MODULE 2 – DECISION TREE LEARNING

1. What is decision tree and decision tree learning?
2. Explain representation of decision tree with example.
3. What are appropriate problems for Decision tree learning?
4. Explain the concepts of Entropy and Information gain.
5. Describe the ID3 algorithm for decision tree learning with example
6. Give Decision trees to represent the Boolean Functions:
   a) A && ~ B
   b) A V [B && C]
   c) A XOR B
   d) [A&&B] V [C&&D]
7. Give Decision trees for the following set of training examples

Day	Outlook	Temperature	Humidity	Wind	PlayTennis
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

8. Consider the following set of training examples.
   a) What is the entropy of this collection of training example with respect to the target function classification?
   b) What is the information gain of a_2 relative to these training examples?

Instance	Classification	a_1	a_2
1	+	T	T
2	+	T	T
3	-	T	F
4	+	F	F
5	-	F	T
6	-	F	T

9. Identify the entropy, information gain and draw the decision trees for the following set of training examples

Gender	Car ownership	Travel cost	Income Level	Transportation (Class)
Male	0	Cheap	Low	Bus
Male	1	Cheap	Medium	Bus
Female	1	Cheap	Medium	Train
Female	0	Cheap	Low	Bus
Male	1	Cheap	Medium	Bus
Male	0	Standard	Medium	Train
Female	1	Standard	Medium	Train
Female	1	Expensive	High	Car
Male	2	Expensive	Medium	Car
Female	2	Expensive	High	Car

10. Discuss Hypothesis Space Search in Decision tree Learning.
11. Discuss Inductive Bias in Decision Tree Learning.
12. What are Restriction Biases and Preference Biases and differentiate between them.
13. Write a note on Occam's razor and minimum description principal.
14. What are issues in learning decision trees

MODULE 4 –BAYESIAN LEARNING

1. Define Bayesian theorem? What is the relevance and features of Bayesian theorem? Explain the practical difficulties of Bayesian theorem.
2. Define is Maximum a Posteriori (MAP) Maximum Likelihood (ML) Hypothesis. Derive the relation for h_{MAP} and h_{ML} using Bayesian theorem.
3. Consider a medical diagnosis problem in which there are two alternative hypotheses: 1. that the patient has a particular form of cancer (+) and 2. That the patient does not (-). A patient takes a lab test and the result comes back positive. The test returns a correct positive result in only 98% of the cases in which the disease is actually present, and a correct negative result in only 97% of the cases in which the disease is not present. Furthermore, .008 of the entire population have this cancer. Determine whether the patient has Cancer or not using MAP hypothesis.
4. Explain Brute force Bayes Concept Learning
5. What are Consistent Learners?
6. Discuss Maximum Likelihood and Least Square Error Hypothesis
7. Describe Maximum Likelihood Hypothesis for predicting probabilities.
8. Explain the Gradient Search to Maximize Likelihood in a Neural Net
9. Describe the concept of MDL. Obtain the equation for h_{MDL}
10. Explain Naïve Bayes Classifier with an Example
11. What are Bayesian Belief nets? Where are they used?
12. Explain Bayesian belief network and conditional independence with example
13. Explain Gradient Ascent Training of Bayesian Networks
14. Explain the concept of EM Algorithm. Discuss what are Gaussian Mixtures

For More Question Papers Visit - www.pediawikiblog.com

Seventh Semester B.E. Degree Examination, Dec.2018/Jan.2019
Machine Learning

Time: 3 hrs. Max. Marks: 80

Note: ***Answer FIVE full questions, choosing ONE full question from each module.***

Module-1

1 a. Specify the learning task for 'A checkers learning problem'. **(03 Marks)**

b. Discuss the following with respect to the above,

(i) Choosing the training experience.

(ii) Choosing the target function and

(iii) Choosing a function approximation algorithm. **(09 Marks)**

c. Comment on the issues in machine learning. **(04 Marks)**

OR

2 a. Write candidate elimination algorithm. Apply the algorithm to obtain the final version space for the training example. **(10 Marks)**

Sl. No.	Sky	Air temp	Humidity	Wind	Water	Forecast	Enjoy sport
1	Sunny	Warm	Normal	Strong	Warm	Same	Yes
2	Sunny	Warm	High	Strong	Warm	Same	Yes
3	Rainy	Cold	High	Strong	Warm	Change	No
4	Sunny	Warm	High	Strong	Cool	Change	Yes

b. Discuss about an unbiased Learner. **(06 Marks)**

Module-2

3 a. What is a decision tree & discuss the use of decision tree for classification purpose with an example. **(08 Marks)**

b. Write and explain decision tree for the following transactions: **(08 Marks)**

Tid	Refund	Martial status	Taxable Income	Cheat
1	Yes	Single	125 K	No
2	No	Married	100 K	No
3	No	Single	70 K	No
4	Yes	Married	120 K	No
5	No	Divorced	95 K	Yes
6	No	Married	60 K	No
7	Yes	Divorced	220 K	No
8	No	Single	85 K	Yes
9	No	Married	75 K	No
10	No	Single	90 K	Yes

OR

4 a. For the transactions shown in the table compute the following :

(i) Entropy of the collection of transaction records of the table with respect to classification.

(ii) What are the information gain of a_1 and a_2 relative to the transactions of the table? **(08 Marks)**

Instance	1	2	3	4	5	6	7	8	9
a_1	T	T	T	F	F	F	F	T	F
a_2	T	T	F	F	T	T	F	F	T
Target class	+	+	_	+	_	_	_	+	_

b. Discuss the decision learning algorithm. **(04 Marks)**

c. List the issues of decision tree learning. **(04 Marks)**

For More Question Papers Visit - www.pediawikiblog.com

For More Question Papers Visit - www.pediawikiblog.com

Module-3

5 a. Draw the perceptron network with the notation. Derive an equation of gradient descent rule to minimize the error. **(08 Marks)**

b. Explain the importance of the terms : (i) Hidden layer (ii) Generalization (iii) Overfitting (iv) Stopping criterion **(08 Marks)**

OR

6 a. Discuss the application of Neural network which is used for learning to steer an autonomous vehicle. **(06 Marks)**

b. Write an algorithm for back propagation algorithm which uses stochastic gradient descent method. Comment on the effect of adding momentum to the network. **(10 Marks)**

Module-4

7 a. What is Bayes theorem and maximum posterior hypothesis? **(04 Marks)**

b. Derive an equation for MAP hypothesis using Bayes theorem. **(04 Marks)**

c. Consider a football game between two rival teams: Team 0 and Team 1. Suppose Team 0 wins 95% of the time and Team 1 wins the remaining matches. Among the games won by team 0, only 30% of them come from playing on teams 1's football field. On the otherhand, 75% of the victories for team 1 are obtained while playing at home. If team 1 is to host the next match between the two teams, which team will most likely emerge as the winner? **(08 Marks)**

OR

8 a. Describe Brute-force MAP learning algorithm. **(04 Marks)**

b. Discuss the Naïve Bayees classifier. **(04 Marks)**

c. The following table gives data set about stolen vehicles. Using Naïve bayes classifier classify the new data (Red, SUV, Domestic) **(08 Marks)**

Table

Color	Type	Origin	Stolen
Red	Sports	Domestic	Yes
Red	Sports	Domestic	No
Red	Sports	Domestic	Yes
Yellow	Sports	Domestic	No
Yellow	Sports	Imported	Yes
Yellow	SUV	Imported	No
Yellow	SUV	Imported	Yes
Yellow	SUV	Domestic	No
Red	SUV	Imported	No
Red	Sports	Imported	Yes

Module-5

9 a. Write short notes on the following:

(i) Estimating Hypothesis accuracy.

