Chapter-6: Machine Learning

Learning is the process by which a system improve from experience. It is an area of AI that focus on the process of self improvement. A computer program is said to be learnt from experience, performance measure (P). It's performance at task T is measured by P which is improved by experience E. This a learning system is characterised by task(T), experience (E) and performance measure (P).

Types of learning.

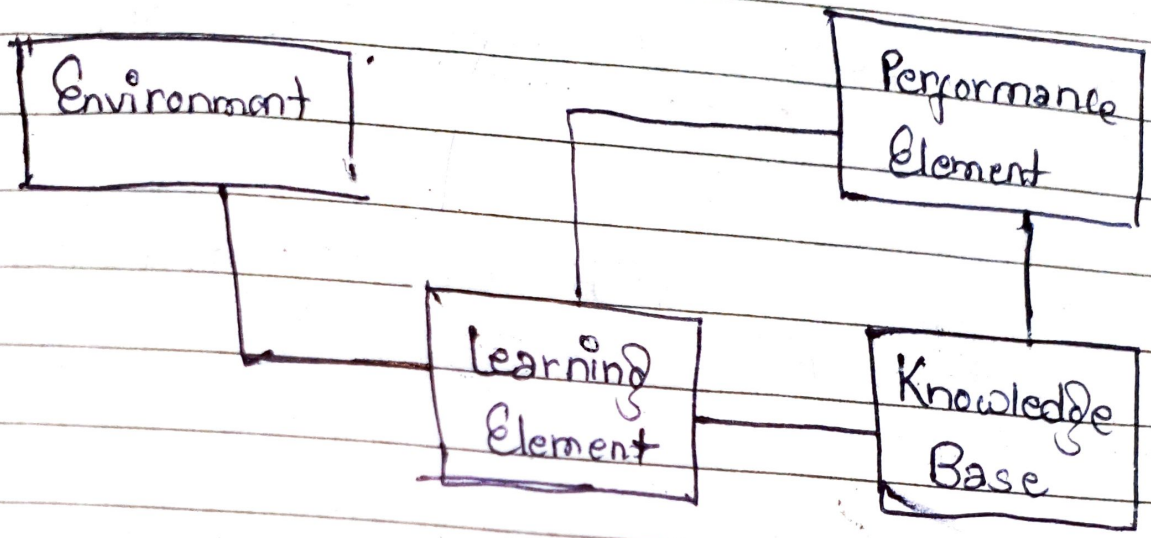
i) Supervise learning:

In this technique, a teacher is available who provides the desired action corresponding to a perception. A set of perception action pair is known as training set.

ii) Unsupervise learning

In this technique there is no label data. The learner only discover some pattern consist of collection of perception. This is also called exploring learning Eg: Finding out malicious n/w attack from a sequence of anolomous data package.

Learning framework:



Environment:

It refers to the name & quality of information given to the learning element. The nature of information depends upon its level: — High level information is abstract & deals with broad class of problem but is complex to design.

- Low-level information is detailed i.e. it deals with single problem only. So, it may not give real

time solution.

- Learning Element:
  It acquire new knowledge through learning elements. Learning may be rote learning, learning by example, explanation based learning.

Knowledge base:
There are three main features of knowledge base:

i) Expressive:
Knowledge should be represented in understandable way.

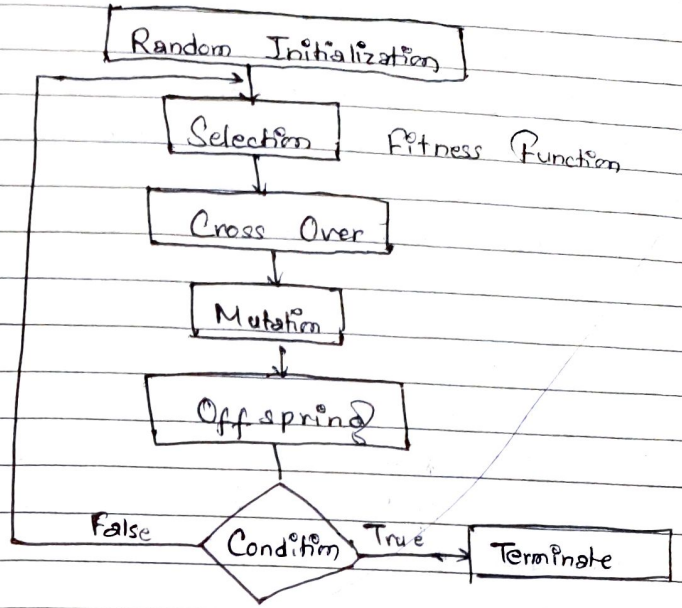
ii) Modifiable:
It must be easy to change the data if required.

