

Roll No.



**PRESIDENCY UNIVERSITY  
BENGALURU  
SCHOOL OF ENGINEERING**

**TEST 1**

**Sem & AY:** Odd Sem 2019-20

**Course Code:** CSE 319

**Course Name:** MACHINE LEARNING

**Program & Sem:** B.Tech (CSE) & VII DE

**Date:** 30.09.2019

**Time:** 9:30AM TO 10:30AM

**Max Marks:** 40

**Weightage:** 20%

**Instructions:**

- (i) All questions are compulsory  
(ii) Write the answers legibly

**Part A [Memory Recall Questions]**

**Answer all the Questions. Each Question carries one mark**

**(10Qx1M=10M)**

**1. Choose the correct answer from the given answers.**

**(C.O.NO.1) [Knowledge]**

- i). Calculating the value of the real estate property can be categorized under
- a. Unsupervised learning      b. Classification      c. Regression  
d. Reinforcement Learning
- ii). Clustering is which type of Machine Learning?
- a. Supervised Learning      b. Unsupervised Learning      c. Reinforcement learning  
d. All of the above
- iii). Which of the following are NOT part of machine learning?
- a. Input Data      b. Generalization      c. Optimization      d. Abstraction
- iv). Machine Learning aims to increase which of the following:
- a. Input Instances      b. Task      c. Accuracy      d. a & c
- v). Which of the following learning modes is high on accuracy in detecting objects from noisy video?
- a. Supervised      b. Unsupervised      c. Reinforcement      d. a & c
- vi). In Reinforcement learning, the machine learns based on the concept of:
- a. Fine tuning      b. Reward      c. Partitioning      d. Cumulative Points

vii). What term/s does not relate to Bayesian framework –

- a. Priori                      b. Posterior                      c. Likelihood                      d. none of a, b, c

viii). In ML experiment if feature space of classifier increases, this nominally requires -

- a. increased training size    b. decreased training size                      c. either a or b                      d. none

ix). This type of learning is to be used when there is no idea about the class or label of a particular data

- a. Reinforcement learning                      b. Supervised learning  
c. Semi-supervised learning                      d. Unsupervised learning

x). If  $P(A) = 0.5$ ,  $P(B) = 0.35$  and  $P(B|A) = 0.2$ , then  $P(A|B) = ?$

- a. 0.285                      b. 0.3                      c. 0.258                      d. 0.29

### Part B [Thought Provoking Questions]

**Answer both the Questions. Each Question carries five marks. (2Qx5M=10M)**

2. A basket is filled with some fresh fruits like apple, banana, cherry, grape, the task is to arrange the same type of fruits at one place. Which learning technique is used? Explain and justify the answer. (C.O.NO.1) [Comprehension]
3. Is Machine Learning needed in all problems? Discuss in detail. (C.O.NO.1) [Comprehension]

### Part C [Problem Solving Questions]

**Answer both the Questions. Each Question carries ten marks. (2Qx10M=20M)**

4. With an example for each, explain all the types of Machine Learning techniques and the concept used in the same. (C.O.NO.2) [Comprehension]
5. There two hypothesis: A particular tumour is of malignant type and A particular tumour is non-malignant type. The prior available data are: Only 0.5% of the population has this kind of tumour which is malignant. The laboratory report has some amount of incorrectness as it could detect the malignancy was present only with 98% accuracy, where as it could detect the malignancy was not present correctly only in 97% of cases. This means the test predicted malignancy was present which actually was a false alarm in 2% of the cases and also missed detecting the real malignant tumour in 3% of the cases.