(ii) Binomial distribution. **(08 Marks)**

b. Discuss the method of comparing two algorithms. Justify with paired to tests method. **(08 Marks)**

OR

10 a. Discuss the K-nearest neighbor language. **(04 Marks)**

b. Discuss locally weighted Regression. **(04 Marks)**

c. Discuss the learning tasks and Q learning in the context of reinforcement learning. **(08 Marks)**

For More Question Papers Visit - www.pediawikiblog.com

MODULE 1 – INTRODUCTION AND CONCEPT LEARNING

1. Define Machine Learning. Explain with examples why machine learning is important.
2. Discuss some applications of machine learning with examples.
3. Explain how some disciplines have influenced the machine learning.
4. What is well- posed learning problems.
5. Describe the following problems with respect to Tasks, Performance and Experience:
    a. A Checkers learning problem
    b. A Handwritten recognition learning problem
    c. A Robot driving learning problem
6. Explain the steps in designing a learning systems in detail.
7. Explain different perspective and issues in machine learning.
8. Define concept learning and discuss with example.
9. Explain the General-to-Specific Ordering of Hypotheses
10. Write FIND-S algorithm and explain with example given below

Example	Sky	AirTemp	Humidity	Wind	Water	Forecast	EnjoySport
1	Sunny	Warm	Normal	Strong	Warm	Same	Yes
2	Sunny	Warm	High	Strong	Warm	Same	Yes
3	Rainy	Cold	High	Strong	Warm	Change	No
4	Sunny	Warm	High	Strong	Cool	Change	Yes

11. What are the key properties and complaints of FIND-S algorithm?
12. Define Consistent Hypothesis and Version Space.
13. Write LIST-THEN-ELIMINATE algorithm.
14. Write the candidate elimination algorithm and illustrate with example
15. Write the final version space for the below mentioned training examples using candidate elimination algorithm.

Example – 1:

Origin	Manufacturer	Color	Decade	Type	Example Type
Japan	Honda	Blue	1980	Economy	Positive
Japan	Toyota	Green	1970	Sports	Negative
Japan	Toyota	Blue	1990	Economy	Positive
USA	Chrysler	Red	1980	Economy	Negative
Japan	Honda	White	1980	Economy	Positive
Japan	Toyota	Green	1980	Economy	Positive
Japan	Honda	Red	1990	Economy	Negative

Example – 2:

Size	Color	Shape	Class
Big	Red	Circle	No
Small	Red	Triangle	No
Small	Red	Circle	Yes
Big	Blue	Circle	No
Small	Blue	Circle	Yes

16. Explain in detail the Inductive Bias of Candidate Elimination algorithm.

MODULE 3 – ARTIFICIAL NEURAL NETWORKS

1. What is Artificial Neural Network?
2. Explain appropriate problem for Neural Network Learning with its characteristics.
3. Explain the concept of a Perceptron with a neat diagram.
4. Explain the single perceptron with its learning algorithm.
5. How a single perceptron can be used to represent the Boolean functions such as AND, OR
6. Design a two-input perceptron that implements the boolean function A Λ $\neg$ B. Design a two-layer network of perceptron's that implements A XOR B.
7. Consider two perceptrons defined by the threshold expression $w_0 + w_1x_1 + w_2x_2 > 0$. Perceptron A has weight values

   $w_0 = 1, w_1=2, w_2=1$

   and perceptron B has the weight values

   $w_0 = 0, w_1=2, w_2=1$

   True or false? Perceptron A is more-general than perceptron B.
8. Write a note on (i) Perceptron Training Rule (ii) Gradient Descent and Delta Rule
9. Write Gradient Descent algorithm for training a linear unit.
10. Derive the Gradient Descent Rule
11. Write Stochastic Gradient Descent algorithm for training a linear unit.
12. Differentiate between Gradient Descent and Stochastic Gradient Descent
13. Write Stochastic Gradient Descent version of the Back Propagation algorithm for feedforward networks containing two layers of sigmoid units.
14. Derive the Back Propagation Rule
15. Explain the followings w.r.t Back Propagation algorithm
    - Convergence and Local Minima
    - Representational Power of Feedforward Networks
    - Hypothesis Space Search and Inductive Bias
    - Hidden Layer Representations
    - Generalization, Overfitting, and Stopping Criterion

MODULE 2

1. Give Decision trees for the following set of training examples

Day	*Outlook*	*Temperature*	*Humidity*	*Wind*	*PlayTennis*
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

Solution :-

* $Entropy(S) \equiv -P_{\oplus} \log_2 P_{\oplus} - P_{\ominus} \log_2 P_{\ominus}$

$$Gain(S, A) \equiv Entropy(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} Entropy(S_v)$$

* Note,

(1) when all members of S belong to the same class then, $Entropy(S) = 0$

(2) If S contains an equal number of positive and negative examples then,

$Entropy(S) = 1$

* The first step is to find the topmost node of the decision tree. ID3 determines the information gain for each attribute, then selects the one with highest information gain.

⇒ Entropy of S : positive examples = 09
negative examples = 05

$$\text{Entropy}([9+, 5-]) = -\left(\frac{9}{14}\right)\log_2\left(\frac{9}{14}\right) - \left(\frac{5}{14}\right)\log_2\left(\frac{5}{14}\right)$$

$$= \underline{0.940}$$

⇒ Information gain of the attribute <u>Outlook</u> is calculated as.

