CBCS SCHEME

USN | | | | | | | | | | | 17CS73

Seventh Semester B.E. Degree Examination, Jan./Feb. 2021
Machine Learning

Time: 3 hrs.   Max. Marks: 100

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

Module-1

1 a. Define machine learning. Mention five applications of machine learning. **(06 Marks)**

b. Explain concept learning task with an example. **(06 Marks)**

c. Apply candidate elimination algorithm and obtain the version space considering the training examples given Table Q1(c).

Eyes	Nose	Head	Fcolor	Hair?	Smile?(TC)
Round	Triangle	Round	Purple	Yes	Yes
Square	Square	Square	Green	Yes	No
Square	Triangle	Round	Yellow	Yes	Yes
Round	Triangle	Round	Green	No	No
Square	Square	Round	Yellow	Yes	Yes

Table Q1(c) **(08 Marks)**

OR

2 a. Explain the following with respect to designing a learning system :

i) Choosing the training experience

ii) Choosing the target function

iii) Choosing a representation for the target function. **(09 Marks)**

b. Write Find–S algorithm. Apply the Find–S for Table Q1(c) to find maximally specific hypothesis. **(06 Marks)**

c. Explain the concept of inductive bias. **(05 Marks)**

Module-2

3 a. Explain the concept of decision tree learning. Discuss the necessary measures required to select the attributed for building a decision tree using ID3 algorithm. **(11 Marks)**

b. Explain the following with respect to decision tree learning :

i) Incorporating continuous valued attributes

ii) Alternative measures for selecting attributes

iii) Handling training examples with missing attribute values. **(09 Marks)**

OR

4 a. Construct decision tree using ID3 considering the following training examples :

Weekend	Weather	Parental availability	Wealthy	Decision class
H_1	Sunny	Yes	Rich	Cinema
H_2	Sunny	No	Rich	Tennis
H_3	Windy	Yes	Rich	Cinema
H_4	Rainy	Yes	Poor	Cinema
H_5	Rainy	No	Rich	Home
H_6	Rainy	Yes	Poor	Cinema
H_7	Windy	No	Poor	Cinema
H_8	Windy	No	Rich	Shopping
H_9	Windy	Yes	Rich	Cinema
H_{10}	Sunny	No	Rich	Tennis

Table Q4(b) **(12 Marks)**

b. Discuss the issues of avoiding overfitting the data, and handling attributes with differing costs. **(08 Marks)**

Module-3

5 a. Discuss the application of neural network which is used to steer an autonomous vehicle. **(06 Marks)**

b. Write Gradient descent algorithm to train a linear unit along with the derivation. **(08 Marks)**

c. Discuss the issues of convergence, local minima and generalization, overfitting and stopping criterion. **(06 Marks)**

OR

6 a. List the appropriate problems for neural network learning. **(05 Marks)**

b. Define preceptron and discuss its training rule. **(05 Marks)**

c. Show the derivation of back propagation training rule for output unit weights. **(10 Marks)**

Module-4

7 a. Explain Bayes theorem and mention the features of Bayesian learning. **(07 Marks)**

b. Prove that a maximum likelihood hypotheses cab be used to predict probabilities. **(08 Marks)**

c. Explain Naïve Bayes classifier. **(05 Marks)**

OR

8 a. Describe MAP learning algorithm. **(08 Marks)**

b. Classify the test data and {Red, SUV, Domestic} using Naine Bayes classifier for the dataset shown in Table Q8(b).

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

Table Q8(b) **(06 Marks)**

c. Write and explain EM algorithm. **(06 Marks)**

Module-5

9 a. Define :

i) Sample error

ii) True error

iii) Confidence intervals. **(06 Marks)**

b. Explain K-nearest neighbor learning algorithm. **(08 Marks)**

c. Write a note no Q – learning. **(06 Marks)**

OR

10 a. Define mean value, variance, standard deviation and estimation bias of a random variable. **(04 Marks)**

b. Explain locally weighted linear regression and radial basis functions. **(10 Marks)**

c. What is reinforcement learning? How it differs from other function approximation tasks? **(06 Marks)**