We will calculate how the prior knowledge of the percentage of cancer cases in a sample population and probability of the test result being correct influence the probability outcome of correct diagnosis. (C.O.NO.1) [Application]



## SCHOOL OF ENGINEERING

Semester: VII

Course Code: CSE 319

Course Name: MACHINE LEARNING

Date: 27-SEP-19

Time: 9:30AM TO 10:30AM

Max Marks: 40

Weightage: 20%

### Extract of question distribution [outcome wise & level wise]

Q. N O.	C.O. NO	Unit/Module Number/Unit  /Module Title	Memory recall type  [Marks allotted]  Bloom's Levels			Thought provoking type  [Marks allotted]  Bloom's Levels			Problem Solving type  [Marks allotted]			Total Mark s
			K			C			A			
1	CO1	MODULE-1, MODULE-2		10								10
2	CO1	MODULE-1					5					5
3	CO1	MODULE-1		5								5
4	CO2	MODULE-1								10		10
5	CO1	MODULE-1									10	10
	Total Marks			15			5			10	10	40

K = Knowledge Level    C = Comprehension Level, A = Application Level

Note: While setting all types of questions the general guideline is that about 60%

Of the questions must be such that even a below average students must be able to attempt, About 20% of the questions must be such that only above average students must be able to attempt and finally 20% of the questions must be such that only the bright students must be able to attempt.

[I hereby certify that All the questions are set as per the above guide lines. Mr. Sanjeev Kaulgud ]

Reviewers' Comments



**Part A**

(10Q x 1M = 10Marks)

Q No	Solution	Scheme of Marking	Max. Time required for each Question
1	i. C ii. B iii. C iv. C v. A vi. B vii. D viii. A ix. D x. A	Total Marks: 10M  1M each for correct answer	10 Mins

**Part B**

(2Q x 5M = 10 Marks)

Q No	Solution	Scheme of Marking	Max. Time required for each Question
2	SUPERVISED LEARNING IS REQUIRED, since the labels are mentioned as the particular fruits.	Total Marks: 5M Correct answer with proper justification and explanation about supervised learning	10 Mins
3	Machine Learning is not required in all conditions like: If the problem to be solved is easy If the accuracy can be compromised etc.	Total Marks: 5M Listing of all conditions where ML is not required	10 Mins

Part C

(2Q x 10M = 20 Marks)

Q No	Solution	Scheme of Marking	Max. Time required for each Question
4	<p>Explanation of all the three types of machine learning with an example each</p> <ol style="list-style-type: none"> <li>1. Supervised Learning</li> <li>2. Unsupervised Learning</li> <li>3. Reinforcement Learning</li> </ol>	<p>Total Marks: 10M</p> <p>Supervised: 4M Unsupervised: 3M Reinforcement: 3M</p>	15 Mins
5	<p>Malignant Tumour = MT, Positive Lab Test = PT, Negative Lab Test = NT</p> <p>h1 = the particular tumour is of malignant type = MT in our example h2 = the particular tumour is of not malignant type = !MT in our example</p> <p>P(MT) = 0.005                      P(!MT) = 0.995 P(PT MT) = 0.98                  P(PT !MT) = 0.02 P(NT !MT) = 0.97                  P(NT MT) = 0.03</p> <p>So, for the new patient, if the laboratory test report shows positive result, let us see if we should declare this as the malignancy case or not.</p> <ul style="list-style-type: none"> <li>• <math>P(h1   PT) = \frac{P(PT h1) \cdot P(h1)}{P(PT)}</math>  <math>= \frac{P(PT MT) \cdot P(MT)}{P(PT)}</math> ..... from (1) – slide 51  <math>= \frac{0.98 \cdot 0.005}{0.0049} = 0.0049 = 0.49\%</math></li> <li>• <math>P(h2   PT) = \frac{P(PT h2) \cdot P(h2)}{P(PT)}</math>  <math>= \frac{P(PT !MT) \cdot P(!MT)}{P(PT)}</math>  <math>= \frac{0.02 \cdot 0.995}{0.0049} = 0.0199 = 1.99\%</math></li> <li>• As P(h2 PT) is higher than P(h1 PT), it is clear that the hypothesis h2 has more probability of being true.</li> <li>• So, <math>h_{MAP} = h2 = !MT</math>.</li> </ul>	<p>Total Marks: 10M</p> <p>Writing the required numerical in proper order</p> <p>5M</p> <p>Applying Bayes theorem and correct answer</p> <p>5M</p>	10 Mins



Roll No.