Value (Outlook) = Sunny, Overcast, Rain

$S_{Sunny} \leftarrow [2+, 3-]$

$S_{Overcast} \leftarrow [4+, 0]$

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

$$\text{Gain}(S, \text{Outlook}) = \text{Entropy}(S) - \left[\left(\frac{5}{14}\right)\text{Entropy}(S_{Sunny}) + \left(\frac{4}{14}\right)\text{Entropy}(S_{Overcast}) + \left(\frac{5}{14}\right)\text{Entropy}(S_{Rain})\right]$$

* Entropy $(S_{sunny}) = -\left(\frac{2}{5}\right)\log_2\left(\frac{2}{5}\right) - \left(\frac{3}{5}\right)\log_2\left(\frac{3}{5}\right)$

$= -(0.4 * (-1.3219)) - (0.6 * (-0.7369))$

$= 0.52876 + 0.44214$

$= 0.970$

* Entropy $(S_{overcast}) = 0$ (because all members belong to same class)

* Entropy $(S_{rain}) = -\left(\frac{3}{5}\right)\log_2\left(\frac{3}{5}\right) - \left(\frac{2}{5}\right)\log_2\left(\frac{2}{5}\right)$

$= -(0.6 * (-0.7369) - (0.4 * (-1.3219))$

$= 0.44214 + 0.52876$

$= 0.970$

$= (0.940) - \left[\left(\frac{5}{14}\right) * 0.9709 + 0 + \left(\frac{5}{14}\right) 0.9709\right]$

$= 0.9409 - [0.3467 + 0.3467]$

$= \underline{0.246}$

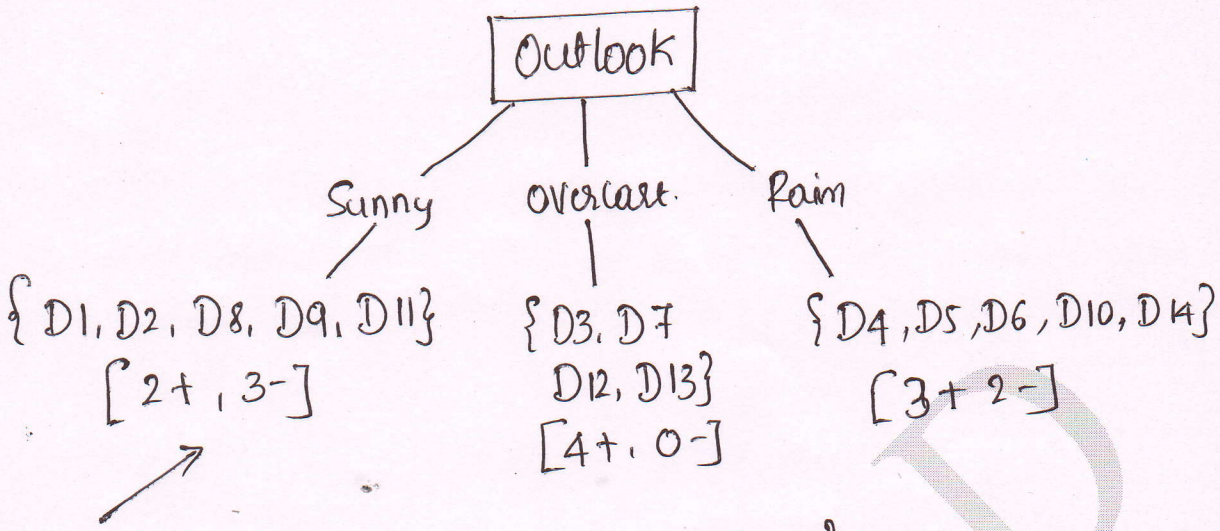
So Gain (S, outlook) = 0.246 ✓

Similarly Gain (S, Temperature) = 0.029

Gain (S, Humidity) = 0.151

Gain (S, wind) = 0.029

So root node will be <u>Outlook</u>



which attribute should be tested here?

$S_{sunny} = \{D1, D2, D8, D9, D11\}$

$Gain(Sunny, Humidity) = 0.970 - \left[\left(\frac{3}{5}\right) * 0.0 + \left(\frac{2}{5}\right) \cdot 0.0\right]$

$= \underline{0.970}$

$Gain(Sunny, Temperature) = 0.970 - \left[\left(\frac{2}{5}\right) * 0 + \left(\frac{2}{5}\right) * 1 + \left(\frac{1}{5}\right) \cdot 0\right]$

$= 0.970 - 0.4$

$= \underline{0.570}$

$Gain(Sunny, Wind) = 0.970 - \left[\left(\frac{3}{5}\right) * (0.918) + \left(\frac{2}{5}\right) \cdot 1\right]$

$= 0.970 - 0.9508$

$= \underline{0.0192}$

So, Attribute Humidity will be descendant node.

$\Rightarrow S_{Rain} = \{D_4, D_5, D_6, D_{10}, D_{14}\}$

$Entropy(S_{Rain}) = -\left(\frac{3}{5}\right)\log_2\left(\frac{3}{5}\right) - \left(\frac{2}{5}\right)\log_2\left(\frac{2}{5}\right)$

$= 0.970$

$Gain(Rain, Temperature) = 0.970 - \left[\left(\frac{0}{5}\right)*0 + \left(\frac{3}{5}\right)*0.918 + \left(\frac{2}{5}\right)*1\right]$

$= 0.0198$

$Gain(Rain, wind) = 0.970 - \left[\left(\frac{3}{5}\right)*0 + \left(\frac{2}{5}\right)*0\right]$

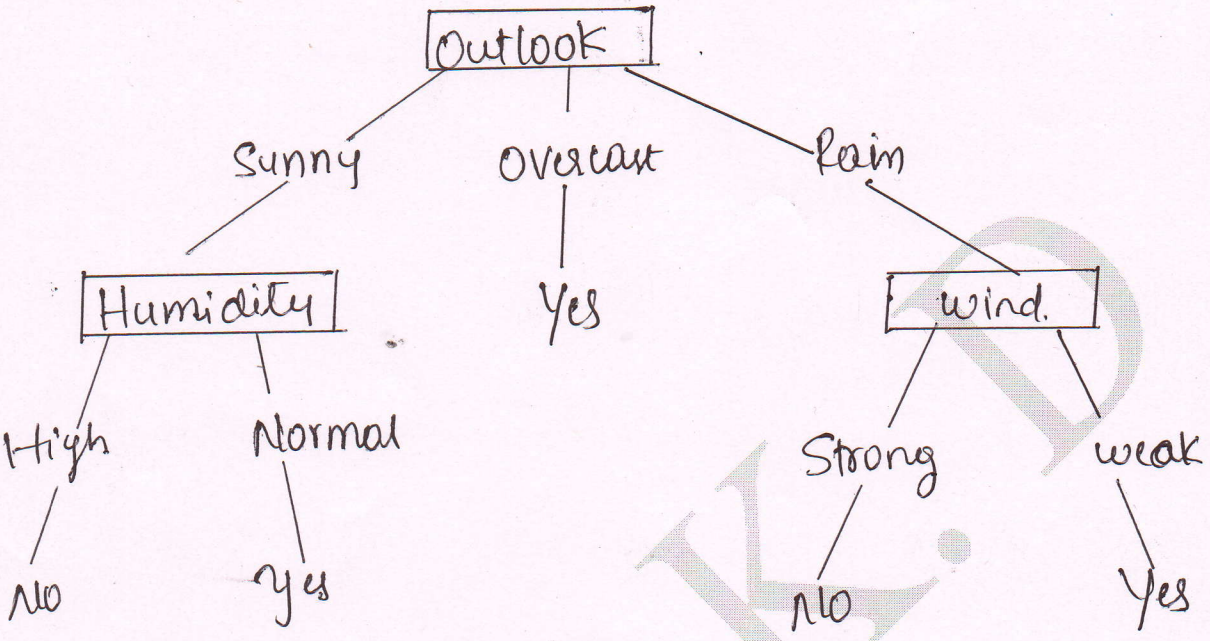
$= 0.970$

$Gain(Rain, Humidity) = 0.97 - \left[\left(\frac{2}{5}\right)*1 + \left(\frac{3}{5}\right)*0.917\right]$

$= 0.0198$

So, highest information gain is ~~code~~ attribute ~~Rain~~ Wind.

