12/03/19

1. find a person can play or not in weather conditions; outlook = rainy, temperature = mild, windy = weak.

Sol: Probability that can play a game.

$P(\text{Outlook} = \text{rainy} \mid \text{play} = \text{yes}) = 3/9$.

$P(\text{temp} = \text{mild} \mid \text{play} = \text{yes}) = 4/9$

$P(\text{humidity} = \text{nominal} \mid \text{play} = \text{yes}) = 6/9$

$P(\text{windy} = \text{weak} \mid \text{play} = \text{yes}) = 6/9$

Probability that can't play a game.

$P(\text{outlook} = \text{rainy} \mid \text{play} = \text{NO}) = 2/5$

$P(\text{temp} = \text{mild} \mid \text{play} = \text{no}) = 2/5$

$P(\text{humidity} = \text{nominal} \mid \text{play} = \text{no}) = 1/5$

$P(\text{windy} = \text{weak} \mid \text{play} = \text{no}) = 2/5$

$P(X \mid \text{play} = \text{yes}) \times P(\text{play} = \text{yes})$

$= \left(\frac{3}{9} * \frac{4}{9} * \frac{6}{9} * \frac{6}{9}\right) \times \frac{9}{14} = 0.0423$

0.0423
0.0045
0.0468

$P(x \mid \text{play} = \text{NO}) \times P(\text{play} = \text{NO})$

$= \frac{2}{5} \times \frac{2}{5} \times \frac{1}{5} \times \frac{2}{5} \times \frac{5}{14} = 0.0045$

$P(A|C) = \frac{P(C|A)P(A)}{P(C)}$ to get $P_C = 0.0468$

$p(play = yes|X) = 0.0423 / 0.0468 = \cancel{0.025}\ 0.903$

$p(play = NO|X) = 0.0045 / 0.0468 = 0.0961$

2. for a given data set patient details with attribute, cold, running nose, headache, fever, flu find whether the person is having flu or not for the given symptoms cold = yes, runnynose = no, headache = mild, fever = yes.

	cold	Runnynose	headache	fever	Classlabel flu
1	yes	NO	mild	yes	No
2	yes	Yes	NO	NO	yes
3	yes	NO	Strong	yes	yes
4	No	yes	mild	Yes	yes
5	No	NO	no	NO	NO
6	NO	yes	Strong	yes	yes
7	NO	yes	Strong	NO	No
8	yes	yes	mild	yes	yes

Cold

	yes	NO	P(yes)	P(NO)
~~yes~~	3	[illegible]	3/5	[illegible]/3
~~NO~~				

P(Yes) = 5/8

P(NO) = 3/8

Runnynose

	yes	NO	f(yes)	f(NO)
no	[illegible]	2	[illegible]/5	2/3

Compare with classlabels.

headache

	yes	NO	f(yes)	f(NO)
mild	2	1	2/5	1/3
~~strong~~				

2/03/19

1. find a person can play or not in weather conditions; outlook = rainy, temperature = mild, windy = weak. Probability that can play a game.

Sol

$P(\text{Outlook} = \text{rainy} \mid \text{play} = \text{yes}) = 3/9.$

$P(\text{temp} = \text{mild} \mid \text{play} = \text{yes}) = 4/9$

$P(\text{humidity} = \text{nominal} \mid \text{play} = \text{yes}) = 6/9$

$P(\text{windy} = \text{weak} \mid \text{play} = \text{yes}) = 6/9.$

Probability that can't play a game.

$P(\text{outlook} = \text{rainy} \mid \text{play} = \text{NO}) = 2/5$

$P(\text{temp} = \text{mild} \mid \text{play} = \text{no}) = 2/5$

$P(\text{humidity} = \text{nominal} \mid \text{play} = \text{no}) = 1/5$

$P(\text{windy} = \text{weak} \mid \text{play} = \text{no}) = 2/5$

$P(x \mid \text{play} = \text{yes}) \times P(\text{play} = \text{yes})$

$= \left(\frac{3}{9} * \frac{4}{9} * \frac{6}{9} * \frac{6}{9}\right) \times \frac{9}{14} = 0.0423$

$$\begin{array}{r} 0.0423 \\ 0.0041 \\ \hline 0.0465 \end{array}$$

$P(x \mid \text{play} = \text{NO}) \times P(\text{play} = \text{NO})$

$= \frac{2}{5} \times \frac{2}{5} \times \frac{1}{5} \times \frac{2}{5} \times \frac{5}{14} = 0.0045$

$P(A \mid C) = \frac{P(C \mid A) P(A)}{P(C)}$ to get $P_C = 0.0468$

$P(\text{Play}=\text{yes}|X) = 0.0423 / 0.0468 = ~~0.025~~ \; 0.9013$

$P(\text{play}=\text{NO}|X) = 0.0045 / 0.0468 = 0.0961$

2. for a given dataset patient details with 8 attribute, cold, running nose, headache, fever, flu find whether the person is having flu or not for the given symptoms cold = yes, runnynose = no, headache = mild, fever = yes.

