19-07-18 (Thursday)

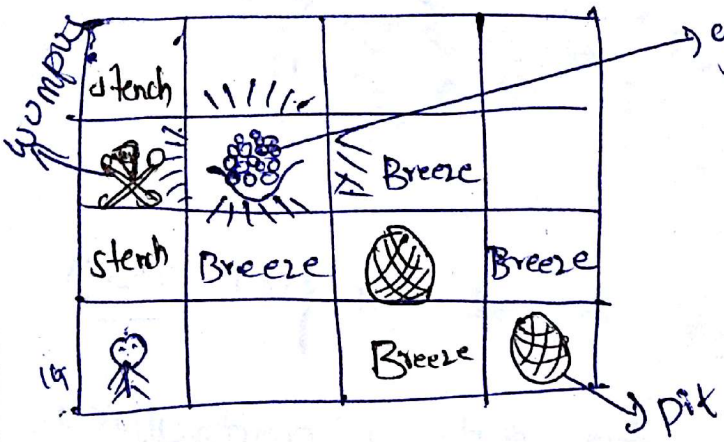
Unit-2 – Knowledge Based systems

Knowledge Based systems

Wumpus game:-

> for this, graph theory is not visible since environment is not suitable

- has to note clues & use whenever it require.



≋ – glittering – gold

Breeze (air) – pit

stench (bad smell) – wumpus.

- has to note clues

Disadvantages over graphs

(0,0)

> has no option for noting clues ~~environment~~

> environment is partially observable

> No Logic

So to make all this possible We are using "discrete mathematics"

this wumpus game is solved by discrete Mathematics so it is called 'knowledge Based systems'

i) Knowledge Base:- Rules & regulations & observations
i.e Rules & Regulations + Background knowledge

ii) Background knowledge:-

iii) Knowledge Representation :- (Discrete / propositional logic)

iv) Inference:- Based on previous knowledge we can deduce new knowledge.

v) TELL :-

vi) Ask :- Querying

Approaches:-

i) procedural Approach:- Total information is written in the form of programming

ii) Declarative approach:- information is represented mathematically.

iii) entailment:- $\alpha \models \beta$ or $\alpha \vdash \beta$

by using α we deduce β.

ex:- $4+5=9 \models 9=4+5$

by using these we can deduce these.

$\alpha \vdash_A \beta$

Using Alogarithm A we deduce α to β.

propositions:-

$\wedge$ – AND

$\vee$ – OR

$\sim$ – NOT

$\rightarrow$ – implies

$\Leftrightarrow$ – double implies

formalizing simple sentences:-

A = Aldo is Italian

B = Bob is english

a) $\sim A$

b) $A \cup B$

c) $A \cap B$

d) $A \cup (\sim A \rightarrow \sim B)$

e) $(A \cap B) \cup (\sim A \cap \sim B) = A \Leftrightarrow B$

Inference rules

Given	Inference	Rulename
1) $P \cap q$	P, q	Simplification
$\sim(P \cup q)$	$\sim P, \sim q$	
2) P, q	$P \cap q$	Conjunction
3) P	$P \cup q$	Disjunction addition
4) $\sim(P \cap q), P$	$\sim q$	Conjunctive Argument
$\sim(P \cap q), q$	$\sim P$	
5) $P \rightarrow q, P$	q	detachment

6)	$p \to q$	$\neg q \to \neg p$	contrapositive
7)	$p \to q, q \to r$	$p \to r$	chain Rule
8)	$p \to q, \neg q$	$\neg p$	Modus Tollens
9)	$p \to q, p$	q	Modus ponens
10)	$(p \vee q), \neg p$	q	Disjunctive syllogism
11)	$(p \to q) \wedge (r \to s)(p \vee r)$	$q \vee s$	constructive Dilemma
12)	$(p \to Q) \wedge (R \to S) \neg Q \cup \neg S$	$P \cup R$	Destructive Dilemma
13)	$\neg(p \cap Q)$	$\neg p \cup \neg Q$	Demorgan's law
	$\neg(p \cup Q)$	$\neg p \cap \neg Q$	
14)	$\neg(\neg A)$	A	Double negation
15)	$A \Leftrightarrow B, A$	B	
16)	$A \Leftrightarrow B, B$	A	
16)	$p \to Q$ $\neg p \to r$	$q \vee r$	Resolution

(f) It is snowing, It is cold
S
It is cold, John is wearing coat
C J
It is snowing.

$S \to C$
$C \to J$
$\frac{S \to \text{true}}{J}$

(P) i) If it is raining, it is not cold
If it is not raining, John is not wearing a coat

ii) It is cold.
John is not wearing the coat

$R \rightarrow \sim C$

$\sim R \rightarrow \sim J$

$C, \sim R, \sim J$

(P) i) If polo is thin, then carlo is not blande or Robert is not tall.

ii) If Roberta is tall, the sandra is a lovely

iii) If sandra is lovely and carlo is blande then pablo is thin.

iv) carlo is blande

Is Roberta tall?

Sol:- $P \rightarrow (\sim C \cup \sim R)$

$R \rightarrow S$

$(S \cap C) \rightarrow P$

C

R?

C - true
S - true

C ∩ S - true then P - true.

P → (¬C ∪ ¬S)

either one of them false

means P is false.

→

a → b
true def true.
false maybe

a <=> b
true def true
false def false

propositional logic:

20-07-18 (Friday)

> $\omega(i,j) \iff (S(i-1,j) \cap S(i+1,j) \cap S(i,j-1) \cap S(i,j+1))$

$S(i,j) \iff ((\omega(i-1,j) \cup \omega(i+1,j) \cup \omega(i,j-1) \cup \omega(i,j+1))$

> $P(i,j) \iff (B(i-1,j) \cap B(i+1,j) \cap B(i,j-1) \cap B(i,j+1))$

$B(i,j) \iff$

> $G(i,j)$

$G(i,j)$

$S(i,j) \iff w(\text{adj rooms})$

false $\iff$ false

$A \iff B$

$\neg A$

$\neg B$

At $(0,0)$

$\neg w(0,0)$

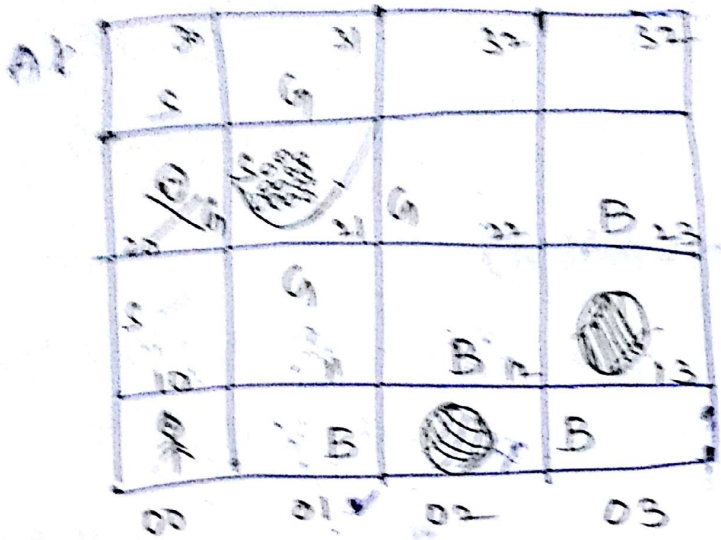
$\neg s(0,0)$

$\neg w(0,1) \cdot \cap \neg w(1,0)$

$S(i,j) = [w(i-1,j) \cup w(i+1,j) \cup w(\ ) \cup w(\ )]$

$\text{Safe}(i,j) \rightarrow \neg(w(i,j))$ { Stench - A, Wumpus - B } at no wumpus

