29.10.19

Concept learning:

The problem of searching through a predefined space of potential hypotheses for the hypothesis that best fits the training examples.

Example:

If some instance x satisfies all the constrains of hypothesis h, then h classifies x as a positive example. $(h(x)=1)$. If the hypothesis that ~~aldo~~ Aldo enjoys his favorite sport only on cold days with high humidity is presented by the expression

(?, Cold, High, ?, ?, ?)

—·—

The most general hypothesis, that every day is a positive example, is presented by (?, ?, ?, ?, ?, ?)

And the most specific possible hypothesis that no day is positive example is Represented by $(\emptyset, \emptyset, \emptyset, \emptyset, \emptyset, \emptyset)$

Indicate by a "?" that any value is acceptable of for this attribute.
Indicate by a "$\emptyset$" that no value is acceptable.

	SKy	Air Temp	Humid	Wind	water	fore cast
h	?	cold	High	?	?	?

— • —

The set of items over which the concept is defined is called the set of instances, which is denoted by 'X'

X in this case is the set of all possible bays represented by the ~~ar~~ attributes - SKy, Air Tamp, Humidity, Wind, Water and forecast.

Scanned by CamScanner

The concept or function to be learned is called the target concept, is denoted by "c".

It can be seen as a boolean valued function defined over "X"

$$c : X \rightarrow \{0, 1\}$$

The terget concept corresponds to the value of the attribute EnjoySport

$$c(x) = 1$$

if EnjoySport = yes

$$c(x) = 0$$

if EnjoySport = No

H is ~~be de~~ use to denote the set of all possible hypothesis that the learner may consider regarding the identity of the target concept.

The goal of a learner is to find a hypothesis h which can identify all objects in X.

$h(x) = c(x)$ for all x in X

The symbol D use to denote the set of available training examples.

Prototypical concept learning task:

Given:

⇒ Instance X : Possible days, each describe by the attributes sky, AirTemp, Humidity, Wind, Water, forecast.

⇒ Target concept c :

$$EnjoySport : X \rightarrow \{0, 1\}$$

⇒ Hypotheses H :

Each hypothesis is described by a conjunction of the attributes given ~~def~~ before. the constraints may be "?" "ϕ" or a spacific value.

$h \rightarrow$?, cold, High, ?, ?, ?

⇒ Training examples D:
positive and negative examples of the target function.

Determine:

A hypothesis h in H with
$h(x) = c(x)$ for all x in X

General to specific ordering of hypothesis:

Many machine learning algorithms rely on the concept General-to-specific ordering of hypothesis.

$h_1 = (\text{true}, \text{true}, ?, ?, ?, ?)$

$h_2 = (\text{true}, ?, ?, ?, ?, ?)$

Any instance classified by h1 will also be classified by h2. So we can say that h2 is more general than h1.

Consider the given attributes

$h_1 = ($ sunny, ?, ?, strong, ?, ? $)$

$h_2 = ($ sunny, ?, ?, ?, ?, ? $)$

using this concept we can find a general hypothesis that can be defined over entire data set X.

Find-S

To find out a single hypothesis defined ~~of~~ on X we can use the concept more-general-than partial ordering. It finds the most specific hypothesis and it considers only the positive example. And it will eliminate all the negative examples.