So, the final tree is.



DEEPAK.D

2. Consider the following set of training examples.

a) What is the entropy of this collection of training example with respect to the target function classification?

b) What is the information gain of a_2 relative to these training examples?

Instance	Classification	a_1	a_2
1	+	T	T
2	+	T	T
3	-	T	F
4	+	F	F
5	-	F	T
6	-	F	T

$\Rightarrow$ $Entropy(S) = -P_{\oplus} \log_2 P_{\oplus} - P_{\ominus} \log_2 P_{\ominus}$

Information gain

$$Gain(S, A) \equiv Entropy(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} Entropy(S_v)$$

a) positive instances = 03

Negative instances = 03

$$Entropy(S) = -\left(\frac{3}{6}\right) \log_2 \left(\frac{3}{6}\right) - \left(\frac{3}{6}\right) \log_2 \left(\frac{3}{6}\right)$$

$Entropy(S) = 1$

when there are equal number of +ve & -ve instances, then $Entropy(S) = 1$.

DEEPAK D

b) $Gain(S, a_2) = Entropy(S) - \left[\frac{4}{6} Entropy(S_T) + \frac{2}{6} Entropy(S_F)\right]$

find

① $Entropy(S_T) = -\left(\frac{2}{4}\right)\log_2\left(\frac{2}{4}\right) - \left(\frac{2}{4}\right)\log_2\left(\frac{2}{4}\right)$

$= \underline{1}$

② $Entropy(S_F) = -\left(\frac{1}{2}\right)\log_2\left(\frac{1}{2}\right) - \left(\frac{1}{2}\right)\log_2\left(\frac{1}{2}\right)$

$= \underline{1}$

$\Rightarrow Gain(S, a_2) = 1 - \left[\frac{4}{6}(1) + \frac{2}{6}(1)\right]$

$= 1 - 1$

$Gain(S, a_2) = \underline{0}$

3. Give decision trees to represent the following Boolean functions.

i) A && ¬B

ii) A V [B && C]

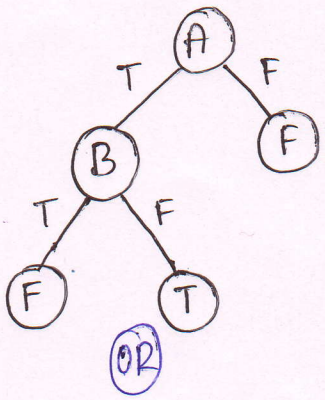
iii) A XOR B

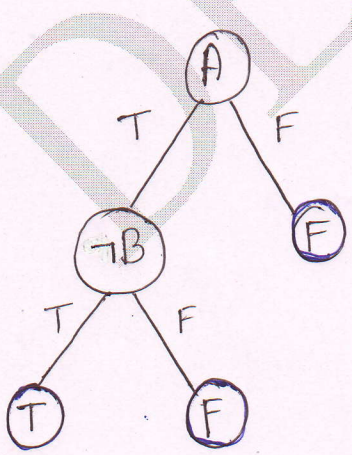
iv) [A && B] V [C && D]

Solution

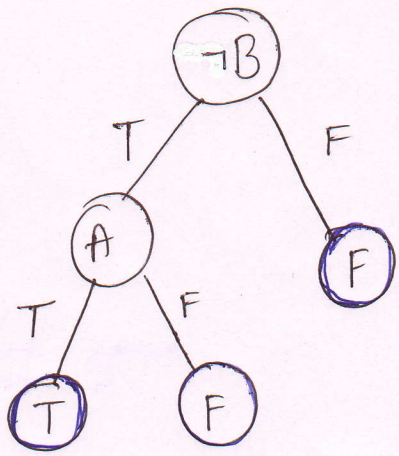
i) A && ¬B

A	B	¬B	A && ¬B
T	T	F	F (-)
T	F	T	T (+)
F	T	F	F (-)
F	F	T	F (-)





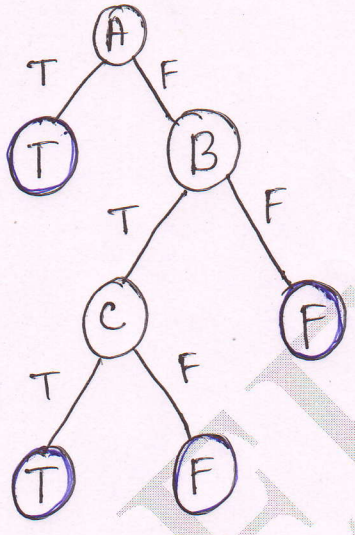
OR



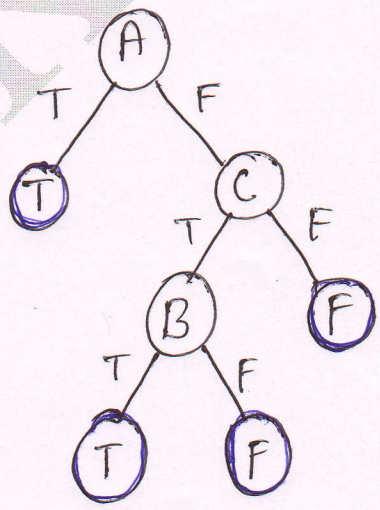
DEEPAK.D

ii) A V [B && C]

A	B	C	B && C	A V [B && C]
T	T	T	T	T
T	T	F	F	T
T	F	T	F	T
T	F	F	F	T
F	T	T	T	T
F	T	F	F	F
F	F	T	F	F
F	F	F	F	F

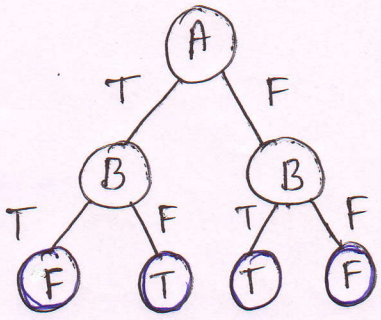


OR.