So at $(0,0)$ there is no wumpus



G - Glitter - by gold
B - Breeze - by pit
Stench - wumpus

At (0,0)

$\sim W(0,0)$

$\sim P(0,0)$

$\sim Gold(0,0)$

$\sim Stench(0,0)$

$\hookrightarrow \sim W(0,1) \cap \sim W(1,0)$

$\sim B(0,0)$

$\hookrightarrow \sim P(0,1) \cap \sim P(1,0)$

Safe(0,1): $\sim W(0,1) \cap \sim P(0,1)$

Safe(1,0): $\sim W(1,0) \cap \sim P(1,0)$

$\sim Glitter(0,0)$

$\hookrightarrow \sim G(0,1) \cap \sim G(1,0)$

Sate

2,1
1,2
1,1
1,0
0,1
00

$\sim W$ = NO
$W(0,1)$ = wumpus at (0,1)
$\sim P$ = NO
$P(1,0)$ = pit at (1,0)

At (0,1)

$\sim S(0,1)$

$\hookrightarrow \sim W(0,2) \cap \sim W(1,1)$

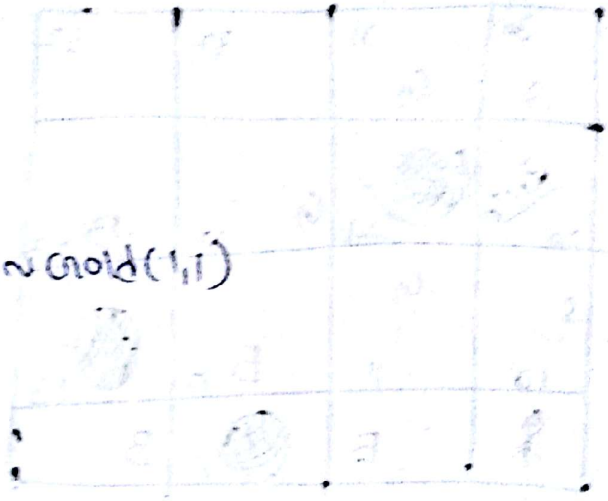
$B(0,1)$

$\hookrightarrow P(0,2) \vee P(1,1)$

Safe(0,2) = X

Safe(1,1) = X

$\sim Gold(0,2) \cap \sim Gold(1,1)$



At(1,0)

$\sim B(1,0)$

↳ $\sim P(1,1) \vee \sim\sim P(2,0) \cap \sim P(0,0)$

S(1,0) :

↳ $W(1,1) \vee W(2,0) \vee W(0,0)$

Safe(1,1) $\Rightarrow \sim W(1,1) \cap \sim P(1,1)$

Safe(2,0) $\Rightarrow$ X

At(1,1)

$\sim S(1,1)$

$\sim\sim W(1,2) \cap \sim W(1,0) \cap \sim W(0,1) \cap \sim W(2,1)$

$\sim B(1,1)$

↳ $\sim P(1,2) \cap \sim P(2,1)$

Safe(1,2) ✓

Safe(2,1) ✓

Glitter(1,1)

Gold(1,2) $\vee$ Gold(2,1)

At (2,1)

Cold (2,1)

At (1,2)

→ NS (1,2)

↳ NW (1,3) ∩ NW (2,2)

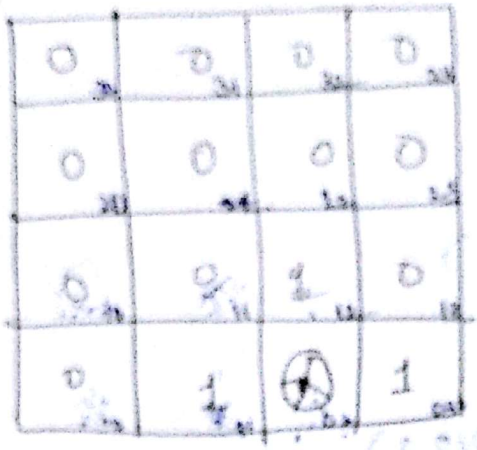
→ B (1,2)

↳ P(~~5,2~~) ∪ P(1,3) ∪ P(2,2)

Safe (1,3) X

Safe (2,2) X

Mini - Mine Sweeper Game development



At

$m(i,j) \Leftrightarrow 1(i-1,j) \cap 1(i+1,j) \cap 1(i,j-1) \cap 1(i,j+1)$

At (0,0)

$\sim 1(0,0)$

$\hookrightarrow \sim m(1,0) \cap \sim m(0,1)$

Safe (1,0)

Safe (0,1)

At (1,0)

$\hookrightarrow \sim m(1,0)$

$\hookrightarrow \sim M(1,1) \cup \sim M(2,0) \cup \sim M(0,0)$

means $\sim M(1,1)$

$\sim M(0,1)$

$M(0,2)$ //

* by inference complexity increases so we should follow alternative which is new representation which is first order logic.
* Inference can be solved by propositional logic

First order logic

1) ∀
2) ∃

bought (Jack, dvd) = Jack bought dvd.
verb subject object

> ∃x (bought (Jack, x)) = ∃ some x which Jack bought / Jack bought something.

> ∀x (bought (Jack, x)) = ∀x Jack bought / Jack bought everything

> ∀x ∃y (Bought (x, y)) = Everyone bought something / all bought something

> ∃x ∀y (Bought (x, y)) = Someone bought everything

> ∀x (Bought (Jack, x) → Bought (susan, x)) = Everything if it is bought by Jack than bought by susan

> ∀x (Bought (Jack, x)) ∩ ∀x (Bought (susan, x)) = Jack bought everything and susan bought everything

> EveryBstudents, loves somestudents

∀x (student (x) → ∃y (student (y) ∩ loves (x, y)) = for all x if there is student x ∃y who is also student and loves some students.

> Every student loved by some student.

relation b/w universal & Existential quantifier

$\forall x \sim P(x) \equiv \sim \exists x P(x)$

$\sim \forall x P(x) \equiv \exists x \sim P(x)$

$\sim \forall x \sim P(x) \equiv \exists x P(x)$

$\forall x P(x) \equiv \sim \exists x \sim P(x)$.

(5) 24-07-18 (Tuesday)

> every greedy king is evil → $\forall x$ (king(x) ∩ greedy(x) → evil(x))

> John is a king → king(John)

> John is greedy → greedy(John)

> Richard is John's brother. (Richard, John)

> Who is evil ? — evil(?)