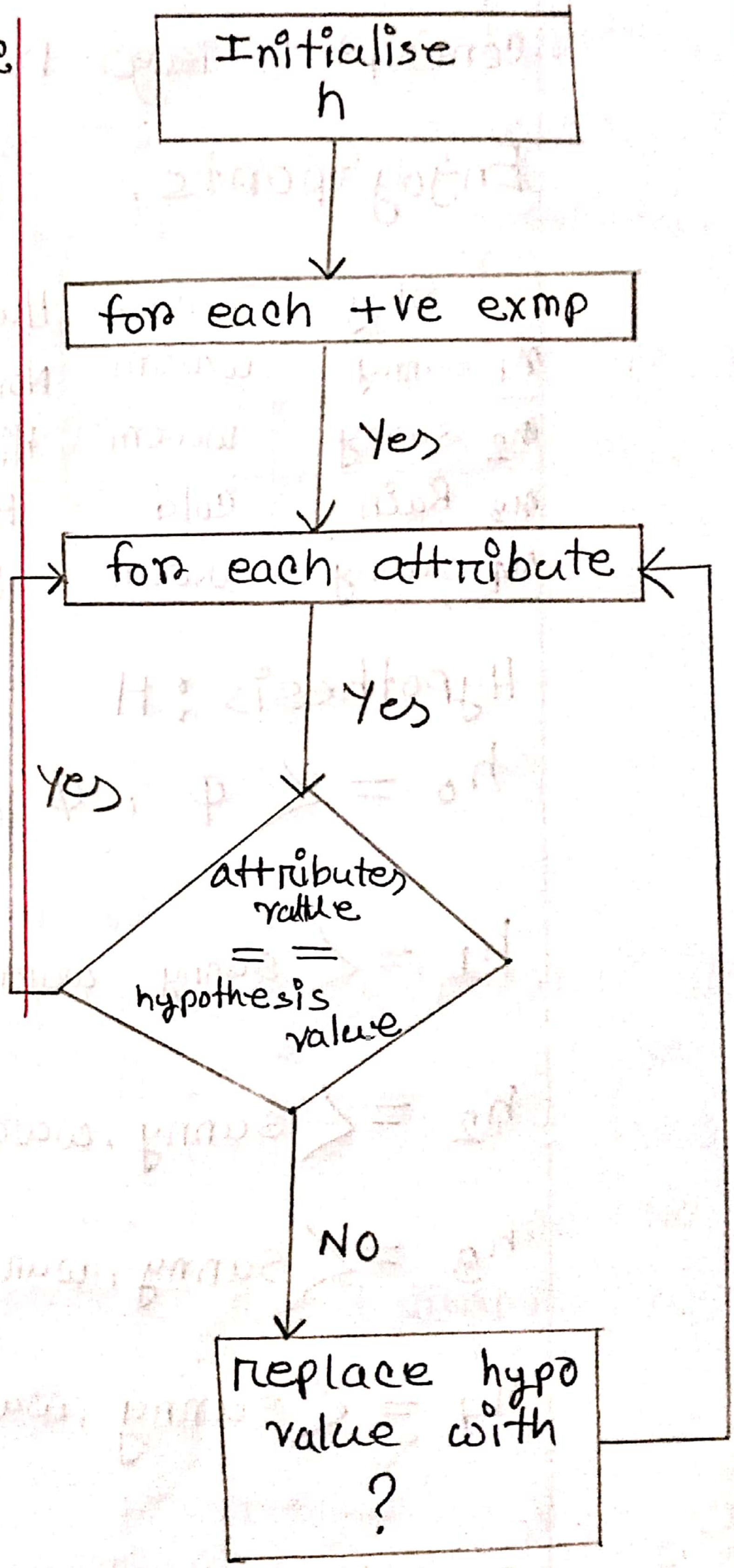
Find-S ~~alg~~ algorithm's limitations can be removed by candidate elimination algorithm.

Find-S Algorithm:

1. Initialize h to the most specific hypothesis in H.

2. For each positive instance x
   - For each attribute ai in h
     * if the constrain ai in h is satisfied by x.
     * Then do nothing
     * Else replace that by "?"

3. Output hypothesis h.



Hypothesis space search by Find-S

Dataset :

Concept — Days on which Person EnjoySports.

	Sky	Temp.	Humidity	wind	water	forcast	Enjoy
x_1	Sunny	warm	Normal	strong	warm	Same	Yes
x_2	sunny	warm	High	strong	warm	Same	Yes
x_3	Rain	Cold	High	strong	warm	Change	NO
x_4	Sunny	warm	High	strong	cool	Change	Yes

Hypothesis : H

$h_0 = \langle \phi, \phi, \phi, \phi, \phi, \phi \rangle$

$h_1 = \langle$ sunny, warm, normal, strong, warm, same $\rangle$

$h_2 = \langle$ sunny, warm, ?, strong, warm, same $\rangle$

$h_3 = \langle$ sunny, warm, ?, strong, warm, same $\rangle$

$h_4 = \langle$ sunny, warm, ?, strong, ?, ? $\rangle$

So this is the hypothesis ~~be~~ depends on the dataset.

