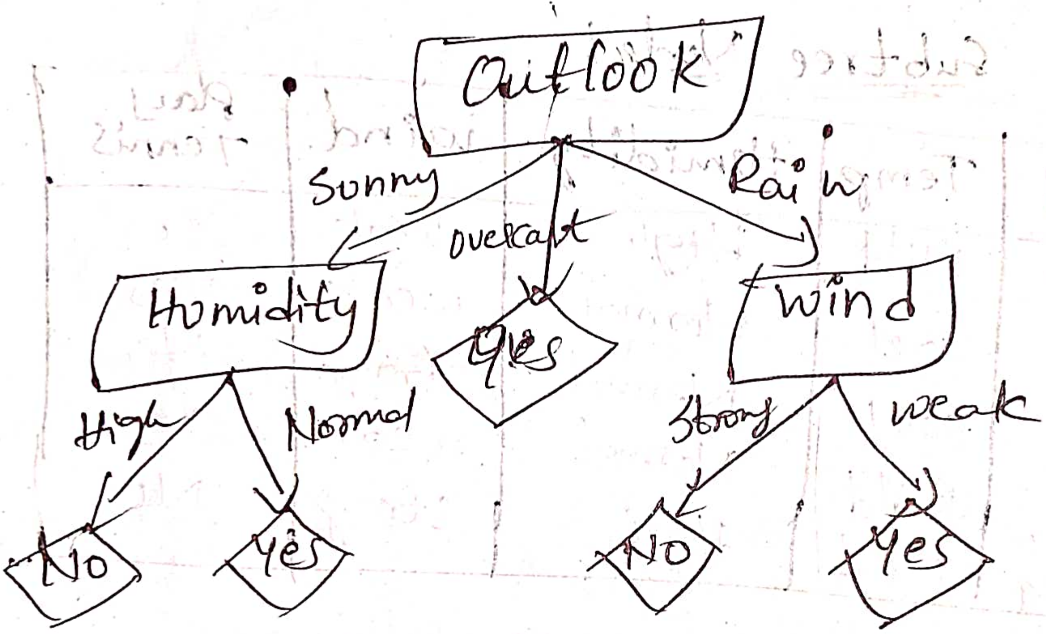
$= \boxed{0.0192}$

ⓒ $S_{strong} \rightarrow [0+, 2-] \Rightarrow E(S_S) = 0$

$S_{weak} \rightarrow [3+, 0-] \Rightarrow E(S_W) = 0$

$Gain(S_R, Wind) = \boxed{0.97}$

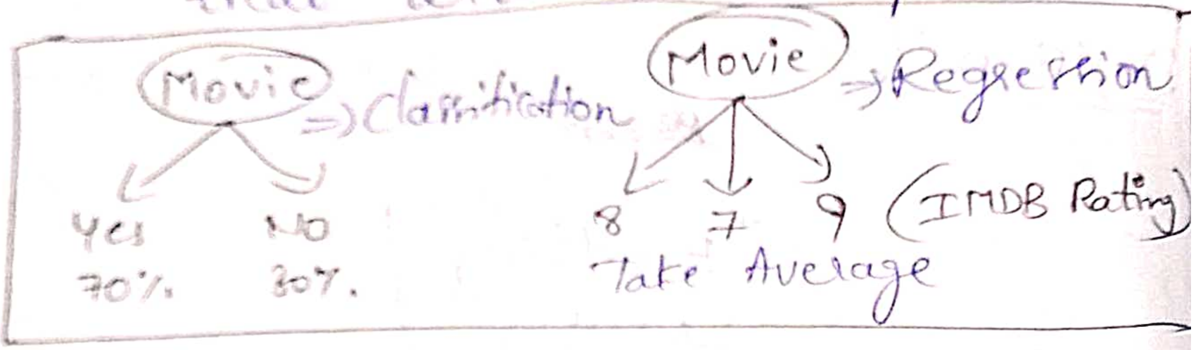
Decision Tree



***Ensemble Learning Techniques:** It is a mix of more ml techniques and apply majority outcome as predicted outcome. It is taking decision from more ml models. Decision is highly accurate.

Basic Techniques:

1) classification → Voting (categorical)
2) Regression → Average (Numerical)
3) Weighted Average → Some ml models more reliable source we should treat with more importance.

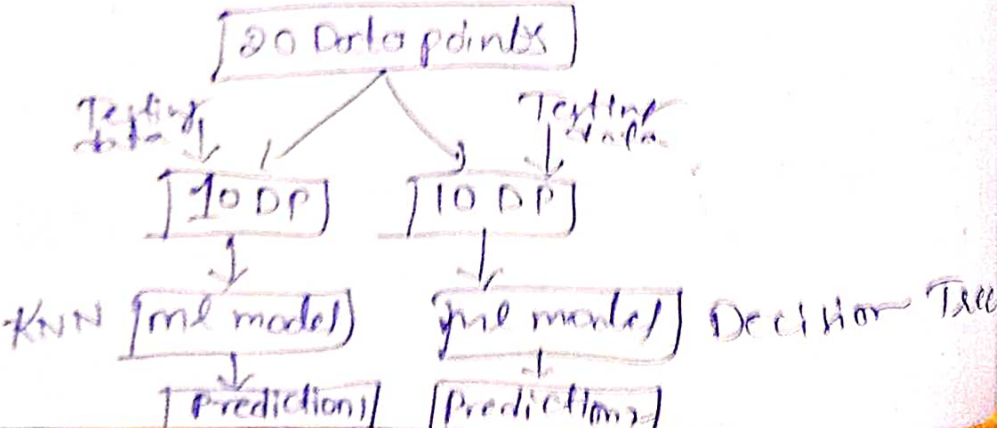


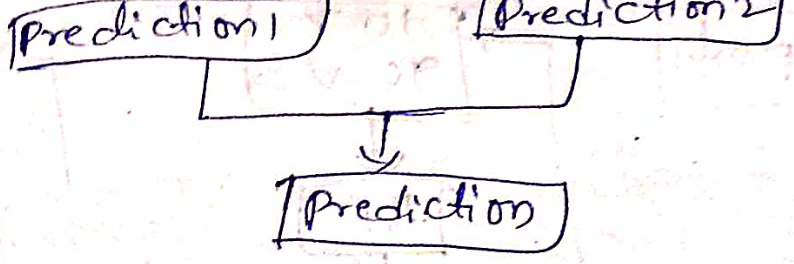
Advanced Ensemble Learning Techniques:

1) Bagging
2) Boosting
3) Stacking

1) **Bagging:** - Input will be divided into subsets.