> To convert first order logic into propositional logic remove universal, Existential instantiation & skolem ex for above examples

to remove universal, Existential instantiation in place of $\forall, \exists$ if there is noun then place it in $\forall, \exists$.

ex: propositional logic

i) king(John) ∩ greedy(John) → evil(John)

ii) king(Richard) ∩ greedy(Richard) → evil(Richard)

iii) king(John)

iv) Greedy(John)

v) Brother(Richard, John)

Who is evil? - John is evil (∵ John is king
John is greedy
every greedy king is evil)

(P) i) Jack own's dog — ∃x (dog(x) ∩ owns(Jack, x))

ii) every dog owner is animal lover → ∀x ∃y (owns(x, animal(y)) → ∀z (animal(z) ∩ loves(x, z))

iii) All cats & Dogs are animal → ∀x (dog(x) ∪ cat(x) → animal(x))

iv) no animal lover kills an animal.

v) Jack (or) Quorisity killed tuna the cat

who killed tuna?

First order logic

own's (Jack, dog)

> ∃x (dog(x) ∩ owns(Jack, x))

> ∀x (dog(x) ∪ cat(x) → animal(x)) (lover)

> cat(Tuna) ⇒ Animal(Tuna)

> ∀x ∃y (owns(x, animal(y)) → ∀z (animal(z) ∩ loves(x, z))

> ∀x ∀y (animal(y) ∩ loves(x, y) → ∼∃z [animal(z) ∩ kill (x, z)])

> kill(Jack, Animal(Tuna)) ∪ kill(coriosity, Animal(Tuna))

1 (P) > every greedy king is evil

> John is a king

> someone is greedy.

> Richard is John's brother

> who is evil? —

Isol: king(John) ∩ greedy(John) → evil(John)

ii sol: king(Richard) ∩ greedy(Richard) → evil(Richard)

iii) king(John)

iv) greedy(P1)

P1 & John or Richard

v) Brother(John, Richard)

who is evil? → P1 & John or Richard

---

2 sol) Converting into FOL (First order logic)

- If there is ∀ means we have to write all its respective noun
- If there is ∃ means write skolem constant like A, B...

dog(D1) ∩ owns(Jack, D1)

dog(D1) ∪ cat(P1) → animal(D1)

dog(Tuna) ∪ cat(Tuna) → animal(Tuna)

dog.

Jack owns a dog

All dog owners are animal lovers

No animal lover kills animals

Jack or curiosity killed an animal.

Who killed the animal?

owns(Jack, A), loves(P) →

Sol:- First order logic

i) owns(Jack, dog)

ii) ∀x (owns(x, dog) → animal_lover(x))

iii) ∀x (animal_lover(x) → ¬kill(x, animal))

iv) kill(Jack, animal) ∪ kill(Curiosity, animal)

Inference:-

i) is in Inference only

ii) has ∀ so by removing ∀ we have to apply for all x.

- owns(Jack, dog) → animal_lover(Jack)
- owns(Curiosity, dog) → animal_lover(Jack)

iii) has ∀ so remove ∀ & apply/substitute vars in x.

- animal_lover(Jack) → ¬kill(Jack, animal)
- animal_lover(Curiosity) → ¬kill(Curiosity, animal)

how to get sol?:

proof
- owns(Jack, dog)
- animal_lover(Jack)
- ¬kill(Jack, animal)

proof sol: kill(Curiosity, animal) ✓

1 (P) The law says, it is crime for an American to sell weapons to hostile nations. The country Nono, an enemy of America has some missiles. All of its missiles are sold to it by colonel west, who is an American. prove that west is criminal.

First order logic

$\exists x \exists y \exists z (American(x) \wedge weapon(y) \wedge hostile(z) \wedge$

$sells(x,y,z) \rightarrow criminal(x))$

$hostile(Nono)$

$\exists y (weapon(y) \wedge owns(Nono,y))$

$\forall y (weapons(y) \wedge owns(Nono,y)) \rightarrow sells(west,y,Nono)$

Inference

$\rightarrow American(west)$

$\rightarrow hostile(Nono)$

$\rightarrow weapon(w1) \wedge owns(Nono,w1)$

$\rightarrow weapon(w1) \wedge owns(Nono,w1) \rightarrow sells(west,w1,Nono)$

proof:-

hostile → None

American → west

weapon → w

sells(west,w,Nono)

Hence west is criminal

Inference in FOL (first order logic)

i) eliminate implication

A	B	A → B
0	0	1
0	1	1
1	0	0
1	1	1

ii) Move ¬ inwards

iii) Standardize variables:- use separate variable if there is interference

iv) skolemization (∃x remove ∃)

v) Drop universal quantification

vi) Distribute ∧ over ∨

$(A \wedge B) \vee C$

$\Rightarrow (A \vee C) \wedge (B \vee C)$

to avoid

Forward chaining:- step-by-step from first repeating cycles till goal state by forward chaining only

1) $A \wedge B \rightarrow F$
2) $A \wedge E \rightarrow G$
3) $B \rightarrow E$
4) $G \rightarrow D$

database (true variables are kept in this)

A, B, F, E, G, D

Prove that if A and B are true, D is true.

1 Backward chaining:-

1) $A \cap B \rightarrow F$
2) $A \cap E \rightarrow G$
3) $B \rightarrow E$
4) $G \rightarrow D$

Backward chaining

Database

A, B, E, G, D

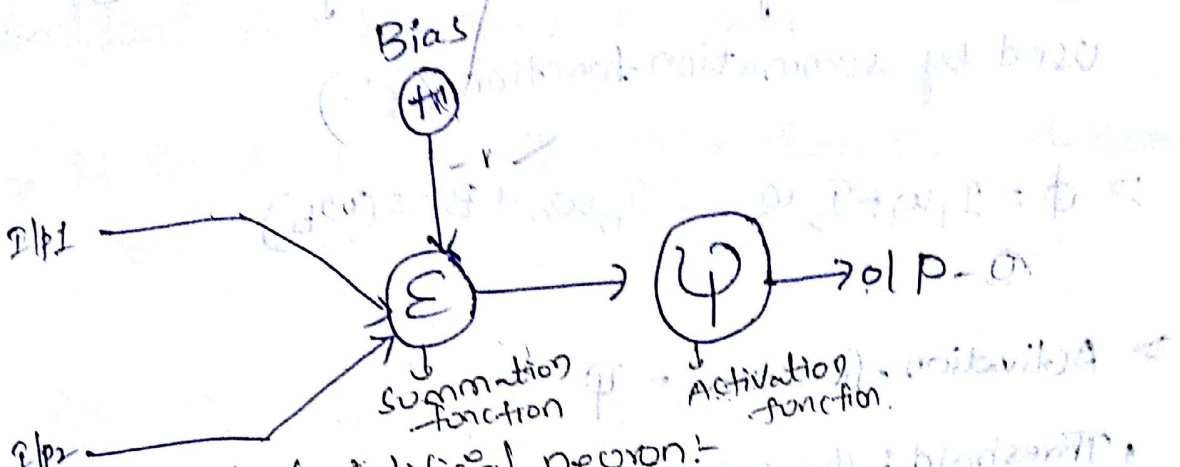
Prove that, if A and B are true, Distributive

So!