Version Space

The candidate elimination algorithm represents the set of all ~~hyhe~~ hypothesis consistent with the training examples.

The subset of all hypothesis is called the version space with respect to the hypothesis space H, and the training examples D. Denoted by $VS_{H.D}$.

Example of version Space :

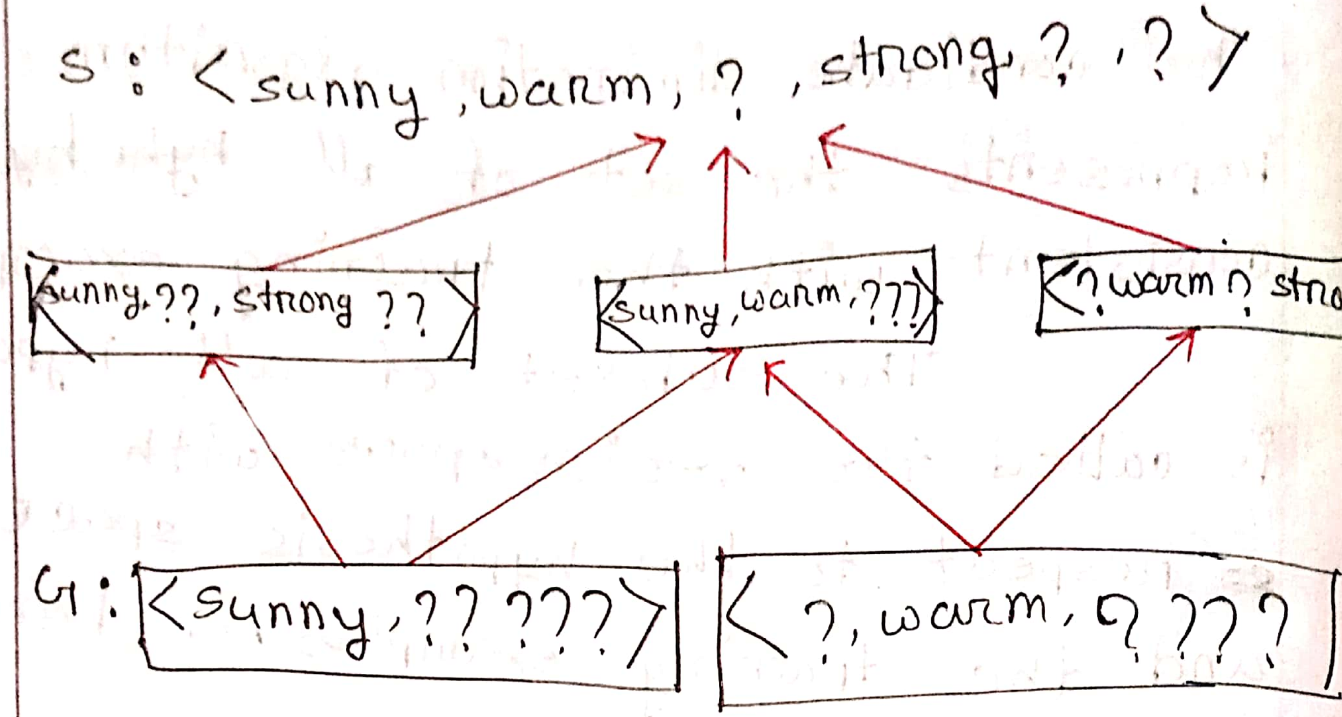
Sky	Temp	Humid	Wind	Water	Forecast	Enjoy Spt
sunny	warm	Normal	strong	warm	same	yes
sunny	warm	High	strong	warm	same	yes
Rainy	cold	High	strong	warm	change	NO
sunny	warm	High	strong	cool	change	yes

① S: <sunny, warm, ?, strong, ?, ?>

<sunny, ??, strong ??>   <sunny, warm, ???>   <? warm ? strong, ??>

G: <sunny ?????>, <?, warm, ? ? ??>

Step 2 :



General boundry: General boundry "G" of version space VS_{HD}, is the ~~ta~~ set of it's maximally general members that are consistant with the given training set.