	cold	Runnynose	headache	fever	Classlabel flu
1	yes	NO	mild	yes	No
2	yes	Yes	NO	NO	yes
3	yes	NO	Strong	yes	yes
4	No	yes	mild	yes	yes
5	No	NO	no	NO	NO
6	NO	yes	strong	yes	yes
7	NO	yes	strong	NO	No
8	yes	yes	mild	yes	yes

Cold

	yes	NO	P(yes)	P(NO)
~~NO~~ yes	3	1	3/5	1/3
NO				

P(Yes) = 5/8

P(NO) = 3/8

Runnynose

	yes	NO	P(yes)	P(NO)
no	1	2	1/5	2/3

Compare with classlabels.

headache

	yes	NO	P(yes)	P(NO)
mild	2	1	2/5	1/3
~~Strong~~				

Scanned by CamScanner

fever	yes	NO	f(yes)	f(NO)
yes	4	1	4/5	1/3

Probability of having flu.

P(cold = yes | flu = yes) = 3/5

P(runny nose = no | flu = yes) = 1/5

P(headache = mild | flu = yes) = 2/5

P(fever = yes | flu = yes) = 4/5.

Probability of not having flu

P(cold = yes | flu = NO) = 1/3

P(runny nose = no | flu = ~~yes~~ NO) = 2/3

P(headache = mild | flu = ~~yes~~ NO) = 1/3

P(fever = yes | flu = ~~yes~~ NO) = 1/3.

P(x | play = yes) x (play = yes)

$= \frac{3}{5} \times \frac{1}{5} \times \frac{2}{5} \times \frac{4}{5} \times \frac{5}{8} = 0.024$

P(x | play = no) x (play = no)

$= \frac{1}{3} \times \frac{2}{3} \times \frac{1}{3} \times \frac{1}{3} \times \frac{3}{8} = 0.009$.

P(C) = 0.024 + 0.009 = 0.033

P(flu = yes | x) = 0.024 / 0.033 = 0.727

P(flu = NO | x) = 0.009 / 0.033 = 0.272.

Scanned by CamScanner

14/03/19

Distance Based algorithms:

classification is determining the what group something belongs in.

1. user sets a value for k.

2. for the data record that needs to be classified, the alg computes the distance b/w the datarecords and all of the reference data records.

3. find the k no. of closed data records

4. the majority of the class of the 'k' datarecords is the predicted class.

1. for a given dataset with the attributes weights as well as heights and the class label. find the person belongs to which class based on given weight 57kgs & height 170cm.

weight (x_1)	height (y_1)	class	distance
51	167	Under weight	6.7
62	182	normal	13
69	176	"	13.4
64	173	"	7.6
65	172	"	8.2
56	174	underweight	4.1
58	169	normal	1.4 (N1)
57	173	"	3 (N3)
55	170	"	2 (N2)

Step1: let K=3

Step2: distance can be find

$$= \sqrt{(57-51)^2+(170-167)^2} = 6.7$$

$$\sqrt{(57-62)^2+(170-182)^2} = 13$$

etc...)

	we	hie	class
80	57	170	normal

The datapoint of 3 neighbours belongs to normal class. Hence the given datapoint can be classified or placed into the normal class.

→ if no. of parameters for a given class label is 2, then it is called as binary classification

Remarks abt K: K must be +ve integer.

→ K must not be a multiple of a given no. of class labels.

→ K=1 the least dist. will be the neighbour for given data record.

2: The given data record contains 3 attributes name, age & gender and the class label is fan following. Find the person belongs to which class when the age=20 & gender=male age = 20, gender = M, fanfolls?

age	gender		fan following	name	dist
32	male	0	X_1	A	12
40	male	0	X_1	B	20
16	fem	1	X_2	C	4.16 (2)
14	f	1	X_1	D	6.1
55	m	0	X_1	E	35
40	m	0	X_1	F	20
20	f	1	X_2	G	1 (1)
15	m	0	X_1	H	5 (3)

$\sqrt{(5)^2+0}$

$\sqrt{25}$

let K=3.

assume the M=0, f=1

$$\sqrt{(20-32)^2+(0-0)^2} = 12. \quad \sqrt{}$$

$$\sqrt{(20)^2} = 20$$

$$\sqrt{16+1} = \sqrt{17}$$

$$\sqrt{36+1} = \sqrt{37}$$

etc...

the given datapoint can be classified or placed into the fanfollowing X_2.