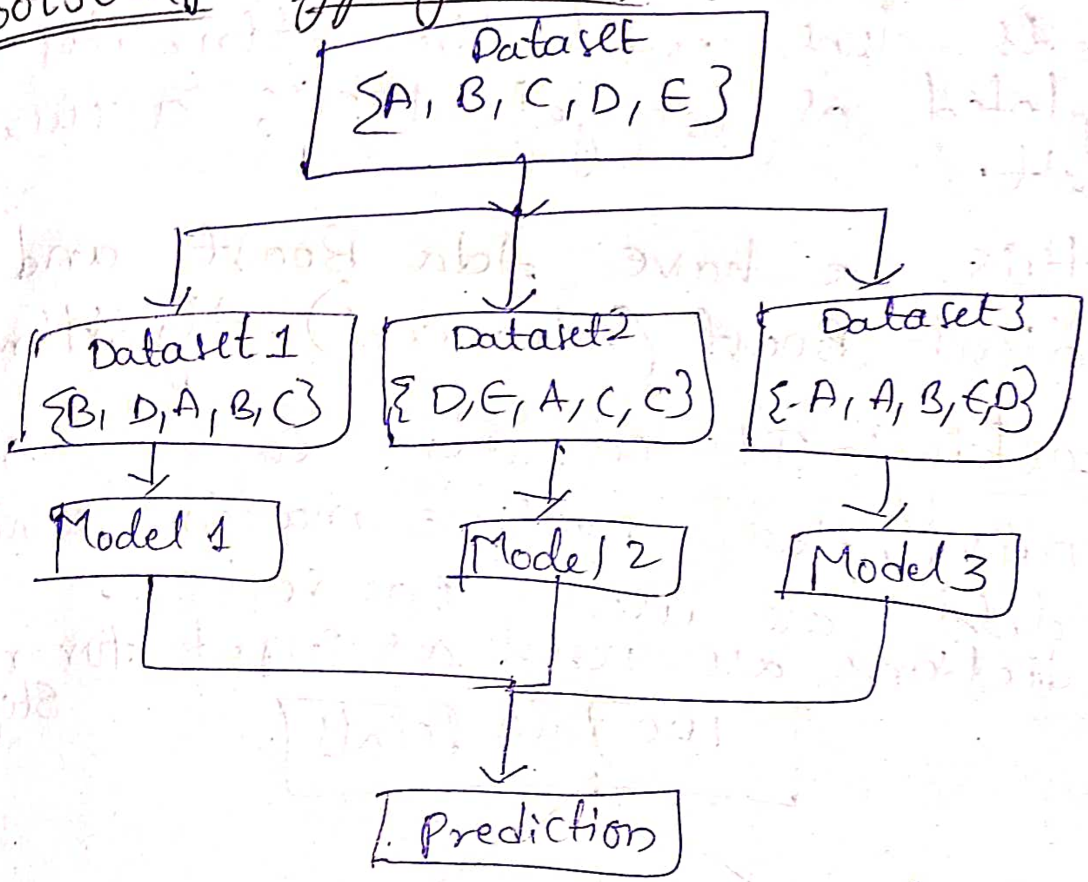
eg:





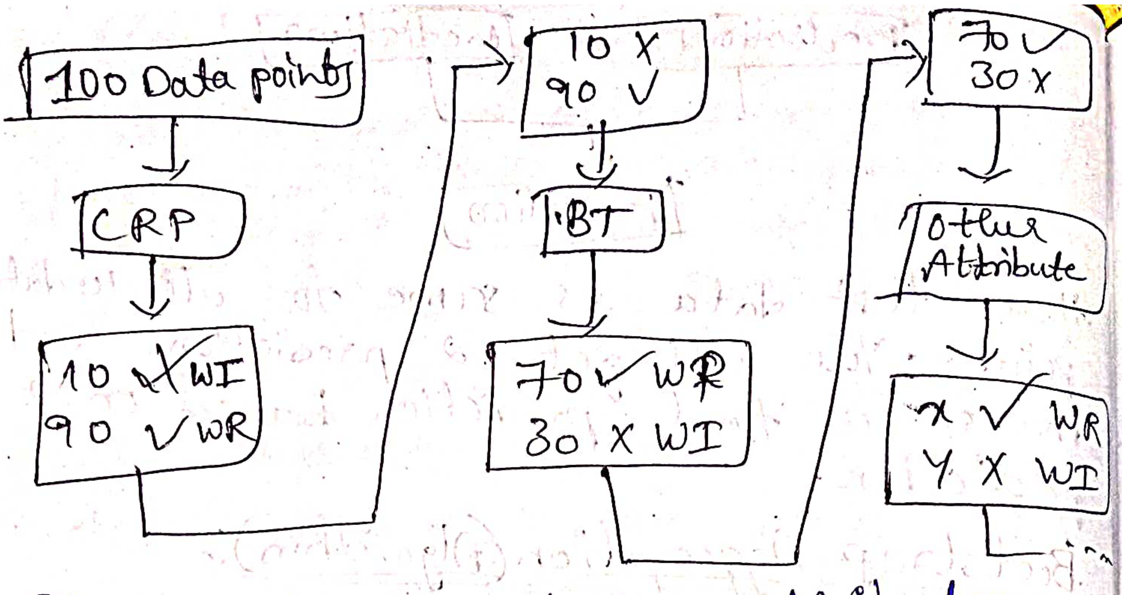
Here test data is same for all 10 data points. You will get 2 predictions and make a final prediction based on 2 prediction.

Bootstrap Aggregation (Algorithm):-



It refers to a technique where multiple dataset are created by randomly sampling with replacement from the original dataset.

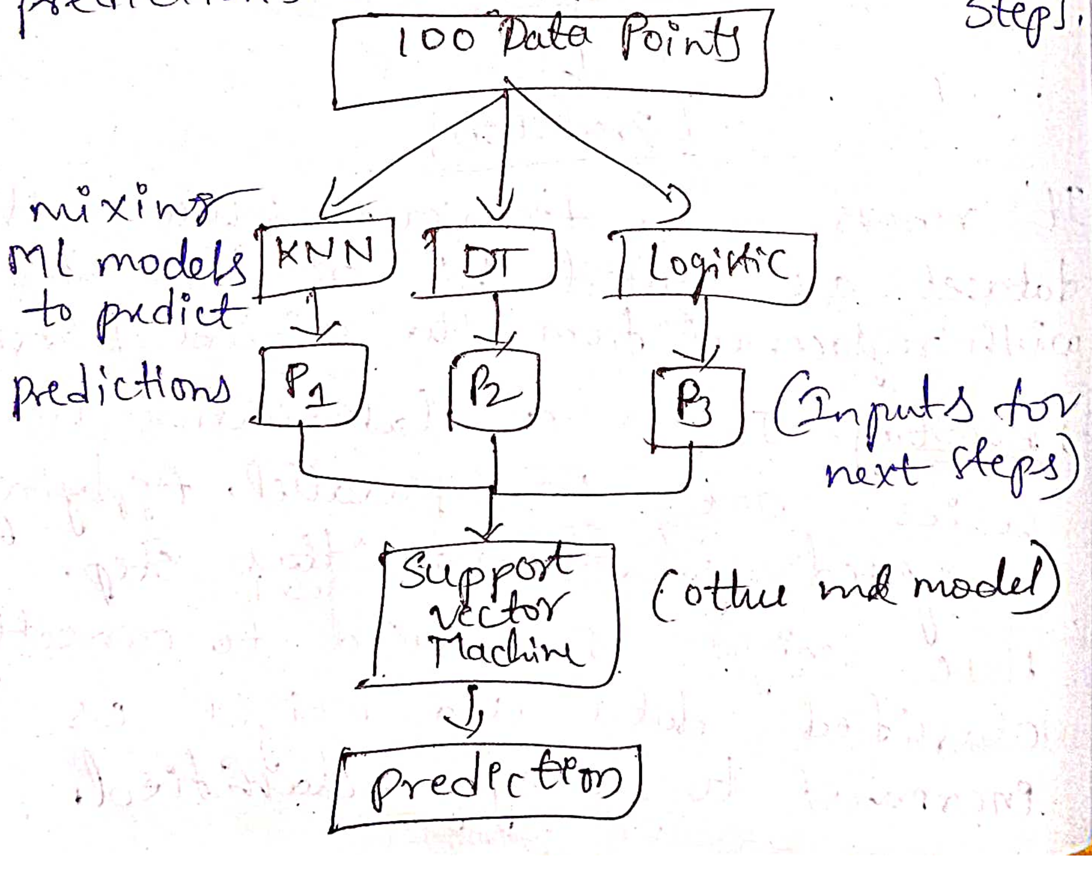
) **Boosting**:- It is a classification in series not in parallel. Applying weighted updates to ~~other~~ next step. Here weight is reduced to correctly classified data and weight is increased to wrongly classified.