Specific boundry: specific boundry "S" of version space VS_{HD}, is the set of its maximally specific members that are consistent with the given training set.

Scanned by CamScanner

Every member of the version space lies between these boundaries.

Candidate Elimination Alg.

The candidate Elimination algo is used ~~for~~ remove the limitations of Find-s algo.

Example trace :

SKY	Temp	Humidity	wind	water	forecast	Enjoy
Sunny	warm	Normal	strong	warm	same	+ve
Sunny	warm	High	strong	warm	same	+ve
Rain	cold	High	strong	warm	change	-ve
Sunny	warm	High	strong	cool	change	+ve

step 1 :

S_0 : $\langle \phi, \phi, \phi, \phi, \phi, \phi \rangle$

G_0 : $\langle ? ? ? ? ? ? \rangle$

step 2 :

S1 : $\langle$sunny, warm, Normal, strong, warm, same$\rangle$

G1 : $\langle ? ? ? ? ? ? \rangle$

Trainig example :

x_1 —

sunny, warm, Normal, strong, warm, same$\rangle$

yes

Step 3 :

S_1 : < sunny , warm , Normal, strong , warm , same>

S_2 : < sunny, warm , ? , strong , warm, same>

G_2 : < ? ? ? ? ? ? >

G_1 : < ? ? ? ? ? ? >

Trainig example :

x_2 : <sunny , warm, High, strong, warm, same ,> yes.

Step 4 :

S_2 : < sunny warm ? strong warm same >

S_3 : < sunny warm ? strong warm same >

G_3 : { <sunny ? ? ? ? ? > < ? warm ? ? ? ? > < ? ? ? ? ? same > }

G_2 : < ? ? ? ? ? >

Training example :

x_3 : < Rainy cold High strong warm change > No

Step 5 :

S3 : <sunny warm ? strong warm same>

S4 : <sunny warm ? strong ? ?>

G2 : {<sunny ? ? ? ? ?> <? warm ? ? ? ?>}

G3 : {<sunny ? ? ? ? ?> <? warm ? ? ? ?> <? ? ? ? ? same>}

Traing example :

x4 : <sunny warm High strong cool, change> yes

Step 5 : resulting version space

S : <sunny warm ? strong ? ?>

$\Downarrow$ $\Downarrow$

<sunny ~~??~~ warm ????> <sunny ??stong ??> <? warm ? strong ??>

$\Uparrow$ $\Uparrow$

G : <sunny ?????> <? warm ????>

Limitations of Find-S Alg:

15.10.19

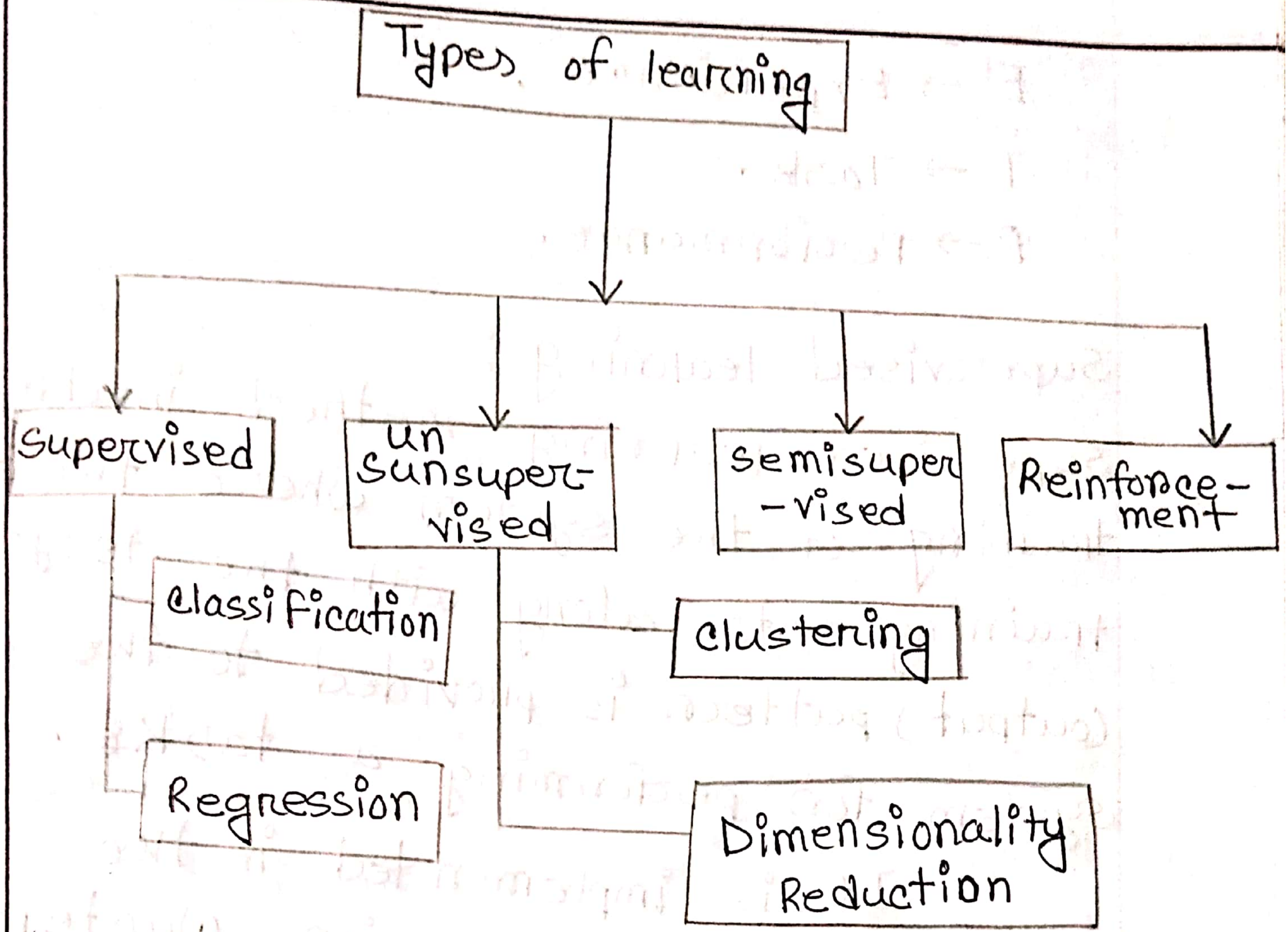
what is Machine learning ?

Machine learning is an application of AI that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. It can access data and use it learn for themselves.

ML is to allow computer systems to learn from experience. ~~without being~~

Machine learning algorithms are three ~~types of~~ categorized :

1. Supervised learning.
2. Un-supervised learning.
3. Reinforcement learning.
4. Semi-supervised learning.



Machine learning:

A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if it is performance tasks in T, as measured by P, improves with experience E.

E → Experience.

T → Task.

P → Performance.

Supervised learning:

Supervised learning method involves the training of the system where the training sets along with the terget (output) pattern is provided to the system for performing a tasks.

It is implemented in the machine learning Regression, cluster and neural networks.

unsupervised learning:

unsupervised learning model does not involve the target output which means no training is provided to the system. The system has to learn by it's own through determining and adapting according to the structural characteristics in the input patterns. It draw conclusions on unlabeled data.

Reinforcement learning:

Reinforcement learning is also an area of machine learning based on the concept of behavioral phychology that works on interacting directly with an environment.

Reinforcement learning tasks on exploitation or exploration.