D -target
G
A
B
E
B

Backward programming is better than forward programming

Machine learning:



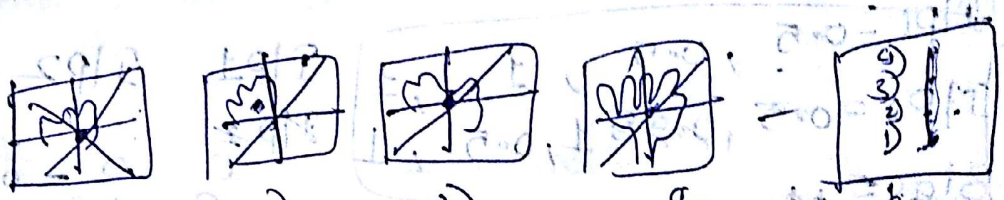
model of Artificial neuron:-

classification:-

i) Classification: Apple, Mango - Apple should recognized by machine.

ii) optical character recognition

ex:- A A A A - should recognize character.



normalization

ex:- i) image recognition → here hand is up so volume raise

ex:- ii) database recognition:- In hospital new patient recognition.

ex:- iii) text:- search engine:- even though we type wrong in google it gives right song information

ex:- iv) machine translation:- one language to other langu

ex:- v) ~~text~~ audio:-

fig:- Model of Artificial neuron.

> for each & every input (I_1, I_2) there is weight (w_1, w_2) which is used by summation function ($\in$)

> $\phi = I_1 w_1 + I_2 w_2 \cdots I_n w_n + \text{Bias}(w_b)$

> Activation function. $-\varphi$

- Threshold : $\phi \geq 0 \Rightarrow 1$
  $\phi < 0 \Rightarrow 0$
- tanh : $\tanh(\phi)$
- Signoid

(P) > By using AND gates we get o/p

I/P1	I/P2	I/P3
0	0	0
0	1	0
1	0	0
1	1	1

I/P1 = 0.5, 0.5, 1, 1
I/P2 = 0.5, 0.5, 1, 1
Bias = -0.75

$w1 = 0.5,\ w2 = 0.5,\ w3 = -0.75$

$\phi = I_1 w_1 + I_2 w_2 \cdots I_n w_n + \text{Bias}(w_b)$

$\phi = I_1(0.5) + I_2(0.5) \cdots + 1(-0.75)$

$\phi = -0.25$

by threshold function $\phi < 0 \Rightarrow 0$

0	0	0
0	1	0
1	0	0
1	1	1

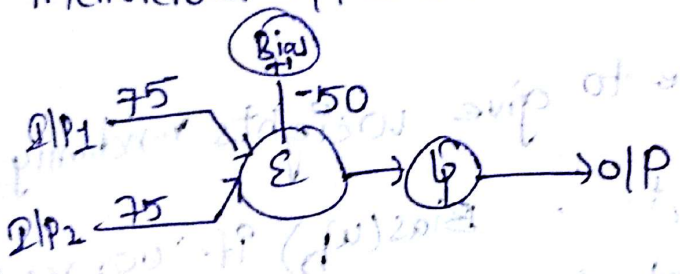
> for AND gate — I/P1 } if These are Individual
— I/P2 }

should be then dominated by Bias function.

> if I/P1 & I/P2 combine then they should dominate Bias function.

OR gate by neural network

> individual i/p's should dominate Bias function

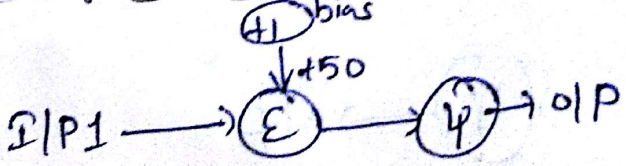


$\phi = I_1 w_1 + I_2 w_2 + \dots I_n w_n + \text{Bias}(w_b)$

threshold fun:- $\phi \geq 0 \Rightarrow 1$

$\phi < 0 \Rightarrow 0$

Not gate by neural network



$0(-100) + 1(50) \Rightarrow 1$

$1(-100) + 1(50) \Rightarrow 0$

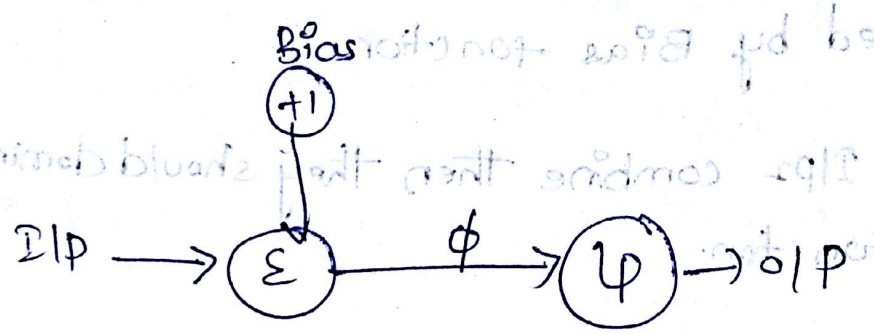
$\phi = I_1 w_1 + I_2 w_2 \dots I_n w_n + \text{Bias}(w_b)$

threshold fun:- $\phi \geq 0 \Rightarrow 1$

$\phi < 0 \Rightarrow 0$

Back propogation Algorithm

(How machines learns? ...)



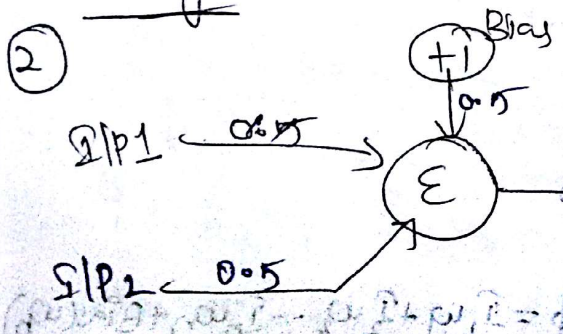
$$\left.\begin{array}{l} 1, \text{ if } \phi \geq 0 \\ 0, \text{ if } \phi < 0 \end{array}\right\} \text{ Threshold Activation.}$$

> firstly we have to give weights randomly. & calculate I, w, t--- Bias(w_b) if our required o/p is not correct then at which i/p our o/p is wrong then we have to add error.