It is done until the attributes are completed or we get 100% accurate result.

In this we have Ada Boost and Gradient Boost (Advanced) Algorithms.

3) Stacking:- It is also called as Blending. Here multiple machine learning models are used in series. predictions are used as input for next steps.



* ADA BOOS

Data set!-

CGPA	Interactiveness
>=9	Yes
<9	No
>=9	No
<9	No
>=9	Yes
>=9	Yes

Step-1:- Initia
item = 1/

Step-2:- Iterate

ⓐ Decision S
decision st
bootstrap
since thee
the full tra
The first de
instances
as shown

①

CGPA
>=9
<9
>=9
<9
>=9
>=9

If CGPA >=
Predicted t

* ADA BOOST ENSEMBLE LEARNING!-

Data set!-

CGPA	Interactiveness	Practical knowledge	Communication skill	Job Profile
>=9	Yes	Good	Good	Yes
<9	No	Good	Moderate	Yes
>=9	No	Average	Moderate	No
<9	No	Average	Good	No
>=9	Yes	Good	Moderate	Yes
>=9	Yes	Good	Moderate	Yes

Step-1!- Initial weight Assigned to each item = 1/6.

Step-2:- Iterate for each weak classifier.

ⓐ **Decision Stump for CGPA!-** Train the decision stump H_{CGPA} with a random bootstrap sample from the training dataset. Since there are only 6 data instances, we the full training dataset.

The first decision stump classifies the instances based on the CGPA attribute as shown in table.

①

CGPA	Predicted Job offer	Actual Job offer	Weight
>=9	Yes ✓	Yes	1/6
<9	no ✗	Yes	1/6
>=9	Yes ✗	no	1/6
<9	no ✓	no	1/6
>=9	Yes ✓	Yes	1/6
>=9	Yes ✓	Yes	1/6

If CGPA ≥9, the data instance is Predicted to have 'Job offer' as 'Yes' else 'No'.

ⓑ Compute the weighted error:- ε_{CAPA} of H_{CAPA} on current training data set T.

$$\boxed{\varepsilon_i = \sum_{j=1}^{N} H_i(d_j)\, wt(d_j)}$$

where, $wt(d_j)$ = weight for jth data va

$H_i(d_j)$ = 0 for correct prediction, &
1 for wrong prediction

$$\varepsilon_{CAPA} = 2 \times \frac{1}{6} = 0.333$$

ⓒ Compute the weight for each weak classifier:-
The weight for each weak classifier.

$$\alpha_{CAPA} = \frac{1}{2} \frac{\ln(1-\varepsilon_{CAPA})}{\varepsilon_{CAPA}}$$

$$\alpha_{CAPA} = \frac{1}{2} \frac{\ln(1-0.333)}{0.333}$$

$$\boxed{\alpha_{CAPA} = 0.347}$$

ⓓ Calculate the normalizing factor:-
The normalizing factor can be calculated as;-

Z_{CAPA} = wt (correct classified instance) $\times e^{-\alpha_{CAPA}}$ × No. of correct classification × No. of wrong classified instance) × wt (wrong classified classification $\times e^{+\alpha_{CA}}$ No. of wrong

$$Z_{CAPA} = \frac{1}{6} \times 4 \times e^{-0.347} + \frac{1}{6} \times 2 \times e^{+0.34}$$

$$\boxed{Z_{CAPA} = 0.9428}$$

ⓔ Update the weight of all data instances:-

weight of correct instances:-

$$wt(d_j)_{i+1} = \frac{wt(d_j)_{CGPA} \text{ of correct Instance} \times e^{-\alpha_{CGPA}}}{Z_{CGPA}}$$

$$wt(d_j)_{i+1} = \frac{\frac{1}{6} \times e^{-0.347}}{0.9428}$$

$$\boxed{(wt(d_j)_{i+1}) = 0.1249}$$

weight of incorrect instances:-

$$wt(d_j)_{i+1} = \frac{wt(d_j)_{CGPA} \text{ of Incorrect Instance} \times e^{+\alpha_{CGPA}}}{Z_{CGPA}}$$

$$wt(d_j)_{i+1} = \frac{\frac{1}{6} \times e^{+0.347}}{0.9428}$$

$$\boxed{(wt(d_j)_{i+1}) = 0.2501}$$

Updated Table:-

CGPA	Predicted Job offer	Actual Job offer	weight
>=9	yes ✓	yes	0.1249
<9	no ✗	yes	0.2501
>=9	yes ✗	no	0.2501
<9	no ✓	no	0.1249
>=9	yes ✓	yes	0.1249
>=9	yes ✓	yes	0.1249

Similarly repeat step-2 for remaining attributes:-

②

Interactions	Predicted Job offer	Actual Job offer	weight
Yes	Yes ✓	Yes	0.1249
No	No ✗	Yes	0.2501
No	No ✓	No	0.2501
No	No ✓	No	0.1249
Yes	Yes ✓	Yes	0.1249
Yes	Yes ✓	Yes	0.1249

$\epsilon_{Interact} = 1 * 0.2501 = 0.2501$

$$\alpha_{Iteract} = \frac{1}{2} \ln \frac{(1-0.2501)}{0.2501} = 0.5490$$

$$Z_{Interact} = 0.1249 * 4 * e^{-0.549} + 0.2501 * 1 * e^{-0.549} + 0.2501 * 1 \times e^{+0.549}$$

$Z_{Ielact} = 0.866$

$\Rightarrow \frac{0.1249 * e^{-0.549}}{0.866}$ $\Rightarrow \frac{0.2501 * e^{-0.549}}{0.866}$

$\boxed{C} \Rightarrow wt(d_j)_{i+1} = 0.0832, \; 0.1667$

$\boxed{W} \Rightarrow wt(d_j)_{i+1} = 0.5001 \Rightarrow \frac{0.2501 * e^{0.549}}{0.866}$

Updated Tables

Interaction	Predicted Job offer	Actual Job offer	weight
Yes	Yes	Yes	0.0832
No	No	Yes	0.5001
No	No	No	0.1667
No	No	No	0.0832
Yes	Yes	Yes	0.0832
Yes	Yes	Yes	0.0832

Practical Knowledge	Predicted Job offer	Actual Job offer	weight
Good	Yes ✓	Yes	0.0832
Good	Yes ✓	Yes	0.5001
Average	No ✓	No	0.1667
Average	No ✓	No	0.0832
Good	Yes ✓	Yes	0.0832
Good	Yes ✓	Yes	0.0832

~~$\varepsilon_{PK} = 1 \times 0.5001 + 3 \times 0.0832$~~

~~$= 0.7497$~~

All are correctly classified and no need to apply any change in weights..