**PRESIDENCY UNIVERSITY  
BENGALURU**

**SCHOOL OF ENGINEERING**

**TEST – 2**

**Sem & AY:** Odd Sem. 2019-20

**Date:** 16.11.2019

**Course Code:** CSE 319

**Time:** 9:30 AM to 10:30 AM

**Course Name:** MACHINE LEARNING

**Max Marks:** 40

**Program & Sem:** B.Tech (CSE) & VII

**Weightage:** 20%

**Instructions:**

- (i) All questions are compulsory
- (ii) Write answers in brief clearly stating assumptions & reasons
- (iii) Draw figures wherever necessary

**Part A [Memory Recall Questions]**

**Answer all the Questions. Each sub Question carries one marks (10Qx1M=10M)**

1. (C.O.NO. 2) [Comprehension]
- I. The probability of observing data points in cluster-holes is best given by -  
(a) 0 (b) 0.1 (c) 0.99 (d) 0.5
  - II. Given big data set, it is easier to perform which as first-step and then second-step -  
(a) Clustering-SVM (b) K-Means - k-NN (c) Regression-Clustering  
(d) Hierarchical Clustering
  - III. SVM is typically given to perform which of following on data points -  
(a) max. (min. distance) (b) data transformation (c) boundary max.  
(d) a, b, c
  - IV. Seed values in K-means clustering are random selections based on -  
(a) application (b) training data (c) side information (d) a, c
  - V. In k-NN, argument k refers to –  
(a) training size (b) attribute size (c) max dim. of Voronoi search (d) b, c

- VI. Which of following is not Regression type learning model -  
 (a) Bivariate (b) Polynomial (c) Multi-variate (d) Hyper-plane
- VII. If  $P(A) = 0$ ,  $P(A^c) = 1$  then the binary Entropy is –  
 (a) 1 (b) 0.66 (c) 0.721 (d) 0
- VIII. Overfitting is avoidable in Decision Tree regression by -  
 (a) Parsing (b) Backtracking and Pruning (c) limiting depth (d) b, c
- IX. Entropy captures which stopping or split criterion used in feature selection -  
 (a) Purity (b) Uncertainty (c) Impurity (d) b, c
- X. A higher degree polynomial regression may learn with -  
 (a) overfitting (b) covering noise (c) underfitting (d) a, b

### Part B [Thought Provoking Questions]

**Answer both the Questions. Each Question carries ten marks. (2Qx10M=20M)**

2. Urns A, B, C have Apples, Oranges and Bananas distributed randomly in urns. Illustrate decision tree-based data model (regression) given - attributes useful in problem definition are: {color, shape, weight, taste} and attribute is permitted up to 3 values of splittings. Discuss Random forest argued as alternate model for above problem. (C.O.NO.2) [Comprehension]
3. Explain bivariate linear regression model which employs age and height attributes of human population. Draw and formulate notion of following for general linear regression model – underfitting, overfitting, noise and bias. Compare effect of Outliers in – SVM Vs. Linear regression Vs. k-NN. (C.O.NO.2) [Comprehension]

### Part C [Problem Solving Questions]

**Answer the Question. The Question carry ten marks. (1Qx10M=10M)**

4. A population of data points show total assignment of 8 := '+' and 10 := '-' points on student performance. This data is split using 'cgpa' criteria/ attribute giving high, medium and low cgpa clusters with following counts in each: high (+) = 4, high(-) = 2; medium (+) = 4, medium (-) = 3; low(+) = 0, low (-) = 5. Compute the entropy of population, entropies of three clusters and information gain from using original population Vs. three clusters. Is there an alternate purity measure other than entropy-based. (C.O.NO.2) [Application]





## SCHOOL OF ENGINEERING

Semester: VII (odd semester, 2019-20)