Error = (Expected o/p − Actual o/p)

OR gate

(2)



1) I/p 1 & I/p 2 — Bias — $w_1 = 0.5$, $w_2 = 0.5$, bias = 0.5

0	0	0
0	1	1
1	0	1
1	1	1

i/p 1	i/p 2	o/p	
0	0	1	X

Error = EV − AV

$0 - 1 = -1$

$= 0*(-1) + 0*(-1) + 1(-1)$

$= 0 + 0 - 1$

$\Delta w = w_t + \lambda$ i/p (Error)

$= wt + \lambda$ I/p (Exp − Actual) ✓

0	0	0 ✓	0.5 −0.5 0.5
0	1	0 X	0.5 −0.5 −0.5
			0.5 0.5 0.5

$0*(1) + 0.5 + 1(-1)*0.5 + 1*(1)*0.5$

0	1	1 ✓	
1	0	1 ✓	

1	1	1 ✓	
0	0	1 X	

1, if $\phi \geq 0$
0, if $\phi < 0$
} Threshold Activation

OR Gate

0	0	0
0	1	1
1	0	1
1	1	1

$\Delta W_i = W_i + \lambda \, i/p \, (Error)$

$\Delta Wt = wt + \lambda \, i/p \, (Exp\ val - Act\ val)$

i/p:- 0 0 weights 0.5 0.5 0.5

$= 0*(0.5) + 0*(0.5) + 1*(0.5)$

$= 0 + 0 + 0.5$

$= 0.5 \geq 0$ so o/p 1

0 0 ≠ 1 x so Add error to i/p's

error = Expected val - Actual value

$= 0 - 1$

$= -1$

now

$= 0*(-1)*(0.5) + 0*(-1)*(0.5) + 1*(-1)*(0.5)$

$= (0) + (0)*(-0.5)$

So weights are

0.5 0.5 -0.5

now

$= 0*(0.5) + 0(0.5) + 1(-0.5)$

$= 0 + 0 - 0.5$

$= -0.5 < 0$ so it 00 then o/p is 0

00 0

01 1 —

$= 0(0.5) + 1(0.5) + 1(-0.5)$

$= 0 + 0.5 - 0.5$

$= 0$ - o/p is 1 ✓

> 10

$= 1 \times (0.5) + 0 \times (0.5) + 1 \times (-0.5)$

$= 0.5 + 0 - 0.5$

$= 0$ -o/p is 1

10 1 ✓

> 11

$1 \times (0.5) + 1 \times (0.5) + 1 \times (-0.5)$

$= 0.5 + 0.5 - 0.5$

$= 0.5 \geq 0$ 1

11 1 ✓

ANN training using python

ANN2.py

features — Target decision

ID	Branch	Marks	conduct	sports	placement
	CSE (0)	89	excellent (3)	Yes (0)	✓
	CSE (0)	64	Bad (0)	Yes (01)	X
	ECE (1)	72	Good (2)	No (1)	✓
	MECH (2)	54	satisfactory (1)	Yes (0)	X
	nominal data	numeric data	ordinal	binary	Binary

1) Sklearn

Encoders
- LabelEncoder — for converting text to number.
- One hot encoder — If there is no need of order (ex: Maths, physics, Chemistry).

no-order = nominal data

order = ordinal

1) Load dataset (ie .csv)

2) eliminate unnecessary fields

3) X = features
   y = Target decision

4) encoding

5) Training.

> to remove order in nominal data we use one-hot encoder

i.e

Branch		
CSE	ECE	MECH
1	0	0
1	0	0 X
0	1	0
0	0	1

i/p layer - 6 ∵ Branch - 3
Marks - 1
Conduct - 1
Sports - 1

3+1+1+1 = 6

o/p layer - 1 ∵ placement - 1

hidden layer = $\frac{\text{i/p layer} + \text{o/p layer}}{2}$

anaconda is reference for all files.

A* Deep learning :

Weather dataset ↑ features

0	1	2	3	↑ Target decision
Outlook	Humidity	Temp	Windy	Play
Rainy	number	Number	True	Yes
Sunny			false	No
overcast				

> here label encoding is required when there is requirement of conversion from text to number.
so for outlook, windy, play we need label encoding

> here ~~for target~~ one-hot encoder is required when there is no order means (subjects, outlook)

Rainy	sunny	overcast
1	0	0
0	1	0
0	0	1

so now i/p layer = 6 ⇒ outlook = 3
humidity = 1
Temp = 1
windy = 1

3+1+1+1 = 6

o/p layer = 1 ∵ play

Hidden layer = $\frac{\text{i/p layer} + \text{o/p layer}}{2}$

hidden layer = $\frac{6+1}{2} = 3.5$

13-08-18 (Tuesday)

Textclassification using ANN:

Passage:

Rama was the king of Ayodhya. Rama's wife Sita is Sita. Rama had 2 sons. Rama killed Ravana.

Question:

Who is Rama's wife?

using Nltk (Natural language tool kit) has the code to workout.

Step1: stopWorks Removal: Removing the unnecessary word's like is, the, had etc whos, is

Step2: Lamemetization: Rama Rama's
study studied
study studing.

Step 3:- unique words!

Rama sita
king two sons
Ayodhya killed
wife Ravana

world to vector matrix:-

term frequency vectors	Rama	king	Ayodhya	wife	sita	two	sons	killed	Ravana
1)	1	1	0	0	0	0	0	0	0
2)	0	0	1	1	0	0	0	0	
3)	0	0	0	0	1	1	0	0	
4)	0	0	0	0	0	0	1	1	
Q:-	1	0	0	1	0	0	0	0	

match the Question with above passage lines

match Question with 1st line of passage: 3 are different
match question with 2nd line of passage: 1 is different
match Question with 3rd line of passage: 2 is different
match Question with 4th line of passage: 3 is different.

least different line is the answer. (i.e 1 is different for 2nd line)

Ans:- Rama wife sita :-

This was the very basic. Antonyms & synonyms should be added more to get accuracy

More Reviews passage:

The direction (1) was good. Movie hero acting (2) was awesome. Best comedy. Best comedy (3). Best Movie (4) of the year. Worst experience (5). Music was bad (6).

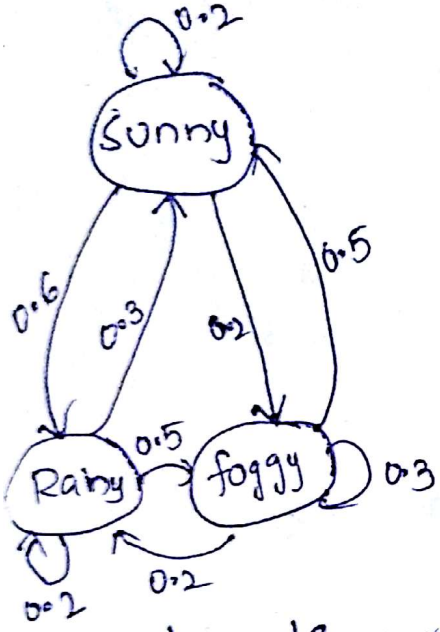
Acting was bad.

Worst movie ever.

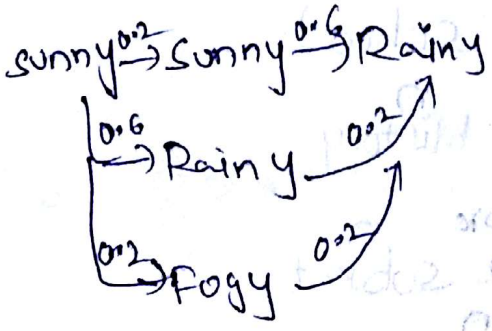
Unique words are: direction, good, Movie, hero, acting, awesome, Best, comedy, year, worst, experience, Music, bad

	direction	good	Movie	hero	acting	awesome	Best	comedy	year	worst	experience	Music	bad
1	1	0	0	0	0	0	0	0	0	0	0	0	0
2	0	1	0	0	0	0	0	0	0	0	0	0	0
3	0	0	1	0	0	0	0	0	0	0	0	0	0
4	0	0	0	1	0	0	0	0	0	0	0	0	0
5	0	0	0	0	1	0	0	0	0	0	0	0	0
6	0	0	0	0	0	1	0	0	0	0	0	0	0
7	0	0	0	0	0	0	1	0	0	0	0	0	0
8	0	0	0	0	0	0	0	1	0	0	0	0	0
9	0	0	0	0	0	0	0	0	1	0	0	0	0
10	0	0	0	0	0	0	0	0	0	1	0	0	0
11	0	0	0	0	0	0	0	0	0	0	1	0	0
12	0	0	0	0	0	0	0	0	0	0	0	1	0
13	0	0	0	0	0	0	0	0	0	0	0	0	1

Hidden Markov Models (parts of speech Toggling in text)



D1 D2 D3



$(0.2 \times 0.6) + (0.6 \times 0.2) + (0.2 \times 0.2)$

> automatic parts of speech is for Markov Model.