(A)

Com skill	Predicted Job offer	Actual Job offer	Weight
Good	Yes ✓	Yes	0.0832
Mod	No ×	Yes	0.5001
Mod	No ✓	No	0.1667
Good	Yes ×	No	0.0832
Mod	No ×	Yes	0.0832
Mod	No ×	Yes	0.0832

$$\varepsilon_{CS} = 1 * 0.5001 + 3 * 0.0832$$

$$= 0.7497$$

$$\alpha_{CS} = \frac{1}{2} \ln \frac{(1-0.7497)}{0.7497} = -0.5485$$

$$Z_{CS} = 0.0832 * 1 * e^{-(-0.5485)} + 0.1667 * 1 * e^{-(-0.5485)} + 0.5001 * 1 * e^{+(-0.5485)} + 0.0832 * 3 * e^{+(-0.5485)} = 0.866$$

$\boxed{C} \Rightarrow wt(ds)_{i+1} = \frac{0.0832 * e^{(-0.5485)}}{0.866} = 0.1663$

$wt(ds)_{i+1} = \frac{0.1667 * e^{-(-0.5485)}}{0.866} = 0.331$

$\boxed{W} \Rightarrow wt(d_j)_{i+1} = \dfrac{0.5001 \times e^{+(-0.5485)}}{0.866}$

$= 0.3337$

$wt(d_j)_{i+1} = \dfrac{0.0832 + e^{+(-0.5485)}}{0.866}$

$= 0.0555$

Updated Table:

Commskill	Predicted Job offer	Actual Job offer	weight
good	yes	yes	0.0832
mod	no	yes	0.3337
mod	no	no	0.3331
good	yes	no	0.0555
mod	no	yes	0.0555
mod	no	yes	0.0555

Step-3 :- Compute the final predicted value for each data instance :-

SI. No	$\alpha_{CGPA} = 0.347$	$\alpha_I = 0.549$	$\alpha_{CS} = -0.5485$	weighted Avg	final Prediction
1	Yes	Yes	Yes	0.3475	Yes
2	No	No	No	0	No
3	Yes	No	No	0.347	Yes
4	No	No	Yes	-0.5485	No
5	Yes	Yes	No	0.896	Yes
6	Yes	Yes	No	0.896	Yes

WA > 0 ⇒ final prediction ⇒ yes

$$H_F(d_j) = \sum_{i=1}^{m} \alpha_i * H_i(d_j)$$

$$H_F(d_j) = \alpha_{CAPA} * Yes + \alpha_{Interact} * Yes + \alpha_{CSI} * Yes = 0.347 * 1 + 0.5490 * 1 - 0.5485 * 1 = 0.3475$$

* Naive Bayes classifier to classify examp. into one of the given instances!

Data Set:- (Continuous) (Discrete)

day	outlook	Temperature	Humidity	wind	PlayTennis
D_1	Sunny	Hot	High	Weak	No
D_2	Sunny	Hot	High	Strong	No
D_3	Overcast	Hot	High	Weak	Yes
D_4	Rain	Mild	High	Weak	Yes
D_5	Rain	Cool	Normal	Weak	Yes
D_6	Rain	Cool	Normal	Strong	No
D_7	Overcast	Cool	Normal	Strong	Yes
D_8	Sunny	Mild	High	Weak	No
D_9	Sunny	Cool	Normal	Weak	Yes
D_{10}	Rain	Mild	Normal	Weak	Yes
D_{11}	Sunny	Mild	Normal	Strong	Yes
D_{12}	Overcast	Mild	High	Strong	Yes
D_{13}	Overcast	Hot	Normal	Weak	Yes
D_{14}	Rain	Mild	High	Strong	No

New instance have to classify yes or no

{Sunny, Cool, high, strong} = ?

* Prior probabilities

P(PlayTennis = yes

P(PlayTennis = no)

* Conditional probab

attributes!

outlook	Yes	
sunny	$\frac{2}{9}$	3
overcast	4/9	
Rain	3/9	2

Temperature	Yes
Hot	$\frac{2}{9}$
Mild	$\frac{4}{9}$
Cool	$\frac{3}{9}$

* classify new e

Naive Bayes Class

$v_{NB} = \underset{v_j \in (Yes, N}{\text{argmax}}$

$= \underset{v_j \in (yes, no)}{\text{argmax}} P(v_j)$

~~= argmax P(v~~

~~$v_j \in (yes, no)$~~

$v_{NB}(yes) = P(y$

$P(l$

$= 0.64 \times \frac{2}{9}$

*Prior probabilities of yes and not

$P(\text{PlayTennis} = \text{yes}) = \frac{9}{14} = 0.64$

$P(\text{PlayTennis} = \text{no}) = \frac{5}{14} = 0.36$

*Conditional probabilities of individual attributes:

outlook	Yes	No
sunny	$\frac{2}{9}$	3/5
overcast	4/9	0
Rain	3/9	2/5

Humidity	Yes	No
High	3/9	4/5
Normal	6/9	1/5

Temperature	Yes	No
Hot	$\frac{2}{9}$	2/5
Mild	$\frac{4}{9}$	2/5
Cool	$\frac{3}{9}$	1/5

Windy	Yes	No
Strong	3/9	3/5
weak	6/9	2/5

*classify new instance either yes or not

Naive Bayes Classifier formula:

$$v_{NB} = \underset{v_j \in (\text{Yes}, \text{No})}{\operatorname{argmax}} P(v_j) \prod_i P(a_i / v_j)$$

$$= \underset{v_j \in (\text{yes}, \text{no})}{\operatorname{argmax}} P(v_j) \; P(\text{outlook} = \text{sunny} / v_j) \times$$

~~argmax P(vj)~~
~~vj ∈ (yes, no)~~

$$P(\text{Temperature} = \text{cool} / v_j) \times$$
$$P(\text{Humidity} = \text{high} / v_j) \times$$
$$P(\text{wind} = \text{strong} / v_j).$$

$$v_{NB}(\text{yes}) = P(\text{yes})\, P(\text{Sunny}/\text{yes})\, P(\text{cool}/\text{yes})\, P(\text{high}/\text{yes})\, P(\text{strong}/\text{yes})$$

$$= 0.64 \times \frac{2}{9} \times \frac{3}{9} \times \frac{3}{9} \times \frac{1}{9} = 0.00526$$

$$v_{NB}(no) = P(no)\,P(sunny/no)\,P(cool/no)\,P(high/no)\,P(strong/no)$$

$$= 0.36 \times \frac{3}{5} \times \frac{1}{5} \times \frac{4}{5} \times \frac{3}{5}$$

$$= 0.0206$$

$$v_{NB}(yes) = \frac{v_{NB}(yes)}{v_{NB}(yes) + v_{NB}(no)}$$

$$= \frac{0.00526}{0.00526 + 0.0206}$$

$$= 0.205$$

$$v_{NB}(no) = \frac{v_{NB}(no)}{v_{NB}(yes) + v_{NB}(no)}$$

$$= \frac{0.0206}{0.00526 + 0.0206}$$

$$= 0.795$$

$$\boxed{\therefore v_{NB}(no) > v_{NB}(yes)}$$

Hence {sunny, cool, high, strong} = no

Q) Consider the
rival teams
Suppose team
time & team
matches. Amo
A only 35%
playing at
on the othe
victories of
while playi

a) If team B
between the
probability
the winner

b) If team B
match betw
will emerge

Sol: X → Winni
Y → Hosti

$P(Y_A) = 0.$

$P(Y_B) = 0.$

Probability t
match et
Probability
match won
0.35

Naive Bayes Classification ~~Discrete~~ (Continuous) :-

Data :-

Person	Height	weight	foot size
Male	6.00	180	12
Male	5.92	190	11
Male	5.58	170	12
Male	5.92	165	10
female	5.00	100	6
female	5.50	150	8
female	5.42	130	7
female	5.75	150	9

Based on the following data determine the gender of a person having height 6ft; weight 130 lbs; and foot size 8 inch.