Course Code: CSE 319

Course Name: Machine Learning

Date: 16-11-2019

Time: 1 hour

Max Marks: 40

Weightage: 20 %

### Extract of question distribution [outcome wise & level wise]

Q.NO	C.O.NO	Unit/Module Number/Unit /Module Title	Memory recall type [Marks allotted] Bloom's Levels			Thought provoking type [Marks allotted] Bloom's Levels			Problem Solving type [Marks allotted]			Total Marks
			K			C			A			
1	CO2	2, 3	i, iii, iv, v, vi	5	L1	ii, viii, ix, x	4	L2	vii	1	L3	10
2	CO2	2	2	5	L1	2	5	L2	-	-	-	10
3	CO2	2	3	3	L1	3	4	L2	3	3	L3	10
4	CO2	2	-	-	-	-	-	-	-	10	L3	10
	Total Marks			13	L1		13	L2		14	L3	40



K = Knowledge Level C = Comprehension Level, A = Application Level

Note: While setting all types of questions the general guideline is that about 60%

Of the questions must be such that even a below average students must be able to attempt, About 20% of the questions must be such that only above average students must be able to attempt and finally 20% of the questions must be such that only the bright students must be able to attempt.

## Annexure- II: Format of Answer Scheme



### SCHOOL OF ENGINEERING

#### SOLUTION

Date: 16-11-2019

Semester: VII, Odd Semester (2019-20)

Time: 1 hour

Course Code: CSE 319

Max Marks: 40

Course Name: Machine Learning

Weightage: 20 %

#### Part A

(10 Q x 2 M = 20 Marks)

Q No	Solution	Scheme of Marking	Max. Time required for each Question
1. i, ii, iii, iv, v, vi vii, viii, ix x	(a), (d), (d)  (d), (c), (d)  (d), (d), (d)  (d)	<u>1 Mark</u> for each correct choice from multiple choices.	2 min. per Q. = 20 min.



**Part B**

(1 Q x 10 M = 10 Marks)

Q No	Solution	Scheme of Marking	Max. Time required for each Question
2.	<p>For formulating and arguing on Decision Tree model using attributes till level 1,2 from parent label of Urn → Urns: A, B, C → Fruit → 4 attributes in level 3, again 3 splits in level 5 for each attribute.</p> <p>Give merits of R. Forest, show original D. Tree split into color, shape (one-tree) and (weight, taste) as another tree that is subjective yet can be used with majority voting, since trees are expected to cover uncorrelated data points given features.</p>	<p><b>10 Marks</b> with following division:</p> <p>7 Marks for formulating and arguing on D. Tree model.</p> <p><b>3 Marks</b> for discussing /comparing Random Forest.</p>	<b>10 min</b>
3.	<p>Develop 2-variable (age &amp; height) model with scatter of points shown on graph. Propose linear equation form with scaling, shifting constants and noise.</p> <p>Show line and curvilinear depiction including most data points. Argue how small size gives underfitting &amp; bias, large size gives high variance or overfitting. Scatter in data points denotes noise. Bias is covered by intercept or constant. Explain in a line the four terms. Introduce outliers and show how decision regions avoid outlier in each – SVM (best), k-NN and Regression.</p>	<p><b>10 Marks</b> with following division:</p> <p>3 Marks</p> <p>4 Marks</p> <p>3 Marks</p>	<b>10 min</b>
<b>Part C</b>			
4.	<p>- generate probabilities and use [summation (-pi. log pi)] to compute four different value of entropies.</p> <p>- use standard result on Information gain to compute from total entropy of initial state minus total entropy of branches.</p> <p>Entropies: total =0.99, high=0.92, low=0.99, med = 0. Info. Gain = 0.3, explain in a line - max. arg Pj (Ci); ith cluster, jth class</p>	<p><b>10 Marks</b> with following division:</p> <p>- 2 Marks each to solution showing entropies = <b>8 Marks</b></p> <p>- <b>2 Marks</b></p>	<b>10 min</b>





Roll No																			
---------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

**PRESIDENCY UNIVERSITY  
BENGALURU**

**SCHOOL OF ENGINEERING**

**END TERM FINAL EXAMINATION**

**Semester:** Odd Semester: 2019 - 20

**Date:** 20 December 2019

**Course Code:** CSE 319

**Time:** 9:30 AM to 12:30 PM

**Course Name:** MACHINE LEARNING

**Max Marks:** 80

**Program & Sem:** B.Tech (CSE) & VII (DE-III)

**Weightage:** 40%

**Instructions:**

(i) Read the all questions carefully and answer accordingly.