iii) A XOR B

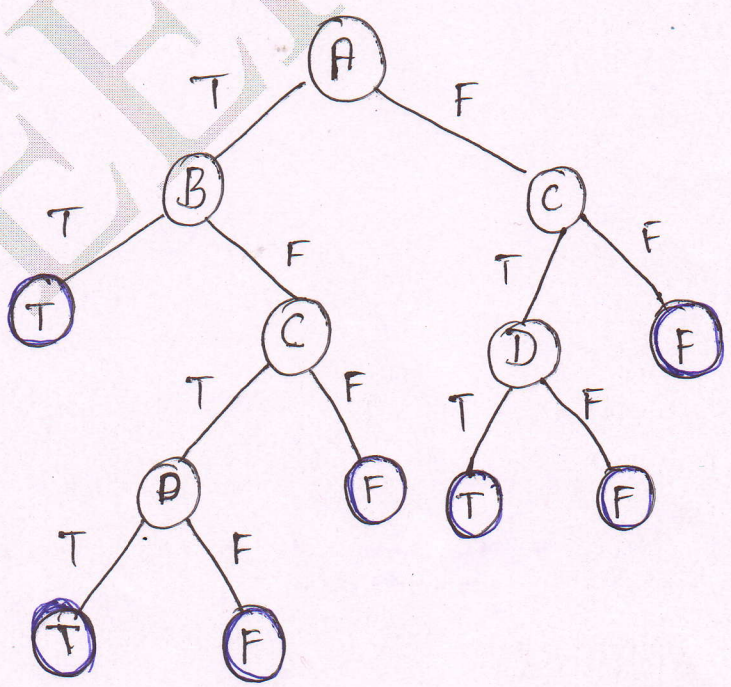
A	B	A XOR B
T	T	F
T	F	T
F	T	T
F	F	F



DEEPAK.D

iv) [A && B] v [C && D]

A	B	C	D	A && B	C && D	(A && B) v (C && D)
T	T	T	T	T	T	T
T	T	T	F	T	F	T
T	T	F	T	T	F	T
T	T	F	F	T	F	T
T	F	T	T	F	T	T
T	F	T	F	F	F	F
T	F	F	T	F	F	F
T	F	F	F	F	F	F
F	T	T	T	F	T	T
F	T	T	F	F	F	F
F	T	F	T	F	F	F
F	T	F	F	F	F	F
F	F	T	T	F	T	T
F	F	T	F	F	F	F
F	F	F	T	F	F	F
F	F	F	F	F	F	F



DEEPAK.D

MODULE - 1

1. Consider the Enjoy Sport concept and instance given below, identify the Specific hypothesis using Find-S algorithm.

Example	Sky	AirTemp	Humidity	Wind	Water	Forecast	EnjoySport
1	Sunny	Warm	Normal	Strong	Warm	Same	Yes
2	Sunny	Warm	High	Strong	Warm	Same	Yes
3	Rainy	Cold	High	Strong	Warm	Change	No
4	Sunny	Warm	High	Strong	Cool	Change	Yes

* First initialize 'h' to the most specific hypothesis in H

$$h = \langle \phi, \phi, \phi, \phi, \phi, \phi \rangle$$

* Consider the first training instance.

$$x_1 = \langle \text{Sunny, warm, Normal, Strong, warm, Same} \rangle, +$$

Observing instance x_1, 'h' is too Specific so replace by the next more general constraint that fits the example.

$$h_1 = \langle \text{Sunny, warm, Normal, Strong, warm, Same} \rangle$$

* Consider Second training instance.

$$x_2 = \langle \text{Sunny, warm, High, Strong, warm, Same} \rangle, +$$

Observe 'x_2' with 'h_1' and replace by more general constraint

$$h_2 = \langle \text{Sunny, warm, ?, Strong, warm, Same} \rangle$$

* Consider third training instance.

$x_3 = \langle$ Rainy, Cold, High, Strong, warm, Change $\rangle$, −

Find-S algorithm ignores negative instances, so. $h_3 = h_2$

$h_3 = \langle$ Sunny, warm, ?, Strong, warm, Same $\rangle$

* Consider fourth training instance.

$x_4 = \langle$ Sunny, warm, High, Strong, warm, Change $\rangle$, +

Compare x_4 with h_3 and replace by more general constraint

$h_4 = \langle$ Sunny, warm, ?, Strong, ?, ? $\rangle$

The final Specific hypothesis for given instances is

$h_f = \langle$ Sunny, warm, ?, Strong, ?, ? $\rangle$

DEEPAK.D

2. Consider the EnjoySport concept and instance given below, identify the general and specific hypotheses using Candidate - Elimination learning algorithm

Example	Sky	AirTemp	Humidity	Wind	Water	Forecast	EnjoySport
1	Sunny	Warm	Normal	Strong	Warm	Same	Yes
2	Sunny	Warm	High	Strong	Warm	Same	Yes
3	Rainy	Cold	High	Strong	Warm	Change	No
4	Sunny	Warm	High	Strong	Cool	Change	Yes

* The boundary sets are first initialized to G_0 & S_0 the most general & most specific hypotheses in H

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

G_0 $\langle ?, ?, ?, ?, ?, ? \rangle$

* Consider the first training instance.

$x_1 = \langle$ Sunny, Warm, Normal, Strong, Warm, Same $\rangle$

$S_0 = \langle \phi \ \phi \ \phi \ \phi \ \phi \ \phi \rangle$

$\downarrow$

$S_1 = \langle$ Sunny, Warm, Normal, Strong, Warm, Same $\rangle$

$G_0 \ G_1 = \langle ? \ ? \ ? \ ? \ ? \ ? \rangle$

DEEPAK.D

* Consider second instance

$x_2 = \langle$ Sunny, warm, High, Strong, warm, Same $\rangle$ +

$S_1 = \langle$ Sunny, warm, Normal, Strong, warm, Same $\rangle$

↓

$S_2 = \langle$ Sunny, warm, ?, Strong, warm, Same $\rangle$

$G_1, G_2 = \langle ?, ?, ?, ?, ?, ? \rangle$

* Consider third instance which is negative.

$x_3 = \langle$ Rainy, Cold, High, Strong, warm, change $\rangle$ −

S_2 $S_3 = \langle$ Sunny, warm, ?, Strong, warm, Same $\rangle$

$G_3 = \langle$ Sunny, ? ? ? ? ? $\rangle$ $\langle$? warm ? ? ? ? $\rangle$
$\langle$? ? ? ? ? Same $\rangle$

↑

G_2 $\langle ?, ?, ?, ?, ?, ? \rangle$

* Consider the fourth instance.

x_4 = <Sunny, warm, High, Strong, cool, change>

S_3 = <Sunny, warm, ?, Strong, warm, Same>

↓

S_4 = <Sunny, warm, ?, Strong, ? ?>

G_4 = <Sunny ? ? ? ? ?> <? warm ? ? ? ?>

↑

G_3 = <Sunny, ? ? ? ? ?> <? warm ? ? ? ?>
<? ? ? ? ? warm>

S_4 and G_4 are the final set of hypotheses which are consistent to training instances

DEEPAK D

3. Consider the "Japanese Economy Car" concept and instance given below, identify the hypotheses using Candidate - Elimination learning algorithm.

Origin	Manufacturer	Color	Decade	Type	Target Value
Japan	Honda	Blue	1980	Economy	Positive
Japan	Toyota	Green	1970	Sports	Negative
Japan	Toyota	Blue	1990	Economy	Positive
USA	Chryler	Red	1980	Economy	Negative
Japan	Honda	White	1980	Economy	Positive

* Initialize G_0 & S_0.