Sol: $P(\text{Male}) = \frac{4}{8} = 0.5$

$P(\text{female}) = \frac{4}{8} = 0.5$

$\text{Mean}(\text{Height})_{\text{male}} = 5.855 \Rightarrow \frac{\sum x_i}{N} (\bar{x})$

$\text{Variance}(\text{Height})_{\text{male}} = 0.035055 \Rightarrow \frac{\sum (x_i - \bar{x})^2}{n-1}$

$\text{Mean}(\text{Weight})_{\text{male}} = 176.25$

$\text{Variance}(\text{weight})_{\text{male}} = 122.92$

$\text{Mean}(\text{Foot size})_{\text{male}} = 11.25$

$\text{Variance}(\text{foot size})_{\text{male}} = 0.91667$

Attribute	Male	female
Mean(height)	5.855	5.4175
Variance(height)	0.035033	0.097225
Mean(weight)	176.25	132.5
Variance (weight)	122.92	055833
Mean(footsize)	11.25	7.5
Variance(footsize)	0.91667	1.6667

New Instance :-

sex	Height(ft)	Weight(lbs)	footsize(inch)
sample	6	180	8

$$\text{Posterior(Male)} = \frac{P(M) * P(H/M) * P(W/M) * P(FS/M)}{\text{Evidence}}$$

$$\text{Posterior(female)} = \frac{P(F) * P(H/F) * P(W/F) * P(FS/F)}{\text{Evidence}}$$

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

$$P(H/M) = \frac{1}{\sqrt{2 \times 3.142 * 0.635033}} e^{\frac{-(6-5.855)^2}{2 * 0.035033}}$$

$$= 1.5789$$

$P(W/M) = 5.9881e^{-6}$

$P(Fs/M) = 1.3112e^{-3}$

$P(H/F) = 2.2346e^{-1}$

$P(W/F) = 1.6789e^{-2}$

$P(Fs/F) = 2.8669e^{-1}$

Posterior(Male) = $6.1984e^{-9}$

Posterior (female) = $5.377e^{-4}$

$\boxed{\therefore \text{Posterior(female)} > \text{Posterior (Male)}}$

Hence the sex of new enstance es "female".

KNN Classifier Example :-

①

Rank	Sepal length	Sepal width	Species	Dest
3	5.3	3.7	Setosa	0.60
6	5.1	3.8	Setosa	0.70
13	7.2	3.0	Virginica	2.00
2	5.4	3.4	Setosa	0.3
1	5.1	3.3	Setosa	0.22
8	5.4	3.9	Setosa	0.82
15	7.4	2.8	Virginica	2.22
10	6.1	2.8	Versicolor	0.9
14	7.3	2.9	Virginica	2.1
9	6.0	2.7	Versicolor	0.8
5	5.8	2.8	Virginica	0.67
12	6.3	2.3	Versicolor	1.36
4	5.1	2.5	Versicolor	0.60
11	6.3	2.5	Versicolor	1.25
7	5.5	2.4	Versicolor	0.75

New classifier Species = ?

Sepal length	Sepal width	Species
5.2	3.1	?

Ans :-

Step-1 :- Find distance for each attribute to new classifier.

$$\text{Destance} = \sqrt{(x-a)^2 + (y-b)^2}$$

Distance (Sepal length, Sepal width)

$$= \sqrt{(5.2 - 5.3)^2 + (3.1 - 3.7)^2}$$

$$= 0.608$$

Distan
0.608
0.707
2.002
0.36
0.22
0.82
2.22
0.94
2.1
0.89
0.67
1.36
0.60
1.25
0.75

ribute

Step-2:- find rank of each example with respect to new example:

The distance having minimum value it is given as least rank (1). (sort in ascending older.

Step-3:- find the nearest neighbour. Given value of k we need to identify the first ranked example.

If K=1 -> setosa
If K=2 -> setosa
If K=3 -> setosa
If K=5 -> setosa

Hence the species of new classifier is "setosa".

② find KNN for the new instance

Mind => 1

Height	Weight	Class	(d) Distance	Rank
167	51	Underweight	6.7	5
182	62	normal	13	8
176	69	normal	13.4	9
173	64	normal	7.6	6
172	65	normal	8.2	7
174	56	Underweight	4.1	4
169	58	normal	1.4	1
173	57	normal	3	3
170	55	normal	2	2

New classifier class => ?

Height	weight	class
170	57	?

If $K = 1$, Normal
If $K = 2$, Normal
If $K = 3$, Normal
If $K = 5$, Normal

Hence the class of the new classifier is "Normal".

③ compute 3 NN for test instance $t1 = (3, 7)$ using Euclidean distance and corresponding weights.

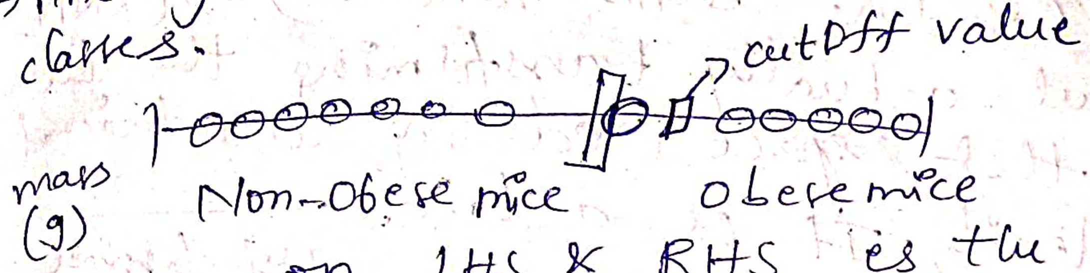
$1/d^2$ Max vote ⇒ 1

Training Instance	x_1	x_2	output	Distance (d)	d^2	vote	Rank
I_1	7	7	0	4	16	0.06	3
I_2	7	4	0	5	25	0.04	4
I_3	3	4	1	3	9	0.11	1
I_4	1	4	1	3.6	12.96	0.08	2

If $K = 3$, Output is "1".

Hence the output of the new classifier is "1".

Support Vector Machines (SVM):- It is an extension of LDA
We have to put limits to seperate data points into different classes.
→ Non-linear seperation of data points.
→ finding cutofff value to seperate into classes.



→ Margin on LHS & RHS is the distance between cutoff value to data point.
→ Midpoint of 2 data points, calculate distance & take midpoint.
→ The shortest distance between the observations & the threshold is called "margin".
→ Still allow mis classification and get high accuracy on test data.
→ move margin to left ⇒ margin to left is reduced.
→ move margin to right ⇒ margin to right is reduced.
→ Sometimes the margin is near to one class.
→ **Soft Margins**:- SVM allows misclassifications.
→ The distance between the observation and threshold is called "soft margin".