3) For the given set of attributes age, income, student, credit rating, class label buys computer, find whether a person buys computer/not with a given values age = youth, income = medium, student = yes, credit rating = fair.

	age	income	student	credit rating	class buys computer
1.	youth	high	no	fair	no
2.	youth	high	no	excellent	no
3.	middle aged	high	no	fair	yes
4.	senior	medium	no	fair	yes
5.	senior	low	yes	fair	yes
6.	senior	low	yes	excellent	no
7.	middle aged	low	yes	excellent	yes
8.	youth	medium	no	fair	no
9.	youth	low	yes	fair	yes
10.	senior	medium	yes	fair	yes
11.	youth	medium	yes	excellent	yes
12.	middle aged	medium	no	excellent	yes
13.	middle aged	high	yes	fair	yes
14.	senior	medium	no	excellent	NO

ages	P_i	n_i	$IG(P_i\ n_i)$
youth	2	3	

$P(yes) = 9/14$

$P(NO) = 5/14$

$$IG(2,3) = -\frac{2}{5}\log_2\left(\frac{2}{5}\right) - \frac{3}{5}\log_2\left(\frac{3}{5}\right)$$

$$0.4\log_2(0.4) - 0.6\log_2(0.6)$$

age: youth

p(age = youth | buys - computers = yes) = 2/9

p(age = youth | buys - Computer = no) = 3/5

income: medium

p(income = medium | buys-computer = yes) = 4/9

P(income = medium | buys-Computer = no) = 2/5.

Student:

P(Student = yes | Buys-Computer = yes) = 6/9

P(Student = yes | Buys-Computer = no) = 1/5.

Credit-rating:

P(Credit-rating = fair | buys-computer = yes) = 6/9

P(Credit-rating = fair | buys-Computer = no) = 2/5

$P(X \mid \text{buys-computer} = \text{yes}) \times P(\text{buys} = \text{yes})$

$= \frac{2}{9} \times \frac{4}{9} \times \frac{6}{9} \times \frac{6}{9} \times \frac{9}{14}$

$= \frac{16}{567} = 0.025$

$P(X \mid \text{buys} = \text{No})\ P(\text{buys} = \text{No}) = \frac{3}{5} \times \frac{2}{5} \times \frac{1}{5} \times \frac{2}{5} \times \frac{5}{14}$

$= 0.0068$

$P(C) = 0.028 + 0.0068$

$= \cancel{0.00967} = 0.0348$

Scanned by CamScanner

for yes:

$$\frac{P(X|boys=yes)\,P(boys=yes)}{P(C)} = 0.028/0.034 = 0.80$$

for NO:

$$\frac{P(X|boys=no)\,P(boys=no)}{P(C)} = 0.0068/0.0348 = 0.19$$

2(2)
4(4)
1(1)
30 (33) ANC
24(28) MPI
27(29) CN
26(28) DMT

23(28)
26(29)
23(28)
30 (33)
MES
16(22)
29(33) MWT

18(20)
30(33)

The probability of yes is having high than probability of no

∴ The person can go for buying a computer

∴ This is false positive.

18/03/19

Decision tree

1. consider the following small datasets for the 2 classes of woods. using info gain construct a decision tree to classify the data set.

Density	Grain	Hardness	Class
Heavy	Small	Hard	Oak
Heavy	large	Hard	Oak
Heavy	Small	Hard	Oak
Light	large	soft	Oak
Light	large	Hard	pine
Heavy	Small	soft	Pine
Heavy	large	"	Pine
Heavy	Small	soft	Pine

→ which attribute the info gain chooses as the root of the tree.

→ what class does the tree infer for the Sample Density = Light, grain : small, hardness = hard

→ what class does the tree infer for the Sample d : light, grain = small, hardness = soft.

→ d : heavy, Grain = small, h = hard.

Assume p = Oak, n : pine.
4 4

$$IG = -\frac{p}{p+n}\log_2\left(\frac{p}{p+n}\right) - \frac{n}{p+n}\log_2\left(\frac{n}{p+n}\right)$$

$$IG(4,4) = -\frac{4}{8}\log_2\left(\frac{4}{8}\right) - \frac{4}{8}\log_2\left(\frac{4}{8}\right) \Rightarrow -\frac{1}{2}\log_2\left(\frac{1}{2}\right) - \frac{1}{2}\log_2\left(\frac{1}{2}\right)$$

if P, n = Same then IG = 1

Entropy for the density.