**Part A [Memory Recall Questions]**

**Answer all the Questions. Each Question carries 1 mark.**

**(20Qx1M=20M)**

**1. Fill in the blanks for the given sentences.**

- i. According to Tom Mitchel's definition of Machine Learning, P refers to \_\_\_\_\_.  
(C.O.No.1) [Knowledge]
- ii. The three phases of Machine Learning are Data Input, \_\_\_\_ and Generalization.  
(C.O.No.1) [Knowledge]
- iii. For Unsupervised Learning, the output is dependent upon the \_\_\_\_\_. (C.O.No.1) [Knowledge]
- iv. When we are trying to predict a real-valued variable, the problem falls under the category of \_\_\_\_\_.  
(C.O.No.2) [Knowledge]
- v. 'What is the probability that a particular object belongs to class 'l' given its observed feature values?' is the format of \_\_\_\_\_.  
(C.O.No.1) [Knowledge]
- vi. When we try to predict a categorical or nominal variable, the technique is known as \_\_\_\_\_.  
(C.O.No.2) [Knowledge]
- vii. A \_\_\_\_\_ is a partitioning of a plane into regions based on distance to points in a specific subset of the plane.  
(C.O.No.2) [Knowledge]
- viii. Ordinal data can be naturally \_\_\_\_\_.  
(C.O.No.2) [Knowledge]
- ix. A \_\_\_\_\_ approach is best used when irrelevant attributes have been removed from the data.  
(C.O.No.2) [Knowledge]
- x. The model learns and updates itself through reward/punishment in case of \_\_\_\_\_.  
(C.O.No.2) [Knowledge]
- xi. \_\_\_\_\_ have the disadvantage of being prone to overfit.  
(C.O.No.2) [Knowledge]
- xii. \_\_\_\_\_ is a line that linearly separates and classifies a set of data. (C.O.No.2) [Knowledge]
- xiii. \_\_\_\_\_ algorithm enables neurons to learn and processes elements in the training set one at a time.  
(C.O.No.3) [Knowledge]

- xiv. \_\_\_\_\_ is a technique for finding similarity groups in data. (C.O.No.2) [Knowledge]  
xv. A neural network having more than one hidden layer is \_\_\_\_\_. (C.O.No.3) [Knowledge]

**2. Write short answers for the following questions.**

- i. List all the Activation functions of Artificial Neural Networks. (C.O.No.3) [Knowledge]  
ii. Which algorithm is called as a Lazy algorithm and why? (C.O.No.2) [Knowledge]  
iii. What is a Bias in Artificial Neural Networks. (C.O.No.3) [Knowledge]  
iv. Who gives the rewards in Reinforcement Learning technique? (C.O.No.3) [Knowledge]  
v. Why is the Backpropagation technique used in Artificial Neural Networks. (C.O.No.3) [Knowledge]

**Part B [Thought Provoking Questions]**

**Answer all the Questions. Each Question carries 10 marks. (3Qx10M=30M)**

3. a. A basket is filled with some fresh fruits, the task is to arrange the same type of fruits at one place without having any prior information about the fruits. Which learning technique is used? Explain and justify the answer. [5 M] (C.O.No.1) [Comprehension]  
b. Explain the concept of Back-Propogation in Artificial Neural Networks. [5 M] (C.O.No.3) [Comprehension]  
4. Describe the steps involved in classification learning. [10 M] (C.O.No.2) [Comprehension]  
5. Explain the working of Artificial Neural Network and explain the layers involved, weights and bias. [10 M] (C.O.No.3) [Comprehension]

**Part C [Problem Solving Questions]**

**Answer all the Questions. Each Question carries 10 marks. (3Qx10M=30M)**

6. Which classification technique is appropriate to decide 'LOAN is to be sanctioned or not?' Explain the related classification technique with appropriate figure. [10 M] (C.O.No.2) [Application]  
7. There two hypothesis: A particular tumour is of malignant type and A particular tumour is non-malignant type. The prior available data are: Only 0.5% of the population has this kind of tumour which is malignant. The laboratory report has some amount of incorrectness as it could detect the malignancy was present only with 98% accuracy, where as it could detect the malignancy was not present correctly only in 97% of cases. This means the test predicted malignancy was present which actually was a false alarm in 2% of the cases and also missed detecting the real malignant tumour in 3% of the cases.