Difference between superrised and un-supervised learning :

Superrised	un-superrised
1. Superrised learning technique deals with the labelled data.	1. un-supervised learning technique deals with unlabelled data.
2. The output data patterns are known to the system.	2. The output data patterns are just based on the collection of perceptions.
3. Supervised learning concept is less complex.	3. un-supervised learning method is more complecated.
4. The supervised learning can conduct offline analysis	4. The un-supervised learning employs real-time analysis.
5. classiffication and Regression	5. un-supervised learning inclules

are the types of problems solved under supervised learning method	clustering and associative rule mining problems.
6. The outcome of the supervised learning technique is more accurate and reliable.	6. unsupervised learning generates moderate but reliable results

Difference between supervised and Reinforcement learning.

Supervised	Reinforcement
1. Supervised learning technique deals with the labeld data.	1. It works on interacting with the invironment
2. It highly supervised.	2. less supervised and depends on the learning agent.

Supervised	Reinforcement
3. Many ~~an~~ algorithms are exist in using this learning.	3. Neither supervised nor un-supervised algorithms are used.
4. Assets are depreciable	4. Liabilities are non-depreciable.
5. Runs on any platform or with any applications.	5. Run with any ~~new~~ hardware or software devices.

Diffrence between classification and Regression :

Classification	Regression
1. Classification is the process of finding a model which helps in ~~sa~~ separating the data into multipule classes.	~~8~~ 1. Regression is the process of finding a model for distinguishing the data into continuous real values instead of using classes.

Classification	Regression
2. Nature of the predicted data ~~in~~ is unordered.	2. Nature of the predicted data is ordered.
3. Mapping function is used for mapping of values to predefined classes.	3. Mapping function is used for mapping of values to continuous output.
4. Involves prediction of discrete values.	4. Involves prediction of continuous values
5. Method of calculation by measuring accuracy	5. Method of calculation by measurement of root mean square error

22.10.19

Formal Def. of Reinforcement :

Reinforcement learning refers to goal-oriented algorithms, which learn how to attain a complex objective or maximize along a particular dimension over many steps. This learning can be understand using the following concepts.

- Agents
- Action (A)
- Discount factor
- Environment
- State (S)
- Reward (R)
- Policy ~~(π)~~ (π)
- Value (V)
- Q value or action value (Q)
- Trajectory.

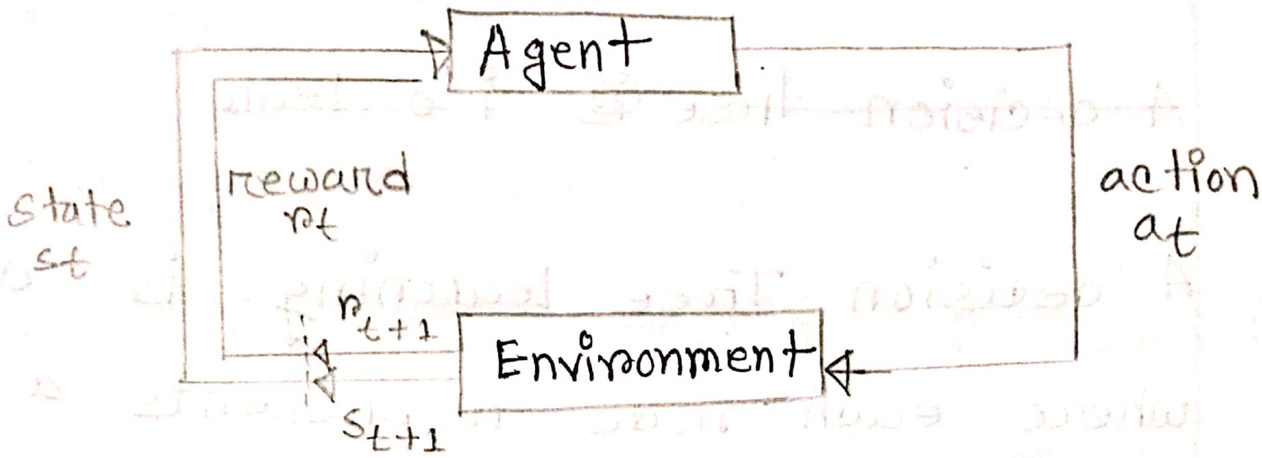


fig: Reinforcement.