E for density we have

	P_i	n_i	$IG(P_i, n_i)$
heavy	3	3	1
light	1	1	1

$IG(3,3) = 1$

$IG(1,1) = 1$

$$E(d) = \sum_{i=1}^{v} \frac{P_i + n_i}{P+n} IG(P_i, n_i)$$

$= 8/8 = 1$

Entropy for Grain

	P_i	n_i	$IG(P_i, n_i)$
large	2	2	1
Small	2	2	1

$IG(D) = IG - E(d)$

$= 1 - 1$

$= 0.$

$IG(2,2) = 1$

$IG(2,2) = 1$

$$E(G) = \sum_{i=1}^{v} \frac{P_i + n_i}{P+n} IG(P_i, n_i)$$

$$\frac{4}{8}(1) + \frac{4}{8}(1) \Rightarrow \frac{1}{2} + \frac{1}{2} = 1$$

$IG(G) = IG - E(G)$

$= 1 - 1 = 0.$

Entropy for hardness

	P_i	n_i	$IG(P_i, n_i)$
Hard	3	1	0.81
Soft	1	3	0.81

$$IG(3,1) = -\frac{3}{4}\log_2\left(\frac{3}{4}\right) - \frac{1}{4}\log_2\left(\frac{1}{4}\right)$$

$$= -0.75\log_2(0.75) - 0.25\log_2(0.25)$$

$$= 0.81$$

$$IG(1,3) = -\frac{1}{4}\log_2\left(\frac{1}{4}\right) - \frac{3}{4}\log_2\left(\frac{3}{4}\right)$$

$$= -0.25\log_2(0.25) - 0.75\log_2(0.75)$$

$$= 0.81$$

$$E(h) = \sum_{i=1}^{4} \frac{P_i + n_i}{p+n} IG(P_i, n_i)$$

$$= \frac{4}{8}(0.81) + \frac{4}{8}(0.81)$$

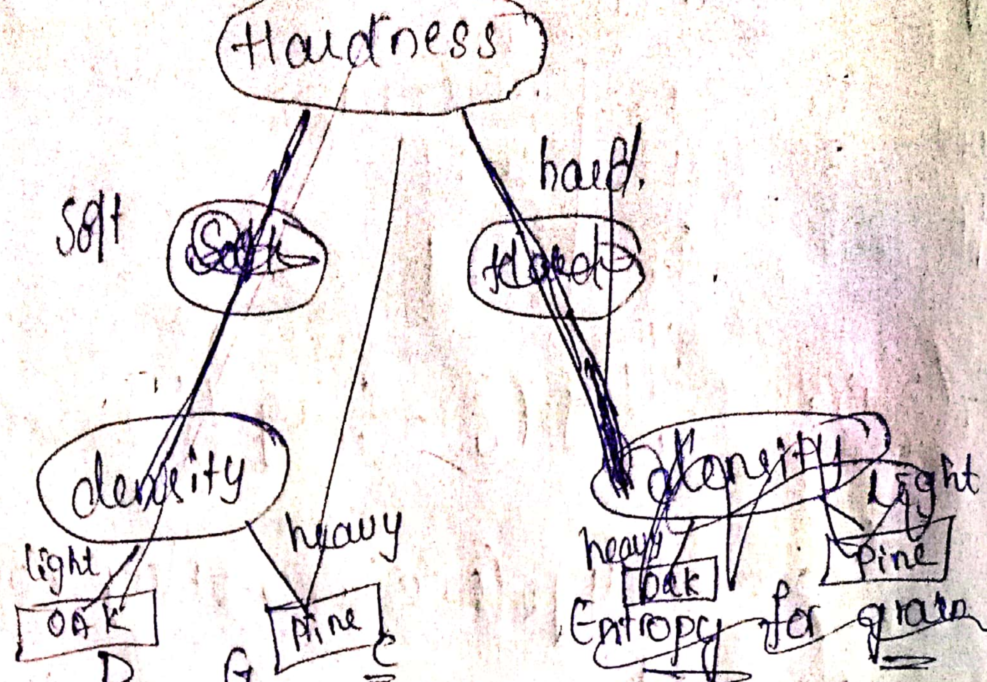
$$= \cancel{1.622}\ 0.81$$

$$G(h) = 1 - E(h)$$

$$= 1 - 0.81$$

$$= \cancel{0.622}\ 0.19$$

$$\left.\begin{array}{l} G(D) = 0 \\ G(G) = 0 \\ G(H) = \cancel{0.622}\ 0.19 \end{array}\right\}$$ highest value is G(H)



Soft

Hardness

Entropy for grain

Hardness	D	G	C
Soft	Light	Large	Oak
Soft	Heavy	Small	pine
Soft	Heavy	Large	"
Soft	Heavy	Small	"

p	n
1	3

$\boxed{IG(1,3) = 0.81}$

Entropy for density

Light	1	0	0
heavy	0	3	0

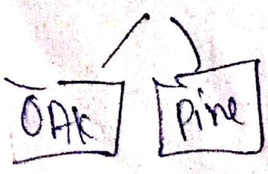
$E(D) = 0$

$Gain(D) = 0.81 - 0.$

now

$G(D) = 0.81$ } max.
$G(G) = 0.31$ }

density



bcoz Heavy for pine in above datatable.