S_0 ⟨Japan ∅ ∅ ∅ ∅ ∅⟩

G_0 ⟨? ? ? ? ?⟩

* Consider the first training instance

x_1 = ⟨Japan, Honda, Blue, 1980, Economy⟩ +

S_0 = ⟨∅ ∅ ∅ ∅ ∅⟩

S_1 = ⟨Japan, Honda, Blue, 1980, Economy⟩

G_0 G_1 = ⟨? ? ? ? ?⟩

* Consider Second ~~training~~ instance.

$x_2 = \langle$ Japan, Toyota, Green, 1970, Sports $\rangle$ —

Specialize G to exclude the negative example.

$S_1 = S_2 = \langle$ Japan, Honda, Blue, 1980, Economy $\rangle$

$G_2 = \langle ?$ Honda $? ? ? \rangle \langle ?? $ Blue $? ? \rangle$
$\langle ? ? ? $ 1980 $? \rangle \langle ? ? ? ? $ Economy $\rangle$

↑

$G_1 = \langle ? ? ? ? ? \rangle$

* Considers third traing instance.

$x_3 = \langle$ Japan, Toyota, Blue, 1990, Economy $\rangle$

Prune G to exclude inconsistent hypothesis with the posite example and generalize S to include with positive example.

$S_3 = \langle$ Japan ? Blue ? Economy $\rangle$

$G_3 = \langle ? ?$ Blue $? ? \rangle \langle ? ? ? ?$ Economy $\rangle$

* Consider fourth instance

$x_4 = \langle$ USA, Chrysler, Red, 1980, Economy$\rangle$ –

Specialize G to include the negative example but stay consistent with S.

$S_4 = \langle$ Japan ? Blue, ? Economy$\rangle$

$G_4 = \langle$? ? Blue ? ?$\rangle$ $\langle$ Japan ? ? ? Economy$\rangle$

* Consider fifth instance

$x_5 = \langle$ Japan, Honda, White, 1980, Economy$\rangle$

Prune G to exclude inconsistent hypotheses with positive example and generalize S.

$S_5 = \langle$ Japan, ? ? ? Economy$\rangle$

$G_5 = \langle$ Japan, ? ? ? Economy$\rangle$

These are the final set of hypotheses which consistent with the training instance.

MODULE 5 – EVALUATING HYPOTHESIS, INSTANCE BASED LEARNING, REINFORCEMENT LEARNING

1. Explain the two key difficulties that arise while estimating the Accuracy of Hypothesis.
2. Define the following terms

   a. Sample error b. True error c. Random Variable

   d. Expected value e. Variance f. standard Deviation
3. Explain Binomial Distribution with an example.
4. Explain Normal or Gaussian distribution with an example.
5. Suppose hypothesis h commits r = 10 errors over a sample of n = 65 independently drawn examples.
   - What is the variance and standard deviation for number of true error rate $error_D(h)$?
   - What is the 90% confidence interval (two-sided) for the true error rate?
   - What is the 95% one-sided interval (i.e., what is the upper bound U such that $error_D(h) \leq 5$ U with 95% confidence)?
   - What is the 90% one-sided interval?

α	0.100	0.050	0.025	0.001
$1-\alpha$	0.900	0.950	0.975	0.999
$z_{1-\alpha}$	1.28	1.64	1.96	3.09

6. What are instance based learning? Explain key features and disadvantages of these methods.
7. Explain the K – nearest neighbour algorithm for approximating a **discrete – valued** function $f : \Re^n \rightarrow V$, with pseudo code
8. Describe K-nearest Neighbour learning Algorithm for **continues (real) valued** target function.
9. Discuss the major drawbacks of K-nearest Neighbour learning Algorithm and how it can be corrected

10. Define the following terms with respect to K - Nearest Neighbour Learning :
    i) Regression ii) Residual iii) Kernel Function.
11. Explain Locally Weighted Linear Regression.
12. Explain radial basis function
13. Explain CADET System using Case based reasoning.
14. What is Reinforcement Learning and explain Reinforcement learning problem with neat diagram.
15. Write Reinforcement learning problem characteristics.
16. Explain the Q function and Q Learning Algorithm assuming deterministic rewards and actions with example.

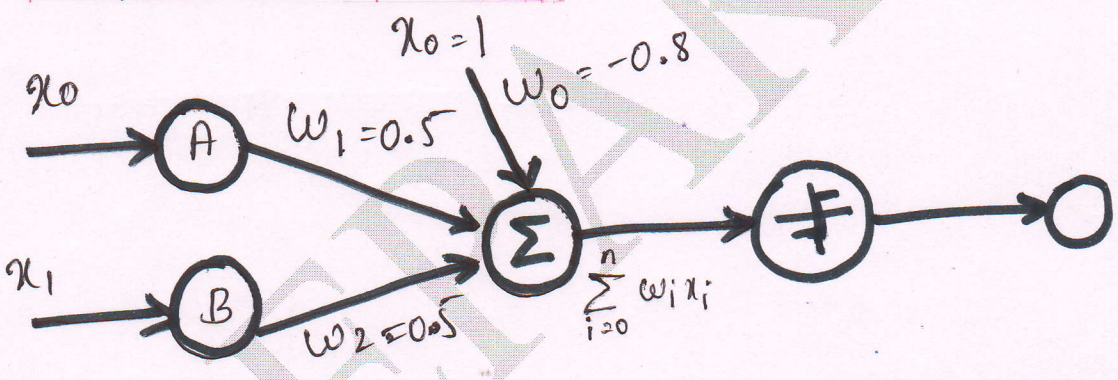
MODULE 3

1. How a single perceptron can be used to represent the Boolean functions such as AND, OR

⇒ Boolean function AND

A	B	A ∧ B
0	0	0
0	1	0
1	0	0
1	1	1

* Let $w_0 = -0.8$
$w_1 = 0.5$
$w_2 = 0.5$



$$O(x_1 \ldots x_n) = \begin{cases} 1 & \text{if } w_0 + w_1 x_1 + w_2 x_2 + \ldots + w_n x_n > 0 \\ -1 & \text{otherwise.} \end{cases}$$

1) if A=0 & B=0 ⇒ x_1 x_2 $0 * 0 - 0.8 + (0.5 * 0) + (0.5 * 0) =$
$= -0.8 < 0$ So, output = 0