(P) To know what is the probability of next day whether ~~sunday~~ sunny, rainy foggy when person brings Umbrella along with launch for the person who is in the ~~dar~~ room

Umbrella probability

sunny - 0.2

foggy - 0.1

rainy - 0.7

today tommorow

sunny - rainy ⇒ umbrella prob ⇒ (0.6×0.7)

Sunny - foggy & " " ⇒ (

Pa

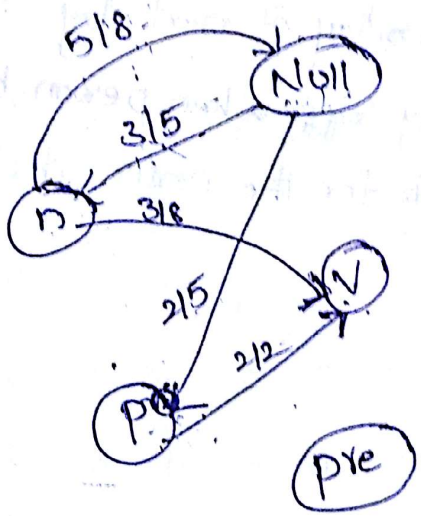
P pro n
Null I go to school.

P v n
Null I remember History.

n v pre n
Null Geography is a subject

n v n
Null school is far

n v n
Null Walking is good



Q) I drink tea

⇒ I → what is the probability of I to be pronoun & noun & verb & preposition from null because (I i

$= \left(\frac{2}{5}\right)\left(\frac{\text{Pronoun}}{I}\right) + \left(\frac{3}{5}\right)\left(\frac{\text{noun}}{I}\right) + \left(\frac{0}{5}\right)\left(\frac{\text{Verb}}{I}\right) + \left(\frac{0}{5}\right)\left(\frac{\text{pre}}{I}\right)$

$= \left(\frac{2}{5}\right)\left(\frac{2}{2}\right) + \left(\frac{3}{5}\right)\left(\frac{0}{2}\right) + \left(\frac{0}{5}\right)\left(\frac{0}{0}\right) + \left(\frac{0}{5}\right)(0)$

$= \left(\frac{2}{5}\right)\left(\frac{2}{2}\right) + 0 + 0 + 0$

$= 2/5$ so I is the ↑ higher probability of being pronoun

⇒ drink

What is the probability of drink as pronoun & noun & verb

preposition from It pronoun)

$$\left(\frac{2}{2}\right)\left(\frac{\text{verb}}{\text{drink}}\right) + \left(\vec{0}\right) + \left(\frac{\text{preposition}}{\text{drink}}\right) + 10)\left(\frac{\text{prep}}{\text{drink}}\right)$$

$$\left(\frac{2}{2}\right)\left(\frac{0}{0}\right) = 0$$

Computational Discourse — NLP

1) pronoun to noun
2) synonyms / Antynongnus.
3) Relation analysis
4) sentimental analysis

27-8-18 (Monday)

ANN Theory

Features of ANN (Artificial Neuron)

1) Robust & Fault tolerant.
2) Flexible
3) Adaptable (Noisy, probabilistic, inconsistent)
4) collective computation

From Biological neuron we designed Artificial neurons.

1) Robust & fault tolerant
2) flexible -
3) Adaptable
4) Collective computation - parallel processing

Comparison of classical VS NN

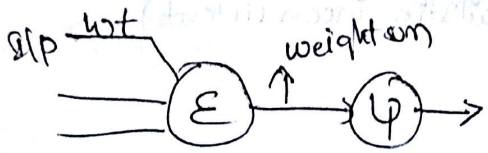
	classical	NN
Speed:	Fast	slow
processing:	sequential	pasallele) (ata million of neurons come into pictoe) - massively
size & complexity:	small & simple	large & complex
storage:	Memory	synaptic weight

	Classical	ANN
Fault tolerance :	less	More
control mechanism :	cpu (central)	Distributed control

ANN Terminology

1) processing unit

- Activation value (Weighted sum)
- Excitatory I/p
- Inhibitory I/p



after adding weights if weights are +ve — Excitatory I/p
" " " " " " -ve — Inhibitory I/p

2) Inter connection :- interconnection b/w I/p & o/p

- I/p state
- o/p state

3) operation.
- Activation dynamics. → total I/p values.

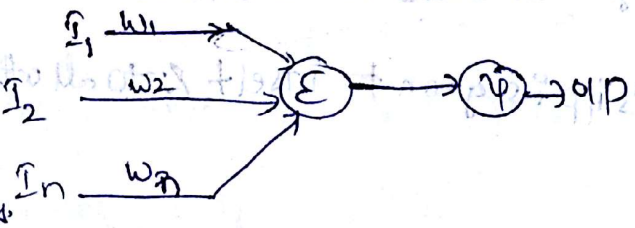
4) update; It says how to modify neuron to get correct o/p
Synchronous/Asynchronous
Deterministic/stochastic.

Changing weights at a time for all neurons — synchronous
Changing weight step by step for all neurons — asynchronous.
dependency of one state to another — Deterministic
opposite of Deterministic — Stochastic

Models of a Neuron

1) Mcculloch - pitts Model

- fixed weighted I/p summation & o/p function.

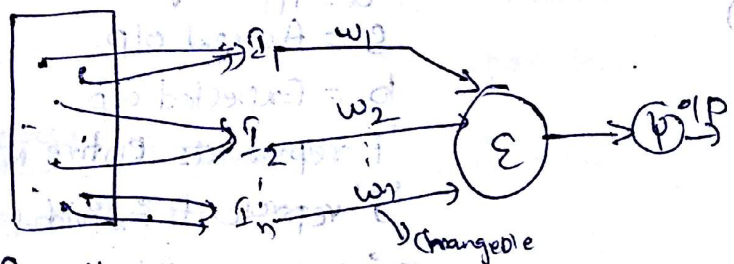


disadvantages

- learning was not possible ∵ fixed weights (which cannot be changed) by that we cannot change to our required o/p.

2) Perceptron Model.

- In a world there is sensory unit (listening, speaking, visible sensing) which are connected to I/p & weights are changeable to change.