Entropy for Grain

large	1	1	1
Small	0	2	0

~~I(0,2)~~ =

$E(G) = \frac{2}{4}(1) + = \frac{1}{2}$

$G(G) = 0.81 - 0.5$

$G(G) = 0.31$

Scanned by CamScanner

hardness	d	Grain	class
hard	heavy	Small	Oak
hard	heavy	large	Oak
hard	heavy	Small	Oak
hard	light	large	pine

Entropy (d) $E(d)$

	P_i	n_i	$IG(P_i, n_i)$
heavy	3	0	0
light	0	1	0

$E(D) = 0$

$G(D) : 0.81 - 0$

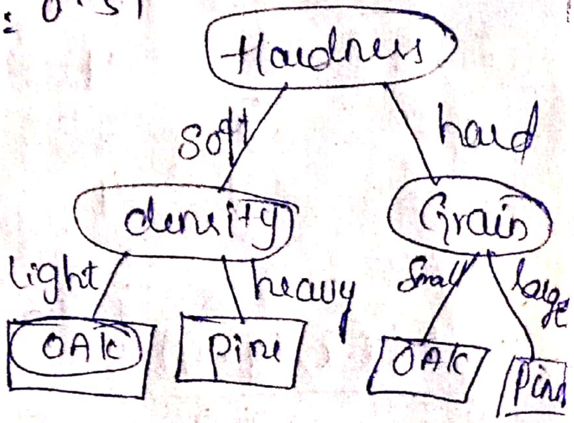
Entropy (G)

small	2	0	0
large	1	1	1

$E(G) = \frac{1}{2}$

$G(G) : 0.31$

$$\left.\begin{array}{l} G(D) = 0.81 \\ G(G) = 0.31 \end{array}\right\} \text{— max}$$



Grain	d	h	Class
large	heavy	hard	Oak
large	light	soft	Oak
large	light	hard	pine
large	heavy	soft	pine

$E(d)$

heavy	1	1	1
light	1	1	1

$IG(d) = 1 - 0$
$= 1$

11/03/18

Statistical Based algorithms

1. Bayes
2. regression line,

→ find the regression line for given data.
By using the regression we can predict the regression line or the best fit line

→ If there is only one dependent variable then it is called as linear regression

$y = mx + c$.

→ In the multiple regression we may have diff no. of dependent variables.

x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})(y-\bar{y})$	$(x-\bar{x})^2$
1	3	-2	-0.6	1.2	4
2	4	-1	0.4	0.4	1
3	2	0	-1.6	0	0
4	4	1	0.4	0.4	1
5	5	2	1.4	2.8	4
$\bar{x} = 3$	$\bar{y} = 3.6$			4.0	10

$$m = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\Sigma(x-\bar{x})^2} = \frac{4.0}{10} = \cancel{0.48}\ 0.4.$$

$y = mx + c$

$\cancel{3.6}\ 1.2 = 0.4(3) + c$

$c = 2.4$

$$\boxed{y = 0.4x + 2.4}$$

$x=1 \quad y = 0.4 + 2.4 \Rightarrow 2.8$

$x=2 \quad y = 3.2$

$x=3 \quad y = 3.6$

$x=4 \quad y = 4.6$

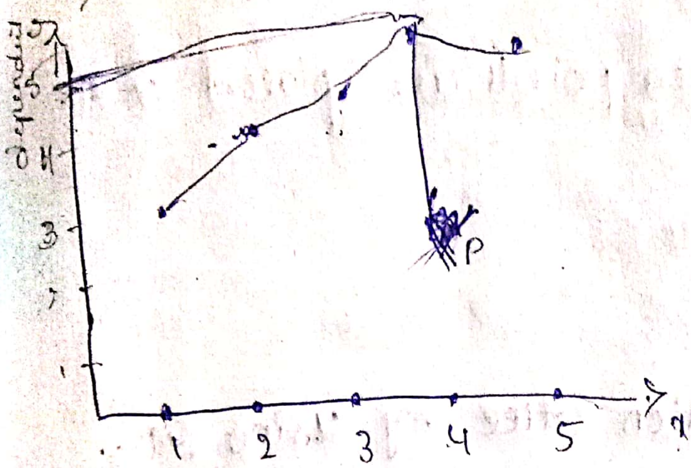
$x=5 \quad y = 4.4$

0.4(2) 0.8 + 2.4 = 3.2

0.4 × 3 = 1.2, 1.2 + 2.4 = 3.6

0.4 × 4 = 1.6, 1.6 + 2.4 = 4.0

0.4 × 5 = 2.0, 2.0 + 2.4 = 4.4



The distance b/w actual point and the predicted line can be called as error.

** The line with the least error is regression line of best fit.

→ R^2 is a statistical measure of how close the data are fitted to the regression line.