2) if A=0 & B=1 ⇒ $-0.8 + (0.5 * 0) + (0.5 * 1)$
$= -0.3 < 0$ So, output = 0

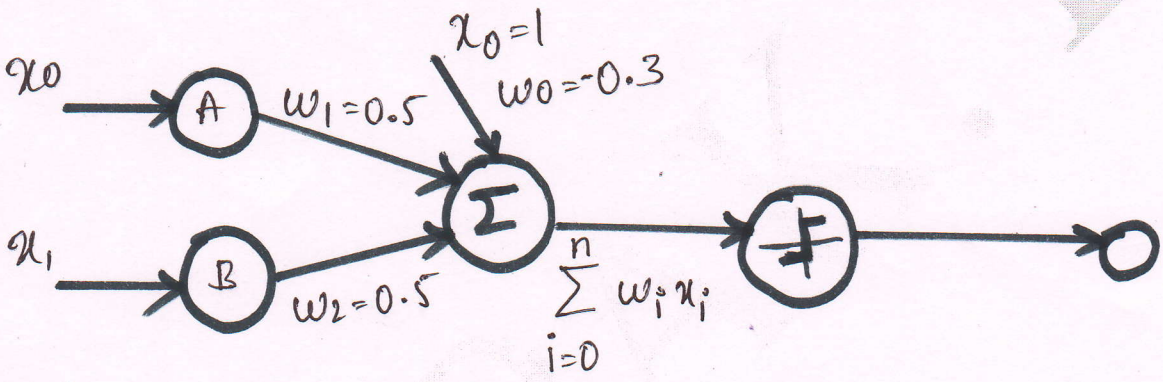
3) if A=1 & B=1 ⇒ $-0.8 + (0.5 * 1) + (0.5 * 0) = -0.3 < 0$
Output = 0

4) if A=1 & B=1 ⇒ $-0.8 + (0.5 * 1) + (0.5 * 1) = 0.2 > 0$
output = 1

⇒ Boolean function OR

A	B	A V B
0	0	0
0	1	1
1	0	1
1	1	1

* Set $w_0 = -0.3$
$w_1 = 0.5$
$w_2 = 0.5$



i) A = 0 B = 0 ⇒ $-0.3 + (0.5 * 0) + (0.5 * 0)$
$= -0.3 < 0$ So output = 0

2) A = 0 B = 1 ⇒ $-0.3 + (0.5 * 0) + (0.5 * 1)$
$= 0.2 > 0$ So output = 1

3) A = 1 B = 0 ⇒ $-0.3 + (0.5 * 1) + (0.5 * 0)$
$= 0.2 > 0$ So, output = 1

4) A = 1 B = 1 ⇒ $-0.3 + (0.5 * 1) + (0.5 * 0)$
$= 0.7 > 0$ So output = 1

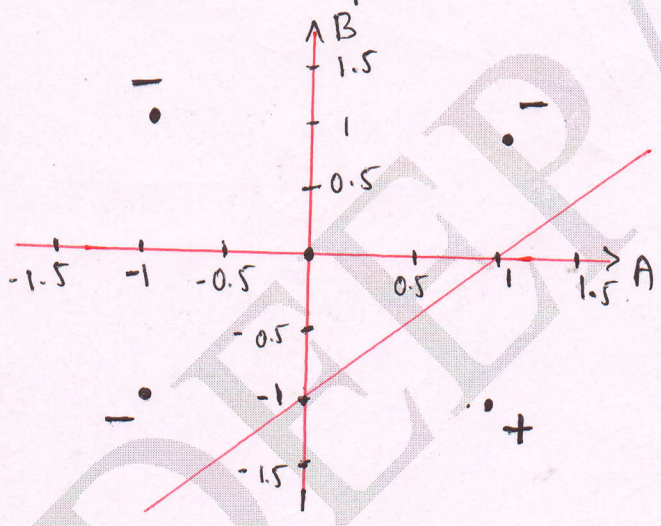
2 (a) Design a two-input perceptron that implements the boolean function A Λ ¬ B. (b) Design a two-layer network of perceptron's that implements A XOR B.

⇒ a) The perceptron has two input A, B and constant 1

A	B	¬B	A Λ ¬B
0 (-1)	0 (-1)	1	0 (-1)
0 (-1)	1	0 (-1)	0 (-1)
1	0 (-1)	1	1
1	1	0	0 (-1)

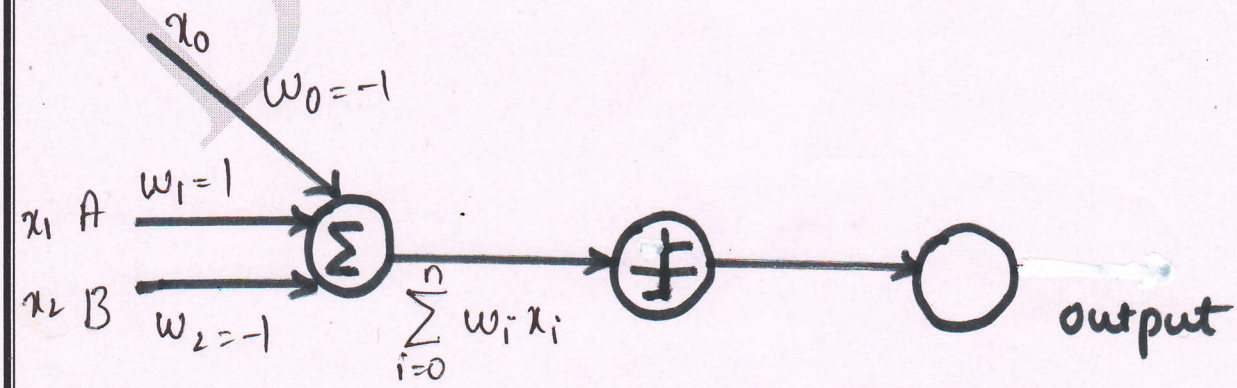
The values of A & B are 1 (true) or -1 or 0 for false.

Decision surfaces



* The line crosses the A axis at 1 and B axis at -1

* The weights are
$w_0 = -1$
$w_1 = 1$ $w_2 = -1$.



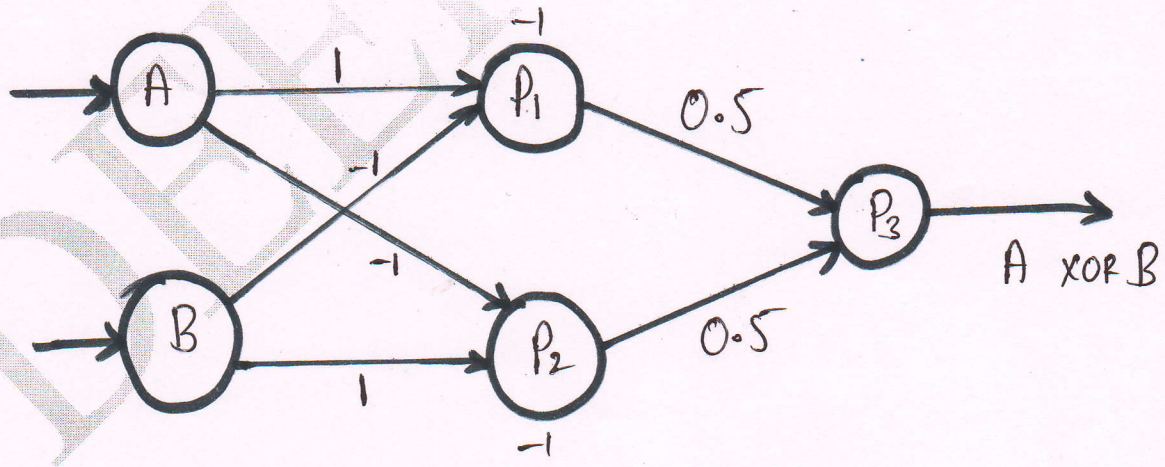
⟹ b) A XOR B cannot be calculated by a single perceptron, so build a two-layer network of perceptrons