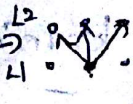
- here classification is only one type that is discrete either yes/no/apple

3) ADALINE Model

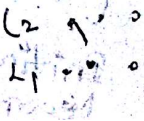
Same as perceptron but they proposed a new model which can calculate continuous values also such as 1.0, 1.2, 1.001, ...

Topology

Geometrical arrangement of neuron.



- Layer 1 to layer 2 — Instar Topology
- Layer 2 to Layer 1 — outstar Topology



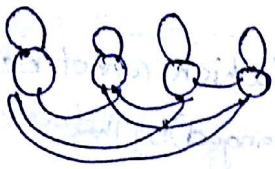
Group of Instar:- group of neurons giving i/p to group of neurons (L1-L2)

Group of outstar :- Same as Group of Instar but (L2-L1)

Bidirectional topology:- If there is no dir. Arrow

Auto Associative :- one's i/p is given to itself & to all other neuron

ex:



Basic learning Laws

$$W(t+1) = w(t) + \Delta w$$

Error = Expected - Actual

$e_i = (b_i - s_i)$

1) Hebbs law

$$\Delta w_{ij} = \eta\, s_i a_j$$

Initial wts ⇒ Near Zero

learning ⇒ unsupervised

η = learning rate constant

a = i/p

S = Actual o/p.

b = Expected o/p

'i' represents Entire NN

'j' represents individual neuron

a_j ⇒ I/p of it's neuron

e_i ⇒ error of entire system

[∵ If there is b_i - Expected o/p means we are saying o/p so it is supervised learning]

2) perceptron learning law

$$\Delta w_{ij} = \eta \cdot (e_i) \cdot (s_i)$$

Initial wts ⇒ Random

learning ⇒ supervised

3) Delta learning

$$\Delta w_{ij} = \eta \cdot (e_i) \cdot \frac{d}{dx}(s_i) \cdot s_i$$

Initial wts ⇒ Random

learning ⇒ supervised

4) Correlation learning

$$\Delta w_{ij} = \eta \cdot b_i \cdot a_j$$

Initial wts ⇒ near zero

learning → supervised

5) Winner-Take-All (Instar) learning

$$\Delta w_{kj} = \eta (a_j - w_{kj})$$

Initial wts ⇒ Random

learning ⇒ unsupervised

6) Outstar learning

$$\Delta w_{kj} = \eta (b - w_{jk})$$

Initial wts ⇒ Random

learning ⇒ supervised.

7) Least Mean Square (LMS) Learning;

$$\Delta w_{kj} = \eta (\text{LMS error}) a_j$$

Initial wts ⇒ Random

learning ⇒ supervised.

LMS error

Case 1: 4unit − 0unit = 4units

Case 2: 4unit − 3unit = 1unit

Case 3: 2unit − 2unit = 0units

here in all cases error is 4units if we go to square it we get correct error

case 1: $3^2 + 1^2 = 10$

case 2: $1^2 + 3^2 = 10$

case 3: $2^2 + 2^2 = 8$ ✓ This having least error in o/p.

> ann2.py → chum modelling tut
- only 1 dataset
↓
weather
diabetis

> Titanic - train & test 2 dataset.

> Question answering system - question - answer

> Text classification - emotions.

(P) hotel problem → service
food
Cleantiness
} Text classification

(P) speaking in phone in different languages according to situation the language should be converted ∴ Text classification

(P) storing info about phone when user ~~ask~~ question it has to give answer ∴ Question Answering system.

(P) season Rainfal Temp Humidity prob soln - Titanic.
?

(P)

Ip layer
No. of Ip Layer
unique words
H/L
Op layer
problem
18 - 18 emotions
3 → service, food, cleanliness - Hotel problems
2 - Respiratory problem
Digestive problem

27-08-Bth
day,

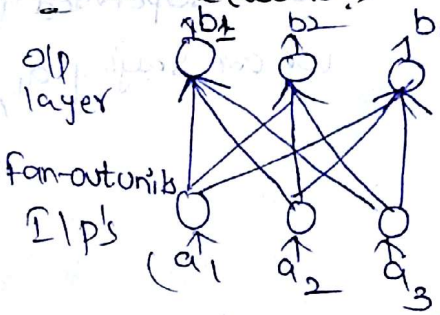
ANN - Theory --- contd

pattern Recognition problem

If new sample is given it has to detect which is done in pattern Recognition problem.

Basic functional units

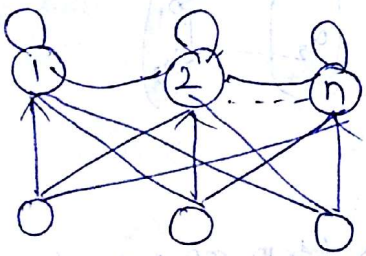
1) Feed forward Network:-



2) Feedback Neural Network

3) competitive learning N/W

Feed forward + feed back Neural



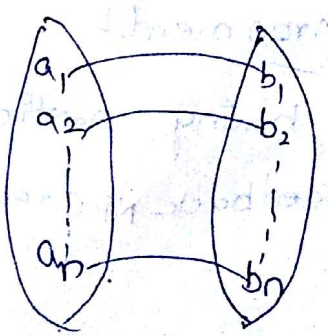
The actions done by these 3 pattern Recognition N/W.

1) Feed forward Network.

a) pattern Association

x	x^2+1
1	2
2	5
3	10

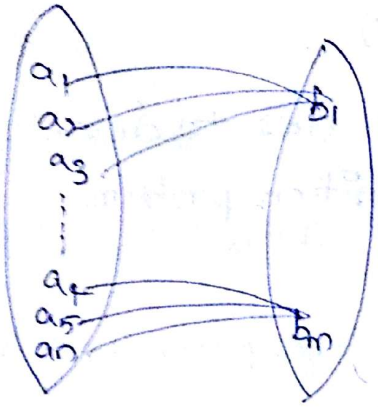
If 1.05 then gives o/p as → 2



ex:- here though sound is not clarity i.e not exactly also it recognizes the correct sound

a1.05 → b1

b) pattern classification

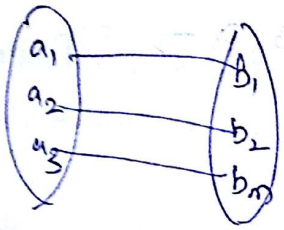


$n \geq m$

though
if apple shape & size is different it recognize it is apple (a_1, a_2, a_3) (b_1)
diff apples → same

here supervised learning becau
we are saying a_1, a_2, a_3 are appl

c) pattern mapping

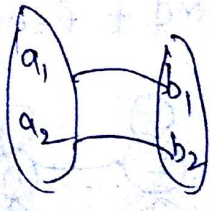


here it gives numbers

2) Feed back Neural network:

a) Auto Association.

here it recognizes itself ie
Cons feed back is required



b) pattern storage:

by Using feedback it stores pattern as same as flipflops storing Electrical signals.

c) Pattern Environment

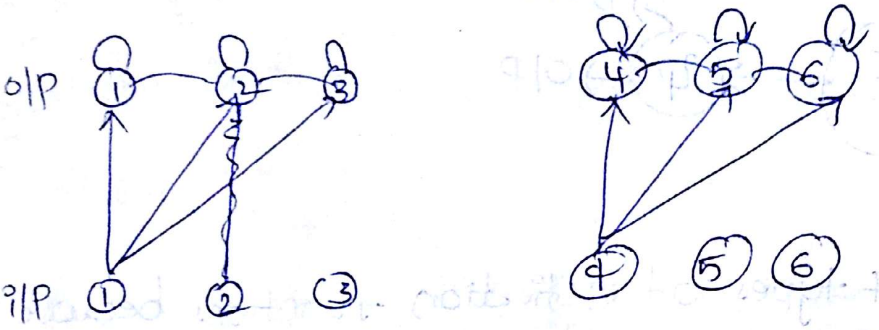
The prob of having particular pattern.

for these also feedback NW required

3) competitive learning NW

a) Temporary pattern storage.