— o —

03.11.19

Decision Tree

~~A decision tree is a lean~~

A decision Tree learning is a tree where each node represents a feature (attribute), each link (branch) represents a decision and each leaf represents an outcome.

There are couple of algorithms to build a decision tree — such as
- Assistant
- C4.5
- CART (classification and Regression trees)
- ID3 (Iterative Dichotomiser 3)

ID3 :

ID3 ~~at~~ is an algorithm used to generate a decision tree from a dataset. ID3 is a procursor of C4.5 algorithm.

Weather Dataset:

Day	outlook	Temp.	Humidity	Wind	Play
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	strong	Yes
D12	overcast	Mild	High	strong	Yes
D13	overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	strong	No

~~ID3~~ ~~App~~

Applying ID3 Algorithm —

Output = entropy

information gain = total entropy

Play = 9 yes and 5 No

$Entropy_{total} = [9+, 5-]$

$= -9/14 \log_2 9/14 - 5/14 \log_2 5/14$

$= 0.94$

Values [Wind] = weak (8), strong (6)

$S_{total} = 9+, 5-$

$S_{weak} = 6+, 2-$

$S_{strong} = 3+, 3-$

$E_{weak} = -6/8 \log_2 6/8 - 2/8 \log_2 2/8$

$= 0.811$

$E_{strong} = -3/6 \log_2 \frac{3}{6} - 3/6 \log_2 3/6$

$= 1$

Information gain of wind —

$= E_{total} - 8/14 \times E_{weak} - 6/14 \, E_{strong}$

$= .94 - 8/14 \times .811 - 6/14 \times 1$

$= .94 - .463 - .42$

$= .940 - .883$

$= \cancel{.048} \; .057$

Values [Humidity] = High (7), Normal (7)

$S_{total} = 9+, 5-$

$S_{high} = 3+, 4-$

$S_{Normal} = 6+, 1-$

$E_{High} = -3/7 \log_2 \frac{3}{7} - 4/7 \log_2 4/7$

$= 0.985$

$E_{Normal} = -6/7 \log_2 6/7 - 1/7 \log_2 1/7$

$= .59$

Information gain of Humidity —

$E_{total} - 7/14 \times E_{High} - 7/14 \times E_{Normal}$

$= '94 - 0'7875$

$= 0'1525$

Values [Temp.] = Hot (4), Mild (6), Cool (4)

$S_{total} = 9+, 5-$

$S_{Hot} = 2+, 2-$

$S_{Mild} = 4+, 2-$

$S_{cool} = 3+, 1-$

$E_{Hot} = -2/4 \log_2 2/4 - 2/4 \log_2 2/4$

$= 1$

$E_{Mild} = -4/6 \log_2 4/6 - 2/6 \log_2 2/6$

$= '917$

$E_{Cool} = -3/4 \log_2 3/4 - 1/4 \log_2 1/4$

$= '8112$

Gain of Temp. —

$E_{total} - 4/14 \times E_{Hot} - 6/14\, E_{Mild} - 4/14 \times E_{cool}$

$= .94 - 4/14 \times 1 - 6/14 - .917 - 4/14 \times .8112$

$= .94 - .2857 - .393 - .231$

$= 0.0303$

Values [Outlook] = Sunny (5), Overcast (4)
Rain (5)

$S_{total} = 9+, 5-$

$S_{sunny} = 2+, 3-$

$S_{overcast} = 4+$

$S_{Rain} = 3+, 2-$

$E_{sunny} = -2/5 \log_2 2/5 - 3/5 \log_2 3/5$

$= .97$

$E_{Overcast} = -4/4 \log_2 4/4 - 0/4 \log_2 0/4$

$= 0$

$E_{Rain} = -3/5 \log_2 3/5 - 2/5 \log_2 2/5$

$= .97$

Gain of Outlook –

$E_{total} - 5/14 \times E_{sunny} - 4/14 E_{overcast} - 5/14 \times E_{Rain}$

$= .94 - 5/14 \times .97 - 4/14 \times 0 - 5/14 \times .97$

$= .94 - .35 - 0 - 0.35$

$= .94 - .70$

$= 0.24$

–.–

Scanned by CamScanner