) Extendability:
Knowledge should be updated and revised in certain periodic time.

Performance Element:
The performance element analyses how complex the learning is & how learning is being performed. Complexity depends upon type of task. The simplest task is classification based on a single rule while the complex task requires the application of multiple rule in sequence.

The learning element should have access to all internal actions of the performance element.

V. Imp Genetic Algorithm



ADT-2

① Step 1:

i. $\forall x$: Horse(x) → mammal(x)
   $\forall x$: Cow(x) → mammal(x)
   $\forall x$: Pig(x) → mammal(x)
ii. $\forall x \forall y$ Offspring(x, y) ∧ horse(y) → horse(x)
iii. Horse (Bluebeard)
iv. Parent (Blueboard, Charlie)
v. $\forall x \forall y$ Offspring(x, y) ⟷ Parent(y, x)
vi. $\forall x \forall y$: mammal(x) → Parent(y, x).

Step 2: In CNF,

i. a. $\neg Horse(x) \vee mammal(x)$

b. $\neg Cow(x) \vee mammal(x)$

c. $\neg Pig(x) \vee mammal(x)$

ii. $\neg (Offspring(x,y) \wedge horse(y)) \vee horse(x)$

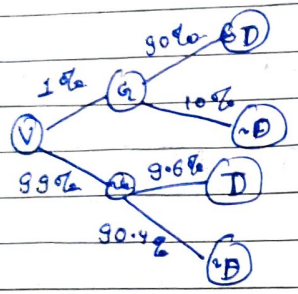
$\equiv \neg Offspring(x,y) \vee \neg horse(y) \vee horse(x)$

iii. Horse (Blueberd)

iv. Parent (Blueberd, Charlie)

v. ------

(2)



$$\text{Answer} = \frac{0.01 \times 0.9}{0.01 \times 0.9 + 0.99 \times 0.096}$$

$$= 0.08$$

Q. $a = abcdefgh$

$\downarrow$

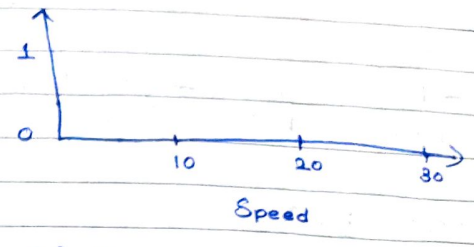
0

$\vdots$

9

$f = (a+b) - (c+d) + (e+f) - (g+h)$

Fuzzy logic:



(1) Fuzzification

(2) Inference Rule

(3) Defuzzification

Fuzzy Operation:

$A = \{(10, 0.1), (20, 0.2), (30, 0.7), (40, 0.6)\}$

$B = \{(10, 0.3), (20, 0.4), (30, 0.5), (40, 0.6)\}$

Union

$A \cup B = \{(10, 0.3), (20, 0.4), (30, 0.7), (40, 0.6)\}$

Intersection

$A \cap B = \{(10, 0.1), (20, 0.2), (30, 0.5), (40, 0.6)\}$

Bold Union:

$A \oplus B = \min(1, m_a + m_b)$

$= \{(10, 0.4), (20, 0.6), (30, 1), (40, 1)\}$

Bold Intersection:

$(A \odot B) = \max(0, m_a + m_b - 1)$

$= \{(10, 0), (20, 0), (30, 0.2), (40, 0.2)\}$

Complement:

$\bar{A} = \{(10, 0.9), (20, 0.8), (30, 0.3), (40, 0.4)\}$

$\bar{B} = \{(10, 0.7), (20, 0.6), (30, 0.5), (40, 0.4)\}$

Iterative Dichotimizer (ID3)

Step:

① Calculate Information Gain

$$I\text{-}G_{before} = \frac{-P}{P+N}\log_2\frac{P}{P+N} - \frac{N}{P+N}\log_2\frac{N}{P+N}$$

② Calculate Entropy of each class.

$$E_{class} = \sum \frac{P_i+N_i}{P+N}\ I\text{-}G_{subclass}$$

$$I\text{-}G_{subclass} = \frac{-P_i}{P_i+N_i}\log_2\left(\frac{P_i}{P_i+N_i}\right) - \frac{N_i}{P_i+N_i}\log_2\frac{N_i}{P_i+N_i}$$

③ Calculate Gain for each class.

$$Gain_{class} = I\text{-}G_{before} - E_{class}.$$

④ Compare all gain then select the root node having highest gain.

⑤ Again start from 1 to 4 for root node. Continue this until all value is satisfied.

E.g.:

Age	Competition	Type	Profit
Old	Yes	Software	Down
Old	No	"	"
Old	N	Hardware	"
Mid	Y	S	"
Mid	Y	H	"
Mid	N	H	Up
Mid	N	S	"
New	Y	S	"
New	N	H	"
New	N	S	"

$$IG_{before} = \frac{-5}{10}\log_2\frac{5}{10} - \frac{5}{10}\log_2\frac{5}{10}$$

$$= 1$$

$$E_{age} = \sum \frac{P_i+N_i}{P+N}\ IG_{subclass}$$

$$I_{old} = \frac{-P_i}{P_i+N_i}\log_2\left(\frac{P_i}{P_i+N_i}\right) - \frac{N_i}{P_i+N_i}\log_2\frac{N_i}{P_i+N_i}$$

$$= \frac{-0}{0+3}\log_2\left(\frac{0}{0+3}\right) - \frac{3}{3}\log_2\left(\frac{3}{3+0}\right)$$

$$= 0$$

$$IG_{mid} = \frac{-2}{2+2}\log_2\left(\frac{2}{2+2}\right) - \frac{2}{2+2}\log_2\left(\frac{2}{2+2}\right)$$

$$= 1$$

$$IG_{new} = 0.$$

$$E_{age} = \frac{0+3}{10}\times 0 + \frac{2+2}{10}\times 1 + \frac{0+3}{10}\times 0$$

$$= \frac{4}{10} = 0.4$$

$$IG_{yes} = \frac{-1}{1+3}\log_2\left(\frac{1}{4}\right) - \frac{3}{1+3}\log_2\left(\frac{3}{4}\right)$$

$$= 0.811$$

$$IG_{new} = \frac{-4}{4+2}\log_2\left(\frac{4}{4+2}\right) - \frac{2}{4+2}\log_2\left(\frac{2}{4+2}\right)$$

$$= 0.918.$$

$$E_{competion} = \frac{1+3}{10}\times 0.811 + \frac{4+2}{10}\times 0.918$$

$$= 0.875$$

PAGE NO. :
DATE :

$$I G_{s.w} = -\frac{3}{3+3}\log_2\left(\frac{1}{2}\right) - \frac{3}{3+3}\log_2\left(\frac{1}{2}\right)$$

$$= 1$$

$$IG_{H/w} = 1$$

$$E_{type} = \frac{3+3}{10} \times 1 + \frac{2+2}{10} \times 1$$

$$= 0.6 + 0.4$$

$$= 1$$

Now,

III :

$$\text{Gain}_{age} = 1 - 0.4 \quad [IG_{before} - E_{class}]$$

$$= 0.6$$

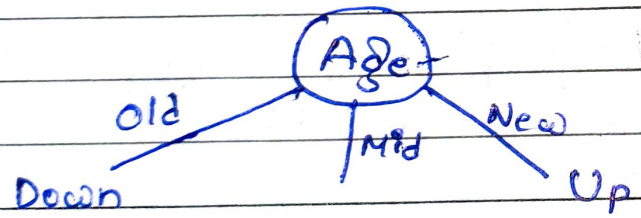
$$\text{Gain}_{competion} = 1 - 0.87$$

$$= 0.13.$$

$$\text{Gain}_{type} = 1 - 1$$

$$= 0.$$

Root node = Age.



Again, $IG_{before} = 1.$