Multidimensional Search space:
in 3D we have plane to seperate
in 2D we have line to seperate.

① Find the hyperplane with maximum margin for the data using SVM algori

x_1	x_2	Class
2	2	-1
4	5	+1
7	4	+1

Ⓐ $n=3$

$\vec{x_1} = (2, 2)$ $y_1 = -1$

$\vec{x_2} = (4, 5)$ $y_2 = +1$

$\vec{x_3} = (7, 4)$ $y_3 = +1$

The hyperplane equation in SVM es

$$f(\vec{x}) = \vec{w} \cdot \vec{x} - b \qquad \left(\vec{w} = \sum_{i=1}^{N} \alpha_i y_i \vec{x_i}\right)$$

Take set of variables based on inpu vectors $\vec{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$

Subject to the conditions

$$\sum_{i=1}^{N} \alpha_i y_i \Rightarrow -\alpha_1 + \alpha_2 + \alpha_3 = 0 \quad \text{———} \quad ①$$

$\alpha > 0$ this is the condition we need to consider.

$$\phi(\vec{\alpha}) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1, j=1}^{N} \alpha_i \alpha_j y_i y_j (\vec{x_i} \cdot \vec{x_j})$$

$$= \sum_{i=1}^{3} \alpha_i - \frac{1}{2} \sum_{i=1, j=1}^{3} \alpha_i \alpha_j y_i y_j (\vec{x_i} \cdot \vec{x_j})$$

$(\vec{x_1} \cdot \vec{x_1}) = 8$, $(\vec{x_1} \cdot \vec{x_2}) = 18$, $(\vec{x_1} \cdot \vec{x_3}) = 22$

$(\vec{x_2} \cdot \vec{x_1}) = 18$, $(\vec{x_2} \cdot \vec{x_2}) = 41$, $(\vec{x_2} \cdot \vec{x_3}) = 48$

$(\vec{x_3} \cdot \vec{x_1}) = 22$, $(\vec{x_3} \cdot \vec{x_2}) = 48$, $(\vec{x_3} \cdot \vec{x_3}) = 65$

$$= (\alpha_1 + \alpha_2 + \alpha_3) - \frac{1}{2}\left(8\alpha_1^2 + 41\alpha_2^2 + 65\alpha_3^2 + 96\alpha_2\alpha_3 - 36\alpha_1\alpha_2 - 44\alpha_1\alpha_3\right)$$

from 1 we get $\alpha_1 = \alpha_2 + \alpha_3$ put it on above equation.

$$\phi(\vec{\alpha}) = 2(\alpha_2 + \alpha_3) - \frac{1}{2}\left(13\alpha_2^2 + 32\alpha_2\alpha_3 + 29\alpha_3^2\right)$$

for $\phi(\vec{\alpha})$ to be maximum we must have $\frac{\partial \phi}{\partial \alpha_2} = 0$, $\frac{\partial \phi}{\partial \alpha_3} = 0$

That is,

$$2 - 13\alpha_2 - 16\alpha_3 = 0 \longrightarrow ②$$

$$2 - 16\alpha_2 - 29\alpha_3 = 0 \longrightarrow ③$$

Solving 2 and 3 we get

$$\alpha_2 = \frac{26}{121}, \quad \alpha_3 = \frac{-6}{121}, \quad \alpha_1 = \frac{20}{121} \left(\because \alpha_1 = \alpha_2 + \alpha_3\right)$$

$$\vec{w} = \sum_{i=1}^{N} (\alpha_i y_i \vec{x_i})$$

$$= \frac{20}{121}(-1)(2, 2) + \frac{26}{121}(1)(4, 5) + \frac{-6}{121}(1)(7, 4)$$

$$= \left(-\frac{40}{121} + \frac{104}{121} - \frac{42}{121}, \frac{-40}{121} + \frac{130}{121} - \frac{24}{121}\right)$$

$$= \left(\frac{22}{121}, \frac{66}{121}\right)$$

$$= \left(\frac{2}{11}, \frac{6}{11}\right)$$

$$b = \frac{1}{2}\left(\min_{i:y_i=+1}(\vec{w}\cdot\vec{x_i}) + \max_{i:y_i=-1}(\vec{w}\cdot\vec{x_i})\right)$$

$$= \frac{1}{2}\left(\min\left((4,5)\cdot\left(\frac{2}{11}, \frac{6}{11}\right), (7,4)\cdot\left(\frac{2}{11}, \frac{6}{11}\right)\right) \max\left((2,2)\cdot\left(\frac{2}{11}, \frac{6}{11}\right)\right)\right)$$

$$= \frac{1}{2}\left(\min\left(\frac{8}{11} + \frac{30}{11}, \frac{14}{11} + \frac{24}{11}\right) + \frac{4}{11} + \frac{12}{11}\right)$$

$$= \frac{1}{2}\left(\frac{38}{11} + \frac{16}{11}\right)$$

$$= \frac{1}{2}\left(\frac{54}{11}\right)$$

$$= \frac{27}{11}$$

SVM classifier is given by

$$f(\vec{x}) = \vec{w}\cdot\vec{x} - b$$

~~$f\left(\left(\frac{2}{11}, \frac{6}{11}\right)\cdot(2,2)\right) - \frac{27}{11}$~~

where $\vec{x} = (x_1, x_2)$

$$= \left(\frac{2}{11}, \frac{6}{11}\right)\cdot(x_1, x_2) - \frac{27}{11}$$

$$= \frac{2x_1}{11} + \frac{6x_2}{11} - \frac{27}{11}$$

$$\therefore \frac{2x_1}{11} + \frac{6x_2}{11} - \frac{27}{11} = 0$$

$$\Rightarrow 2x_1 + 6x_2 - 27 = 0$$

$$\Rightarrow \boxed{x_1 + 3x_2 - \frac{27}{2} = 0}$$

This is the hyperplane with maximum margin for the given data.

1Q)

Data Point	Class label
(1, 2)	A
(2, 3)	A
(3, 3)	B
(6, 5)	B

P(4, 4) using k = 3 for KNN algorithm using euclidean distan

Sol:-

$d_1 = \sqrt{(4-1)^2 + (4-2)^2} = \sqrt{9+4} = \sqrt{13} = 3.60$

$d_2 = \sqrt{(4-2)^2 + (4-3)^2} = \sqrt{5} = 2.236$

$d_3 = \sqrt{(4-3)^2 + (4-3)^2} = \sqrt{2} = 1.414$

$d_4 = \sqrt{(6-4)^2 + (5-4)^2} = \sqrt{5} = 2.236$

Based on d we give rank on ascending order of distance(d).

Data point	Class label	Destance	Rank
(1, 2)	A	3.605	4
(2, 3)	A	2.236	2
(3, 3)	B	1.414	1
(6, 5)	B	2.236	3

Here K = 3, so majority is B. Therefo the class of (4, 4) is "B".

28)

Patient	Smoking (x)	Heart disease (y) (Target Variable)
1	1	1
2	1	0
3	0	0
4	1	1
5	0	0
6	0	1

Predict the probability of heart disease for any patient based on whether they smoke $(X=1)$ or don't smoke. Estimated parameters for logistic regression model:

$\beta_0 = -0.693$, $\beta_1 = 1.386$

(A) $\log(\text{odds}) = z = -0.693 + 1.386x$

$S(x) = \frac{1}{1+e^{-x}}$ $\qquad x = 0.5$

$P = \frac{1}{1+e^{-z}}$

$z = -0.693 + 1.386(X)$

$X = 1$ if the patient smokes and
$X = 0$ if the patient doesn't smokes.