We will calculate how the prior knowledge of the percentage of cancer cases in a sample population and probability of the test result being correct influence the probability outcome of correct diagnosis. [10 M] (C.O.No.2) [Application]



8. For the data given below in Table 1, find the total entropy of 'Job Offered?'.  
 Also, find the Total Entropy and Information Gain for the same, by splitting the data based on 'Communication' and 'CGPA' [10 M] (C.O.No.2) [Application]

CGPA	Communication	Aptitude	Programming Skills	Job Offered?
High	Good	High	Good	Yes
Medium	Good	High	Good	Yes
Low	Bad	Low	Good	No
Low	Good	Low	Bad	No
High	Good	High	Bad	Yes
High	Good	High	Good	Yes
Medium	Bad	Low	Bad	No
Medium	Bad	Low	Good	No
High	Bad	High	Good	Yes
Medium	Good	High	Good	Yes
Low	Bad	High	Bad	No
Low	Bad	High	Bad	No
Medium	Good	High	Bad	Yes
Low	Good	Low	Good	No
High	Bad	Low	Bad	No
Medium	Bad	High	Good	No
High	Bad	Low	Bad	No
Medium	Good	High	Bad	Yes

Table 1: Given Data Set





## SCHOOL OF ENGINEERING

### END TERM FINAL EXAMINATION

#### Extract of question distribution [outcome wise & level wise]

Q.NO	C.O.NO (% age of CO)	Unit/Module Number/Unit /Module Title	Memory recall type	Thought provoking type	Problem Solving type [Marks allotted]	Total Marks
			[Marks allotted] Bloom's Levels	[Marks allotted] Bloom's Levels		
			K	C	A	
1		Module 1,2,3	15			15
2		Module 1,2,3	5			5
3		a. Module 1 b. Module 4		5 5		10
4		Module 2		10		10
5		Module 4		10		10
6		Module 2			10	10
7		Module 1			10	10
8		Module 2			10	10
		<b>Total Marks</b>	<b>20</b>	<b>30</b>	<b>30</b>	<b>80</b>

K = Knowledge Level C = Comprehension Level, A = Application Level

Note: While setting all types of questions the general guideline is that about 60%

Of the questions must be such that even a below average students must be able to attempt, About 20% of the questions must be such that only above average students must be able to attempt and finally 20% of the questions must be such that only the bright students must be able to attempt.

I hereby certify that all the questions are set as per the above guidelines.

Faculty Signature:

Reviewer Commend:

## Format of Answer Scheme



### SCHOOL OF ENGINEERING

#### SOLUTION

**Semester:** Odd Semester: 2019 - 20

**Course Code:** CSE 319

**Course Name:** MACHINE LEARNING

**Program & Sem:** B.Tech & VII (DE-I)

**Date:** 20 December 2019

**Time:** 9:30 AM to 12:30 PM

**Max Marks:** 80

**Weightage:** 40 %

#### Part A

(20Q x 1M = 20Marks)

Q No	Solution	Scheme of Marking	Max. Time required for each Question
1	i. Performance ii. Abstraction iii. Coded algorithms iv. Regression v. Posterior probability vi. Classification vii. Voronoi diagram viii. Ordered ix. Nearest neighbor x. Reinforcement learning algorithm xi. Decision trees xii. Hyperplane xiii. Perceptron xiv. Clustering xv. Deep Neural Network	1 M each for correct answer	15 Mins
2	i. Sigmoid, TanH, ReLu ii. k-NN algorithm, because it starts from beginning without any prior knowledge of about data. iii. Bias: It is an extra input to neurons and it is always 1, and has it's own connection weight. iv. Environment gives the rewards. v. Backpropagation is used to repeatedly adjust the weights so as to minimize the difference between actual output and desired output	1 M each for correct answer	5 Mins