* Express A XOR B in terms of other logical connectives

$$A \text{ XOR } B = (A \wedge \neg B) \vee (\neg A \wedge B)$$

* Define the perceptron P_1 and P_2 for $(A \wedge \neg B)$ & $(\neg A \wedge B)$

* Composing the outputs of P_1 & P_2 into a perceptron P_3 that implements $O(P_1) \vee O(P_2)$



3. Consider two perceptrons defined by the threshold expression $w_0 + w_1x_1 + w_2x_2 > 0$.

Perceptron A has weight values

$w_0 = 1, w_1=2, w_2=1$

and perceptron B has the weight values

$w_0 = 0, w_1=2, w_2=1$

True or false? Perceptron A is more-general than perceptron B.

Solution

True, Perceptron A is more-general than Perceptron B.

$\Rightarrow O(x_1 \ldots x_n) = w_0x_0 + w_1x_1 + w_2x_2 + \ldots + w_nx_n$

$\Rightarrow$ If $B(\langle x_1, x_2 \rangle) = 1$ & $w_0 = 0, w_1 = 2, w_2 = 1$

$0 + 2x_1 + x_2 > 0 \Rightarrow 0 + 2 + 1 > 0$

where, x_0 is constant which is equal to 1 i.e., $x_0 = 1$

$\Rightarrow A(\langle x_1, x_2 \rangle) = 1$ & $w_0 = 1, w_1 = 2, w_2 = 1$

$1 + 2x_1 + x_2 > 0 \Rightarrow 1 + 2 + 1 > 0$

Here, Perceptron A is more general than perceptron B because every instance of x_1 & x_2 that satisfies perceptron B also satisfies perceptron A.

Save trees use e-Question Paper Go green
DOWNLOAD THIS FREE AT www.vturesource.com

CBCS SCHEME

USN | | | | | | | | | | 15CS73

Seventh Semester B.E. Degree Examination, June/July 2019
Machine Learning

Time: 3 hrs. Max. Marks: 80

Note: Answer any FIVE full questions, choosing ONE full question from each module.

Important Note : 1. On completing your answers, compulsorily draw diagonal cross lines on the remaining blank pages.
2. Any revealing of identification, appeal to evaluator and /or equations written eg, 42+8 = 50, will be treated as malpractice.

Module-1

1 a. Define machine learning. Describe the steps in designing learning system. **(08 Marks)**
b. Write Find-S algorithm and explain with example. **(04 Marks)**
c. Explain List-Then-Eliminate algorithm. **(04 Marks)**

OR

2 a. List out any 5 applications of machine learning. **(05 Marks)**
b. What do you mean by hypothesis space, instance space and version space? **(03 Marks)**
c. Find the maximally general hypothesis and maximally specific hypothesis for the training examples given in the table using candidate elimination algorithm. **(08 Marks)**

Day	Sky	Air Temp	Humidity	Wind	Water	Forecast	Enjoy Sport
1	Sunny	Warm	Normal	Strong	Warm	Same	Yes
2	Sunny	Warm	High	Strong	Warm	Same	Yes
3	Rainy	Cold	High	Strong	Warm	Change	No
4	Sunny	Warm	High	Strong	Cool	Change	Yes

Module-2

3 Construct decision tree for the following data using ID3 algorithm.

Day	A1	A2	A3	Classification
1	True	Hot	High	No
2	True	Hot	High	No
3	False	Hot	High	Yes
4	False	Cool	Normal	Yes
5	False	Cool	Normal	Yes
6	True	Cool	High	No
7	True	Hot	High	No
8	True	Hot	Normal	Yes
9	False	Cool	Normal	Yes
10	False	Cool	High	No

(16 Marks)

OR

4 a. Explain the concept of decision tree learning. Discuss the necessary measure required to select the attributes for building a decision tree using ID3 algorithm. **(08 Marks)**
b. Discuss the issues of avoiding over fitting the data, handling continuous data and missing values in decision trees. **(08 Marks)**

Module-3

5 a. Explain artificial neural network based on perception concept with diagram. **(06 Marks)**
b. What is gradient descent and delta rule? Why stochastic approximation to gradient descent is needed? **(04 Marks)**
c. Describe the multilayer neural network. Explain why back propagation algorithm is required. **(06 Marks)**

www.vturesource.com
ALL BRANCHES | ALL SEMESTERS | NOTES | QUESTON PAPERS | LAB MANUALS
A Vturesource Go Green initiative

Save trees use e-Question Paper

Go green

DOWNLOAD THIS FREE AT www.vturesource.com

OR

6 a. Derive the back propagation rule considering the output layer and training rule for output unit weights. **(08 Marks)**

b. What is squashing function 3 why is it needed? **(04 Marks)**

c. List out and explain in briefly representation power of feed forward networks. **(04 Marks)**

Module-4

7 a. Explain maximum a posteriori (MAP) hypothesis using Bayes theorem. **(06 Marks)**

b. Estimate conditional probabilities of each attributes {colour, legs, height, smelly} for the species classes: {M, H} using the data given in the table. Using these probabilities estimate the probability values for the new instance – (Colour = Green, Legs = 2, Height = Tall and Smelly = No) **(10 Marks)**

No	Colour	Legs	Height	Smelly	Species
1	White	3	Short	Yes	M
2	Green	2	Tall	No	M
3	Green	3	Short	Yes	M
4	White	3	Short	Yes	M
5	Green	2	Short	No	H
6	White	2	Tall	No	H
7	White	2	Tall	No	H
8	White	2	Short	Yes	H

OR

8 a. Explain Naive Bayes classifier and Bayesian belief networks. **(10 Marks)**

b. Prove that how maximum likelihood (Bayesian learning) can be used in any learning algorithms that are used to minimize the squared error between actual output hypothesis and predicted output hypothesis. **(06 Marks)**

Module-5

9 a. Explain locally weighted linear regression. **(08 Marks)**

b. What do you mean by reinforcement learning? How reinforcement learning problem differs from other function approximation tasks. **(05 Marks)**

c. Write down Q-learning algorithm. **(03 Marks)**

OR

10 a. What is instance based learning? Explain K-Nearest neighbour algorithm. **(08 Marks)**

b. Explain sample error, true error, confidence intervals and Q-learning function. **(08 Marks)**

* * * * *

www.vturesource.com

ALL BRANCHES | ALL SEMESTERS | NOTES | QUESTON PAPERS | LAB MANUALS

A Vturesource Go Green initiative