$z = -0.693 + 1.386 = 0.693$ (if $X=1$)
$z = -0.693$ if $(X=0)$

$P = \frac{1}{1+e^{-0.693}} = \frac{1}{1.5} = 0.666$

The probability of heart disease who smoke is 0.66.

$P = \frac{1}{1+e^{0.693}} = \frac{1}{2.997} = 0.333$

Therefore the probability of heart disease for a patient who doesn't smoke is 0.33

* Decision Tree using GINI INDEX!

Data :-

Weekend	Weather	Parents	Money	Decision
W_1	Sunny	Yes	Rich	Cinema
W_2	Sunny	No	Rich	Tennis
W_3	Windy	Yes	Rich	Cinema
W_4	Rainy	Yes	Poor	Cinema
W_5	Rainy	No	Rich	Stay In
W_6	Rainy	Yes	Poor	Cinema
W_7	Windy	No	Poor	Cinema
W_8	Windy	No	Rich	Shopping
W_9	Windy	Yes	Rich	Cinema
W_{10}	Sunny	NO	Rich	Tennis

Solution:

The target variable has four possible outc
Gini index of target variable is:-

$$\text{Gini}(S) = 1 - \left(\left(\frac{6}{10}\right)^2 + \left(\frac{2}{10}\right)^2 + \left(\frac{1}{10}\right)^2 + \left(\frac{1}{10}\right)^2\right)$$

$$= 1 - \frac{36+4+1+1}{100}$$

$$= 1 - \frac{42}{100}$$

$$= 0.58$$

4 Cinema,
2 Tennis,
1 stay in,
1 shopping.

① Gini index for money:-

Money = poor, 3 examples in cinema

7 → Rich
3 → poor

~~Gini(S)~~

$$\text{Gini}(S) = 1 - \left(\frac{3}{3}\right)^2 = 0$$

Money = Rich, 2 tennis, 3 cinema, 1 shopp
1 stay in

$$\text{Gini}(S) = 1 - \left[\left(\frac{2}{7}\right)^2 + \left(\frac{3}{7}\right)^2 + \left(\frac{1}{7}\right)^2 + \left(\frac{1}{7}\right)^2\right]$$

$$= 1 - \frac{4+9+1+1}{49} = \frac{49-15}{49} = \frac{34}{49}$$

Weighted Average (Money)

$= 0 \times \left(\frac{3}{10}\right) + 0.694 \times \left(\frac{7}{10}\right) = 0.486$

② Gini Index for parents:

parents = Yes, 5 Yes ⇒ 5 Cinema

5→Yes
5→No

$Gini(S) = 1 - \left(\frac{5}{5}\right)^2 = 0$

parents = No, 2 Tennis, 1 Cinema, 1 shopping, 1 stay in

$$Gini(S) = 1 - \left(\left(\frac{2}{5}\right)^2 + \left(\frac{1}{5}\right)^2 + \left(\frac{1}{5}\right)^2 + \left(\frac{1}{5}\right)^2\right)$$

$$= 1 - \left(\frac{4+1+1+1}{25}\right) = 1 - \frac{7}{25} = \frac{18}{25} = 0.72$$

Weighted Average (Parents)

$= 0\left(\frac{5}{10}\right) + 0.72\left(\frac{5}{10}\right) = 0.36$

③ Gini Index for weather:

Weather = Sunny, 2 Tennis, 1 Cinema

Sunny → 3
Rainy → 3
Windy → 4

$$Gini(S) = 1 - \left(\left(\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2\right)$$

$$= 1 - \left(\frac{4+1}{9}\right) = \frac{4}{9} = 0.444$$

Weather = Windy, 3 cinema, 1 shopping

$$Gini(S) = 1 - \left(\left(\frac{3}{4}\right)^2 + \left(\frac{1}{4}\right)^2\right)$$

$$= 1 - \left(\frac{9+1}{16}\right) = \frac{6}{16} = 0.375$$

Weather = Rainy, 2 Cinema, 1 stay in

$$Gini(S) = 1 - \left(\left(\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2\right)$$

$$= 1 - \left(\frac{4+1}{9}\right) = \frac{4}{9} = 0.444$$

Weighted Average (Weather)

$= 0.444\left(\frac{3}{10}\right) + 0.375\left(\frac{4}{10}\right) + 0.444\left(\frac{3}{10}\right)$

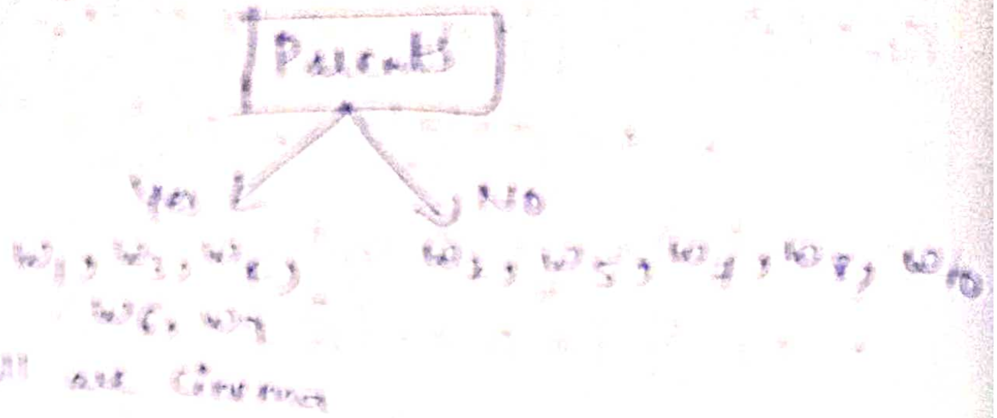
$= 0.416$

for weather ⇒ Gini Index ⇒ 0·416

for parents ⇒ Gini Index ⇒ 0·36

for Money ⇒ Gini Index ⇒ 0·486

Parent is selected as it has smalle Gini Index.



① Gini Index of parents ⇒ No | Weather Attrib

~~Sunny ② tennis, 1 Stay in, 1 Cinema~~, Shopping

$Gini(S) = 1 - \left(\frac{2}{2}\right)^2 = 0$ (2 sunny, no, tennis)

Rainy ①

$Gini(S) = 1 - \left(\frac{1}{1}\right)^2 = 0$

Windy ②

$Gini(S) = 1 - \left(\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2\right) = 0.5$

Weighted Average ⇒ $0.5\left(\frac{2}{5}\right) = 0.2$

② Gini Index of Money ⇒ NO | Money

Rich ④

$Gini(S) = 1 - \left(\left(\frac{2}{4}\right)^2 + \left(\frac{1}{4}\right)^2 + \left(\frac{1}{4}\right)^2\right)$