→ It is also known as coefficient of determination.

$$R^2 = \frac{\Sigma (y_p - \bar{y})^2}{\Sigma (y - \bar{y})^2}$$

y_p is predicted values

x	y	y_p	$y_p - \bar{y}$	$(y_p - \bar{y})^2$	$y - \bar{y}$	$(y - \bar{y})^2$
1	3	2.8	-0.8	0.64	-0.6	0.36
2	4	3.2	-0.4	0.16	0.4	0.16
3	2	3.6	0	0	-1.6	2.56
4	4	4.6	0.4	0.16	0.4	0.16
5	5	4.4	0.8	0.64	1.4	1.16
	$\bar{y} = 3.6$			$\Sigma = 1.6$		$\Sigma = 5.2$

$R^2 = \frac{1.6}{5.2} = 0.32.$

$R^2 = 0.32$ (is not a good fit line) that means the data points are far away from the regression line.

Suppose $R^2 = 1$ the data points are placed in a regression line

25/03/19:

1. Construct a descision tree by using ID_3.

The procedure for constructing a decision by using ID_3 is similar to procedure of J_{48}. But in the ID_3 entropy can be used as a splitting parameter.

→ whereas in J_{48} gain ratio of gain can be used as splitting parameter.

→ $C_{4.5}$ Enhances improves the ID_3 alg in the following ways.

It is similar to Id_3.

→ (i) missing data: whether decision tree is built missing data is ignored.

→ (ii) pruning: there are 2 primary pruning strategies proposed in $C_{4.5}$ the 1^{st} one i

Subtree replacement, the second one is Subtree raising.

Subtree replacement: A subtree is replaced by a leaf node. if this replacement results in a error rate close to that of original tree.

→ Subtree replacement works from the bottom of the tree upto the root.

Subtree raising: It replaces a subtree by its most used subtree.

(iii) Splitting: C4.5 uses the largest gain ratio for the splitting

$$\text{gain ratio}(D,S) = \frac{\text{Gain}(D,S)}{H\left(\frac{|D_1|}{|D|} \cdots \frac{|D_S|}{|D|}\right)}$$

$$\text{Gain}(D) = H(D) - \sum_{i=1}^{S} P(D_i)\, H(D_i)$$

$$H(P_1, P_2 - P_3) = \sum_{i=1}^{S}\left(P_i - \log\left(\frac{1}{P_i}\right)\right)$$

CART: Classification and regression tree is a technique that generates a binary decision tree. As with ID3 entropy is used as a measure to choose the best splitting attribute.

Neural Network Based alg

Gender	Height	class
F	1.6m	short
M	2m	tall
F	1.9m	medium
F	1.88m	medium
F	1.7m	short
M	1.45m	medium
F	1.7m	Tall

Advantages: Neural n/w's are more robust than the decision tree because of the weights

→ NN improves its performance by learning.

→ There is a low error rate and thus a high degree of accuracy once the appropriate training has been performed.

→ NN are more robust than decision trees in the noisy environments.

Neural N/w: A neural n/w is a set of connected i/p & o/p units or nodes in which each connection has a weight associated with it.

disadvantages of ANN:

Neural n/w's are difficult to understand

→ Generating rules from NN is not straight forward.

→ learning As with decision trees. Overfitting may result.

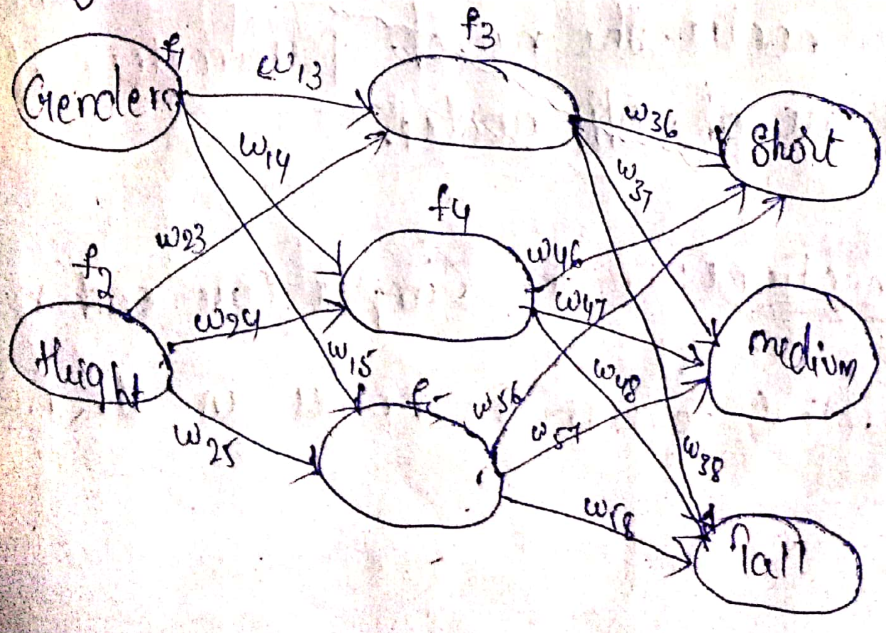
→ The learning phase may fail to converge

→ ANN may be quite expensive to use.