To store pattern temporary we use this



if i/p 1,2,3 o/p also

1,2,3 if i/p changes to 4,5,6 then o/p also 4,5,6 means temporarily at particular time what i/p is present it gives that o/p. i.e

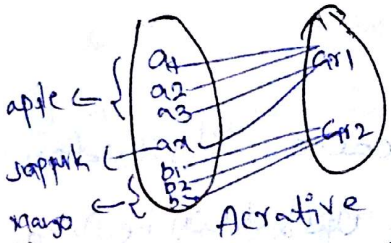
for temporary storage

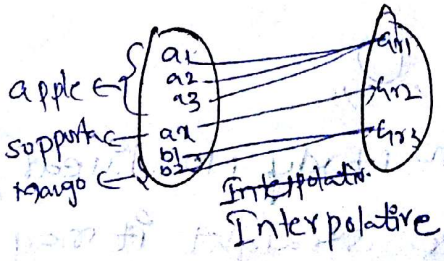
b) pattern clustering :- same as pattern classification but here unsupervised learning we are not giving prior (knowledge)

Acrative :- If 2 apples, 3 mangos are present if supports

Interpolative :- comes than it matches with apples based on shape

↳ If 2apples, 3mangoes are present if supports is matched with separate group



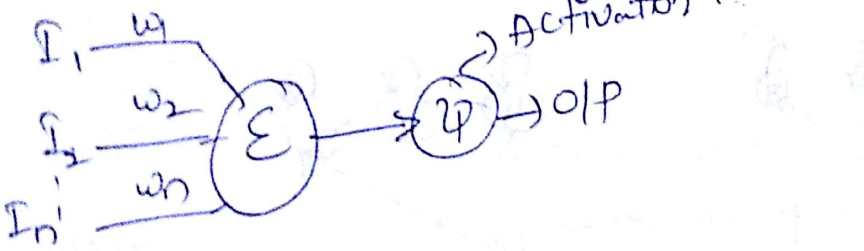


c) Feature Mapping:-

30-08-16 (Tues

Types of Activation functions

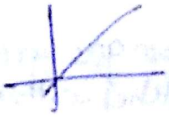
To produce final o/p we use Activation function.



> There is diff types of Activation functions because different projects require different Activation function based on our Activation function project we design

Activation function

1) linear
2) Non linear

$ax+by+c$ = 

$f(x) = x$ Range $-\infty, +\infty$

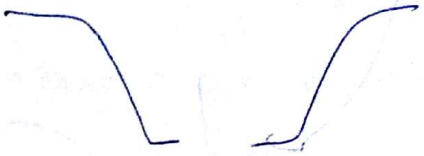
Non-linear

Why we need to study Non-linear, because all situation need not be straight it may be up & down i.e curve

↳ Differentiation:-

the o/p should be Differentiable then error will be decrease

Monotonic



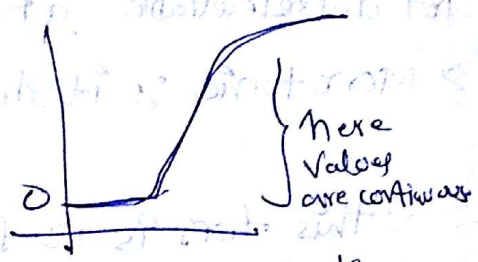
non-Monotonic

Non-linear

1) Sigmoid function

$$f(z) = \frac{1}{1+e^{-z}}$$

Range ⇒ 0 to 1 (continuous)



∴ this function is very useful in probability because it (Non-linear) has continuous values

> it is differentiable ∵ it has slope

> it is monotonic ∵ it is decreasing function.

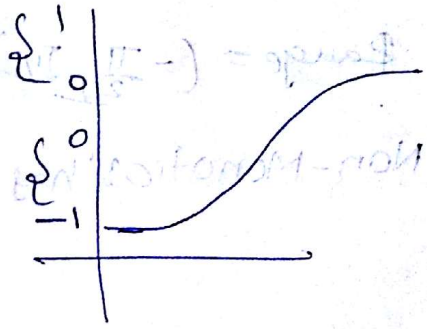
2) Tanh or hyperbolic tangent.

> $f(x) = \tanh(x)$.

> Range = −1 to +1

> Differentiable

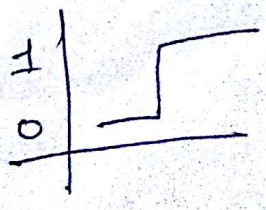
> Monotonic



∵ this is very useful in knowing whether the person is suffering from cancer or not.

3) Threshold function (Binary step)

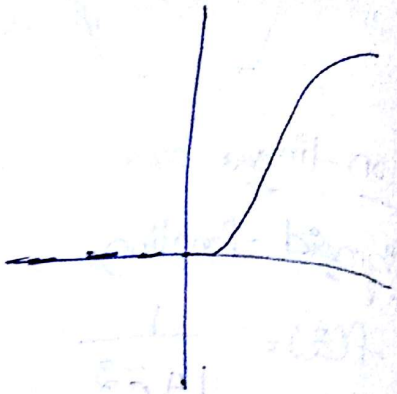
$$f(x) = \begin{cases} 1 \text{ if } x \geq 0 \\ 0 \text{ if } x < 0 \end{cases}$$



It says whether 0 or 1 whereas in sigmoid function from 0 to 1 these are continuous values.

c) ReLU (Rectified Linear unit)

$$f(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } x \geq 0 \end{cases}$$



> Differentiable ∴ has slope

not differentiable at $x=0$ after that

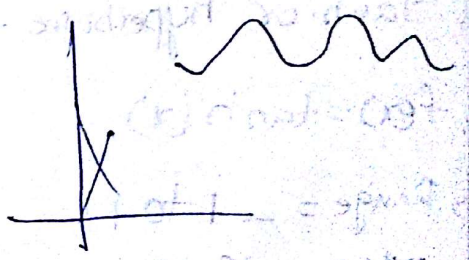
> Monotonic ∴ it decreasing fun

∴ this fun is useful in ~~wh~~ these is no difference in −1 whether test is done/not but in +1 we have to test ~~for~~ the cancer ~~patient~~

b) Arc tan

$f(x) = \tan^{-1}(x)$

Range $= (-\frac{\pi}{2}, \pi/2)$



Non-Monotic ∴ has no slope.

ANN2
Titanic
weather
Questionnaire
Text classification