$= 1 - \frac{6}{16} = \frac{10}{16} = 0.625$

Poor 1

$Gini(S) = 1 - \left(\frac{1}{1}\right)^2 = 0$

Weighted Average ⇒ $0.625\left(\frac{4}{5}\right) = 0.5$

for parents = No|Weather → Gini Index => 0.2

for parents = No|Money → Gini Index => 0.5

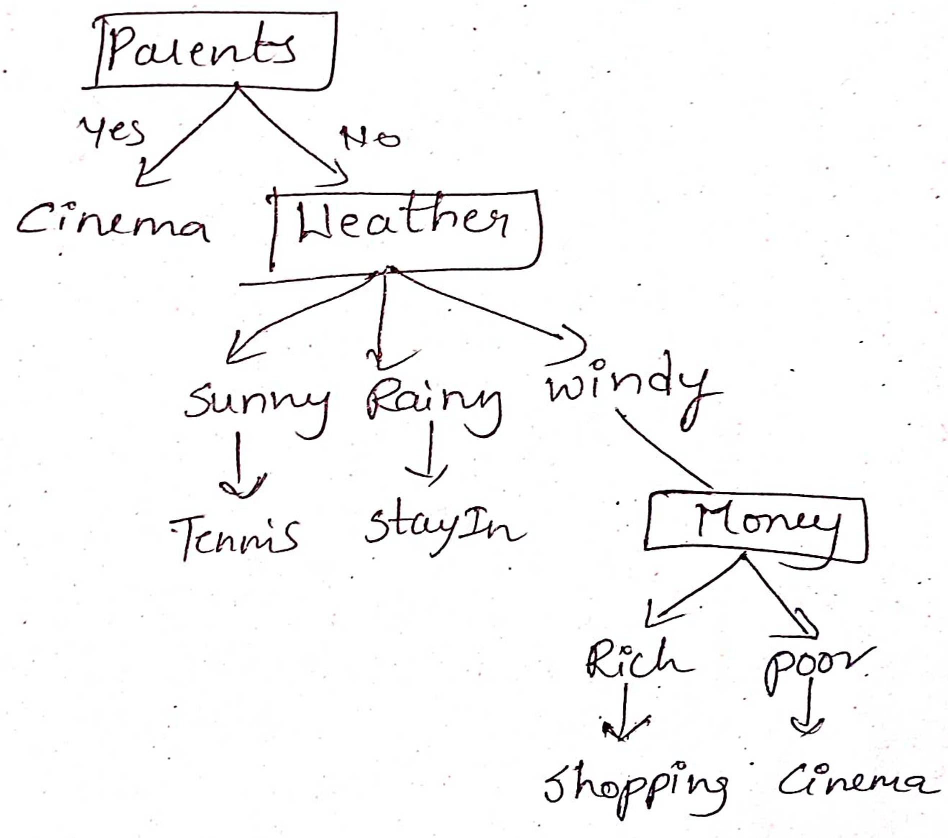
Weather is selected as it has smallest Gini Index

Now, for parents = No & weather = sunny, we have all enstances as Tennis.

other for parents = No & weather = Rainy, we have all enstances as stay In.

another for parents = No & weather = windy, we have ~~all~~ one enstances as ~~spti~~ cinema and other as shopping.

Decision Tree:



Formulas

1) Logistic Regression:

$$Z = \log(\text{odds}) = b_0 + b_1(x)$$

$$P = \frac{1}{1+e^{-Z}}$$

2) Linear Discriminant Analysis (LDA):

1) calculate mean of each class. μ_1, μ_2
2) covariance matrix of each class.

$$S_i = \Sigma (x-\mu_1)(x-\mu_1)^T / N-1$$

3) calculate within scatter matrix $S_W = S_1 + S_2$
~~4)~~ calculate $S_B = \Sigma(\mu_1 - \mu_2)(\mu_1 - \mu_2)^T$
~~5)~~ find eigen values for $S_W^{-1} S_B W = \lambda W$

$$|S_W^{-1} S_B - \lambda I| = 0$$

6) find $\boxed{w^* = S_W^{-1}(\mu_1 - \mu_2)}$
7) $\alpha_1 x_1 + \alpha_2 x_2$ put (x_1, x_2) is given in question. (t)
8) find $\alpha_1 x_1 + \alpha_2 x_2$ of μ_1 and μ_2, find mid
if midpoint $> t$ then it is class A, el
class B.

3) KN Neighbour (KNN):

1) find distance between new instance clas
to all examples in data set using
Euclidean Distance $= \sqrt{(x_1-x_2)^2 + (y_1-y_2)^2}$.
2) compute rank based on distance using
ascending order.
3) Based on given K the majority clas
is assigned to new instance.

4) Decision Tree (ID3 Algorithm):

$$E(S) \rightarrow \text{Entropy} = -\sum_{i=1}^{n} P_i * \log_2(P_i)$$

$$\text{Binomial Classification}(S) = -P_+ \log_2(P_{+i}) - P_- \log_2(P$$

$$\text{Information Gain} = E(S) - \Sigma \frac{|S_v|}{|S|} \text{Entropy}(S_v)$$

5) **ADA Boost Algorithm (Ensemble Learning Technique):-**

1) Initial weight is assigned to each ~~weak~~ item

2) Iterate for each weak classifier (attribute).

ⓐ Weighted Error $\varepsilon_i = \sum_{j=1}^{N} H_i(d_j)\, wt(d_j)$

$H_i(d_j) \Rightarrow 0$ for correct prediction
$\quad\quad\quad\;\; 1$ for Wrong prediction

$wt(d_j) \Rightarrow$ weight for j^{th} data value.

ⓑ Weight of each weak classifier

$$\alpha_{attribute} = \frac{1}{2}\,\frac{\ln(1-\varepsilon_i)}{\varepsilon_i}$$

ⓒ Normalizing factor $Z_i = wt(CCI) * n(CCI) * e^{-\alpha_i} + wt(WCI) * n(CCI) * e^{\alpha_i}$

ⓓ Weight of all data instances (updated Weights)

$$wt(d_j)_{i+1} = \frac{wt(d_j)_i\; CCI * e^{-\alpha_i}}{Z_i}$$

$$wt(d_j)_{i+1} = \frac{wt(d_j)_i\; WCI * e^{\alpha_i}}{Z_i}$$

3) Calculate the final predicted value for each data instance:-

$$H_f(d_j) = \sum_{i=1}^{m} \alpha_i * H_i(d_j)$$

6) **Naive Bayes:-** (For descrete data)

1) Probability of yes and no.
2) Conditional probability of each seperate instance in each attributes.
3) Conditional probability of each seperate instance for new instance.
4) if $P(yes) > P(no)$ then it is no else yes.

(For Continuous data)

1) Mean and variance for each attribute with respect to seperation attribute.
2) probability of yes and no.
3) Probability of posterior values for new instance.
4) $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

7) **Support Vector Machines (SVM):-**

1) $\vec{\alpha} = (\alpha_1, \alpha_2, \alpha_3 \cdots)$
2) $\sum_{i=1}^{N} \alpha_i y_i \Rightarrow 0$
3) $\phi(\vec{\alpha}) = \sum^{N} \alpha_i - \frac{1}{2} \sum^{N} \alpha_i \alpha_j y_i y_j (\vec{x_i} \cdot \vec{x_j})$
4) $\vec{w} = \sum_{i=1}^{N} \alpha_i y_i \vec{x_i}$
5) $b = \frac{1}{2}(\min(\vec{w_i}\,\vec{x_i}), \max(\vec{w_i}\,\vec{x_i}))$
6) $f(\vec{x}) = \vec{w}\,\vec{x} - b$