* * * * *

Scheme & Solutions

Subject Title : Machine learning **Subject Code :** 17CS73

Question Number	Solution	Marks Allocated
1a)	Defn A program that improves its performance at some task through Experience	1M
	Applns any five : (i) Learning to recognize spoken words (ii) Learning to drive an autonomous vehicle (iii) Learning to classify new astronomical structures (iv) Learning to play world-class backgammon (v) Predicting recovery rates of pneumonia patients (vi) Detecting fraudulent use of credit cards	1x5=5M
b)	Concept Learning – Problem of searching through a predefined space potential hypotheses for the hypothesis that best fits the training Examples — 01	1½x4
	Target Concept - defn — 01	
	Learner Task - predicting the value for new instance — 02	
	Representation of Hypothesis — 02	
c)	$S_0 : \{<\phi, \phi, \phi, \phi, \phi>\}$	
	$G_0 : \{<?, ?, ?, ?, ?>\}$	
	$S_1 : \{<R, T, R, P, Y>\}$ $x_1 = <R, T, R, P, Y> +$	
	$G_1 : \{<?, ?, ?, ?, ?>\}$	
	$S_2 : \{<R, T, R, P, Y\}$ $x_2 = <S, S, S, G, Y> -$	
	$G_2 : \{<R, ?, ?, ?, ?>, <?, T, ?, ?, ?>, <?, ?, R, ?, ?>$ $<?, ?, ?, P, ?>, <?, ?, ?, P, ?>\}$	

Subject Title : Machine learning   Subject Code : 17CS73

Question Number	Solution	Marks Allocated

$S_3 : \{<?, T, R, ?, Y>\}$ $x_3 = <S, T, R, Y, Y>$ +

$G_3 : \{\cancel{<R, ?, ?, ?, ?>} <?, T, ?, ?, ?> <?, ?, R, ?, ?> \cancel{<?, ?, ?, P, ?>}\}$

$S_4 : \{<?, T, R, ?, Y>\}$ $x_4 = <R, T, R, G, N>$ −

$G_4 : \{<?, T, ?, ?, Y> <?, ?, R, ?, Y>\}$

$S_5 : \{<?, ?, R, ?, Y>\}$ * $x_5 = <S, S, R, Y, Y>$ +

$G_5 : \{\cancel{<?, T, ?, ?, Y>} <?, ?, R, ?, Y>\}$

version space Since S_5 & G_5 are same, previous result are taken to represent. Since 4^{th} instance is −ve, we will take from S_3 & G_3

$G : \boxed{\{<?, ?, R, ?, ?> <?, T, ?, ?, ?>\}}$

$<?, ?, R, ?, Y> \quad <?, T, R, ?, ?> \quad <?, T, ?, ?, Y>$

$S : \boxed{\{<?, T, R, ?, Y>\}}$

2 a) (i) Choosing the Training Experience < Direct / Indirect

Degree to which learner controls the sequence
− No freedom, semi-free, full freedom
− Distribution of training examples

(ii) Choosing target function ~~$V : B \rightarrow R$~~

Choose move : $B \rightarrow M$

alternate $V : B \rightarrow R$

(iii) Representation $\hat{V}(b) : w_0 + w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4 + w_5x_5 + w_6x_6$

3×3 = 9

Subject Title : Machine learning Subject Code : 17CS73

Question Number	Solution	Marks Allocated
2 b)	Algorithm: 1. Initialize h to the most specific hypothesis in H 2. For each +ve training instance x • For each attribute constraint a_i in h If the constraint a_i is satisfied by x Then do nothing Else replace a_i in h by the next more general constraint that is satisfied by x 3. Output hypothesis h	— 3M
	$h_0 = \langle \phi, \phi, \phi, \phi, \phi \rangle$ $h_1 = \langle R, T, R, P, Y \rangle$, $h_2 = \langle R, T, R, P, Y \rangle$ $h_3 = \langle ?, T, R, ?, Y \rangle$ $h_4 = \langle ?, T, R, ?, Y \rangle$ $h_5 = \langle ?, ?, R, ?, Y \rangle$	— 3M
c)	A Biased Hypothesis Space — 2 An Unbiased learner — 2 The Futility of Bias-free learning - 1	5M
3 a)	Concept of decision tree learning — 5 $\text{Entropy}(S) = -P_{\oplus}\log_2 P_{\oplus} - P_{\ominus}\log_2 P_{\ominus}$ $\text{Entropy}(S) = \sum_{i=1}^{c} -P_i \log_2 P_i$ — 3 $\text{Gain}(S,A) = \text{Entropy}(S) - \sum_{v \in \text{Value}(A)} \frac{\lvert S_v \rvert}{\lvert S \rvert} \text{Entropy}(S_v)$ — 3 With Examples	11M
b)	(i) Incorporating Continuous-valued attributes — 3	3
	(2) Alternative measures $\text{SplitInfo}(S,A) = -\sum_{i=1}^{c} \frac{\lvert S_i \rvert}{\lvert S \rvert} \log_2 \frac{\lvert S_i \rvert}{\lvert S \rvert}$ $\text{GainRatio}(S,A) = \frac{\text{Gain}(S,A)}{\text{SplitInfo}(S,A)}$	— 3
	(iii) Training Examples with missing values	— 3