→ I/p att values must be numeric

E.g. for numeric n/w dataset

name	Gender	Height	class
K	f	1.6m	short
Tim	m	2m	tall
maggie	f	1.9m	middium
Bob	m	1.88m	medium
Dave	M	1.7m	short
steven	m	2.2m	tall.



Scanned by CamScanner

There are Several issues to be examined

attributes (no. of. source nodes): att. Can be used as i/ps (or) Source nodes.

Issues

→ Determining which attributes to use as i/ps is an issue.

→ no. of. hidden layers: in the simpler case there is only one hidden layer.

Hidden layer: the layer which is in b/w source node and destination node.

no. of. hidden nodes: the nodes of hidden layer Called as hidden nodes.

Choosing the best no. of. hidden nodes for hidden layer is one of the difficult pro's in neural n/w's.

Training data: no. of. Sinks (or) destination of o/p node the class parameter Can be used as the destination nodes. The no. of. parameters of a class are no. of. o/p nodes.

~~Choose~~

InterConnections: In the Simpler Case each node is connected to all the nodes in the next level.

weight: The weight assigned to the edges indicates the relative weight b/w 2 nodes.
-> initial weights are usually assumed to be a small +ve number and are assigned randomly.

Activation functions: many diff types of activation functions can be used.

learning technique: The technique for adjusting the weights is called learning technique. many approaches can be used but the most common approach is some form of back propogation.

Stop: the learning may stop when all the training tuples have propogated through the n/w.

Perceptron: A single neural n/w is called perceptron.

-> A perceptron is a single neuron with multiple i/ps & one o/p.

-> The activation function can be used as a sigmoidal function

-> A simple perception can be used to classify into 2 classes.

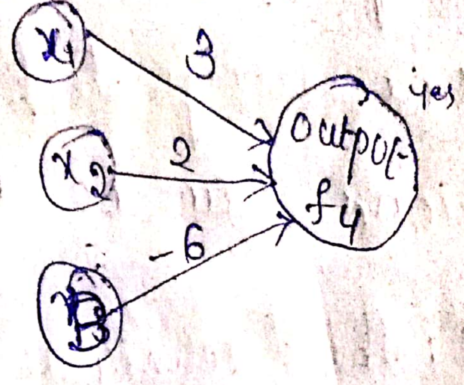
-> using a unipolar activation function an o/p of 1 would be used to classify into 1 class while o/p of 0 ~~can~~ would be classified into another class.

Perceptron

Sigmoidal

if 6 7 4 2

2 8 5 3 2



activation fun of the total o/p of y_i

$S = 3x_1 + 2x_2 - 6.$

$3+4-6 = 1$ yes.

$6+6-6 = 6$ yes

$12+10-6 = 16$ yes

$\begin{array}{r} 21 \\ 16 \\ \hline 37 \end{array} - 6 = 31$

16 = yes

if $S > 0$ $\quad f_4 = \begin{cases} 1 \\ 0 \end{cases}$

otherwise

let us take simple dataset

n	z	class	predicted
1	2	yes	yes (TP)
2	3	No	yes (FP)
4	5	yes	yes (TP)
7	8	yes	yes (TP)
6	2	NO	yes (FP)

$\frac{\text{Correctly predicted}}{\text{Total}} \Leftarrow \frac{3}{5} = 0.6 = 60\%$ predicted.

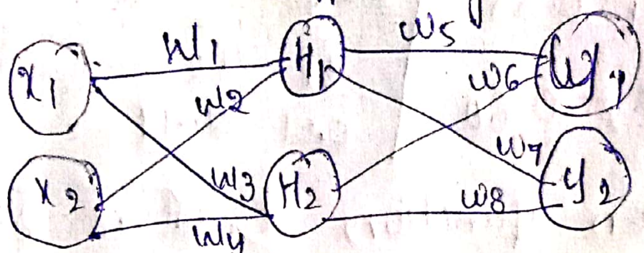
Scanned by CamScanner

Propagation: Normal approach used for processing is called propagation.

Processing: Taking i/p & producing o/p is called processing.

Given a tuple of values i/p to the neural n/w,

$X = \{x_1, x_2, \ldots x_n\}$ one value is i/p at each node in the i/p layer.