Subject Title : Machine learning Subject Code : 17CS73

Question Number	Solution	Marks Allocated

4 a)

$E(S) = 1.571$

$Gain(S, weather) = \boxed{0.6957}$ ← Highest $Gain(S, parents) = 0.3779$

$Gain(S, wealthy) = 0.2816$

$Gain(S_{Sunny}, Parents) = \boxed{0.918}$ – Highest

$Gain(S_{Sunny}, wealthy) = 0$

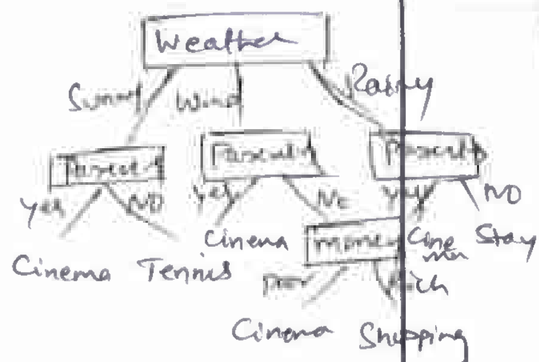
12 M

$Gain(S_{Windy}, Parents) = 0.31125$

$Gain(S_{Windy}, wealthy) = 0.12625$

$Gain(S_{Rainy}, Parents) = 0.918$

$Gain(S_{Rainy}, money) = 0.918$



b)

avoiding overfitting data – 5

Handling attributes with differing costs – 3

8 M

5 a)

Representation – 2

Explanation – 4

6 M

b)

Gradient-Descent(training examples, η)

$\langle \vec{x}, \vec{t} \rangle$

- Initialize each w_i to some small random value
- Until termination condn is met, Do
  - Initialize each Δw_i to 0
  - For each $(\vec{x}, \vec{t})$ in training examples, Do
    - Input the instance $\vec{x}$ to the unit & compute o/p o
    - For each linear unit weight w_i, Do

      $\Delta w_i \leftarrow \Delta w_i + \eta (t - o) x_i$
  - For each linear unit weight w_i, Do

    $w_i \leftarrow w_i + \Delta w_i$

4M

Derivation – 4M

Subject Title : Machine learning **Subject Code :** 17CS73

Question Number	Solution	Marks Allocated
5(c)	Convergence & Local minima — Generalization, overfitting & stopping criterion —	3 3
6 a)	5 x 1 —	5M
b)	Perceptron diagram, Equation, Explanation — Training rule —	3M 2M
c)	Backpropagation Rule : $\Delta w_{ji} = -\eta \frac{\partial E_d}{\partial w_{ji}}$, $E_d = \frac{1}{2}\sum_{k \in outputs}(t_k - o_k)^2$ $\frac{\partial E_d}{\partial w_{ji}}$ Finally $\Delta w_{ji} = -\eta \frac{\partial E_d}{\partial w_{ji}} = \eta (t_j - o_j) o_j (1 - o_j) x_{ji}$	10M
7 a)	Bayes Theorem — 3M Features — 4M	7M
b)	Proof: $h_{ML} = \underset{h \in H}{\arg\max} \sum_{i=1}^{m} d_i \ln h(x_i) + (1 - d_i) \ln(1 - h(x_i))$	8M
c)	Explanation —	5M
8a)	1. $P(h \mid D) = \frac{P(D \mid h) P(h)}{P(D)} = \begin{cases} \frac{1}{\lvert VS_{H,D} \rvert} & \text{if } h \text{ is consistent w.r.t } D \\ 0 & \text{otherwise} \end{cases}$ 2. $h_{MAP} = \underset{h \in H}{\arg\max}\, P(h \mid D)$ derivation	8M
b)	$P_{yes} = 5/10 = 0.5$, $P_{No} = \frac{5}{10} = 0.5$ $P(yes) \cdot P(Red/yes)\, P(SUV/yes) \cdot P(Domestic/yes) = 0.5 \times 0.6 \times 0.2 \times 0.4 = 0.024$ $P(No) \cdot P(Red/No) \cdot P(SUV/No) \cdot P(Domestic/No) = 0.5 \times 0.4 \times 0.6 \times 0.6 = 0.072$ (Red, SUV, Domestic) — No, Normalize $\frac{0.072}{0.024 + 0.072} = \boxed{0.75}$	6M
c)	Algorithm — 3, Explanation - 3	3+3
9	a) 2 x 3 = 6 b) Algorithm + Explanation — 4+4 c) Q-learning Explanation —	8 6
10	a) 4 x 2 - 8M b) ~~Line~~ 2 x 5 — 10M c) defn - 2 diff — 4	6M