$y_1 = H_1 w_5 + H_2 w_6$

$y_2 = H_1 w_7 + H_2 w_8$

$H_1 = x_1 w_1 + x_2 w_2$

$H_2 = x_1 w_3 + x_2 w_4$

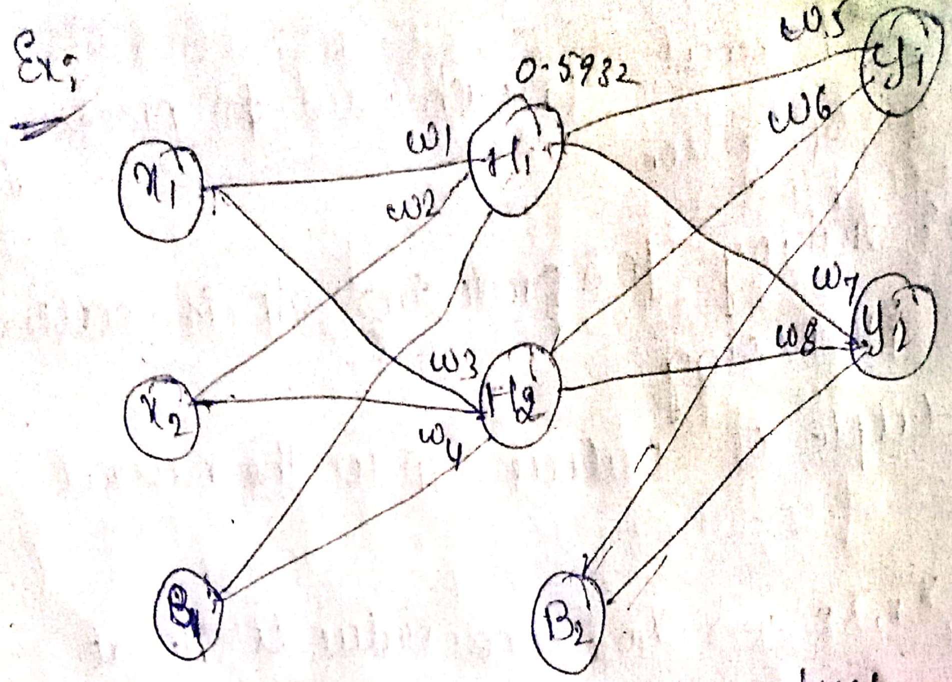
Then the summation and activation functions are applied at each node with an o/p value created for each o/p ~~created~~ edge from that node.

→ These values are inturn send to the subsequent nodes.

→ This process continues until a tuple of o/p values $y = \{y_1, y_2, y_3 \ldots y_n\}$

Back propogation: It is a learning technique that adjust weights ~~edges coeffs~~ in the neural n/w by propagating wat changes backward from sink to the same nodes

Ex:



$w_1 = 0.15$
$w_2 = 0.20$
$w_3 = 0.25$
$w_4 = 0.30$
$w_5 = 0.40$
$w_6 = 0.45$
$w_7 = 0.50$
$w_8 = 0.55$

		Target value	
$x_1 = 0.05$	$B_1 = 0.35$	T_1	T_2
$x_2 = 0.10$	$B_2 = 0.60$	0.01	0.99

the activation function used is sigmoidal

$$\text{i.e., } = \frac{1}{1+e^{-x}}$$

Cal the value H_1

$$H_1 = x_1 w_1 + x_2 w_2 + B_1$$

$$= 0.05 \times 0.15 + 0.10 \times 0.20 + 0.35$$

$$= 0.3775$$

$$\text{O/p of } H_1 = \frac{1}{1+e^{-H_1}} = \frac{1}{1+e^{-0.3775}}$$

$$= 0.5932.$$

$$H_2 = x_1 w_3 + x_2 w_4 + B_1$$

$$= 0.05(0.25) + 0.10(0.30) + 0.35$$

$$= \cancel{0.3425}\ \cancel{0.468}\ 0.3425$$

o/p of $H_2 = \dfrac{1}{1+e^{-0.5988}} = 0.5968.$

$y_1 = H_1 w_5 + H_2 w_6 + B_2$

$= 0.5932(0.40) + 0.5968(0.45) + 0.60$

$= \cancel{0.9288} \quad \cancel{0.5932} = 1.10584$

o/p of $y_1 = \dfrac{1}{1+e^{-y_1}} = \dfrac{1}{1+e^{-1.105}} = 0.75.$

$y_2 = H_1 w_7 + H_2 w_8 + B_2$

$= 0.5932(0.50) + (0.5968(0.55) + 0.60$

$= 1.22485$

o/p of $y_2 = \dfrac{1}{1+e^{-y_2}} = \dfrac{1}{1+e^{-1.225}} = 0.77290$

The o/p of y_1 & y_2 are not matching with T_1, T_2.

Hence back propagate i.e., adjust the ~~wait~~ weights.

values of all the weights.

→ the error for each o/p node can be cal by

$$\sum_{i=1}^{m} \frac{(\text{Target} - \text{output of } y_1)}{m} \quad m = \text{no. of o/p nodes}$$

$E_{total} = \frac{1}{2}(T_1 - \text{o/p of } y_1) + \frac{1}{2}(T_2 - \text{output of } y_2)$