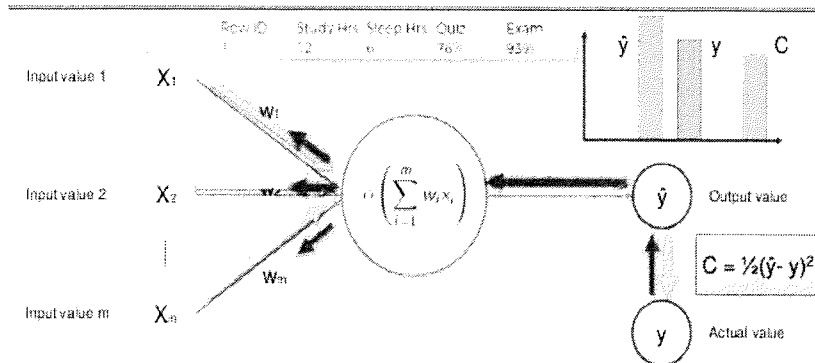
#### Part B

(3Q x 10M = 30 Marks)

Q No	Solution	Scheme of Marking	Max. Time required for each Question
3	a. A basket is filled with some fresh fruits, the task is to arrange the same type of fruits at one place without having any prior information about the fruits. Which learning technique is used? Explain and justify the answer  Unsupervised learning is to be used to solve the problem.	Correct answer with proper justification  5M	15 Mins

b. Explain the concept of Back-Propogation in Artificial Neural Networks.

**Backpropagation**, a procedure to *repeatedly adjust the weights* so as to minimize the difference between actual output and desired output.



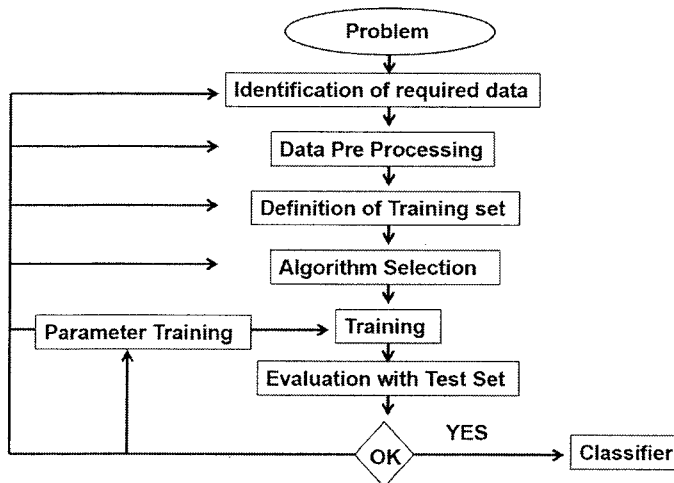
Backpropagation:  
3 M

15 Mins

Diagrams:  
2M

4

Describe the steps involved in classification learning.



Block Dig:  
5M

20 Mins

All phases of learning algorithm carries  
5M

5

The leftmost layer of the network is called the **input layer**, and the rightmost layer the **output layer**

The middle layer of nodes is called the **hidden layer** because its values are not observed in the training set.

Any neural network has 1 input and 1 output layer. The number of hidden layers, for instance, differ between different networks depending upon the complexity of the problem to be solved.

Each of the hidden layers can have a different activation function, for instance, hidden layer1 may use a sigmoid function and hidden layer2 may use a ReLU, followed by a Tanh in hidden layer3 all in the same neural network.

**Weights(Parameters)**—A weight represent the strength of the connection between units. If the weight from node 1 to node 2 has greater magnitude, it means that neuron 1 has greater influence over neuron 2.

**Bias(Offset)**—It is an extra input to neurons and it is always 1, and has it's own connection weight.

Explanation about ANN  
5M

20 Mins

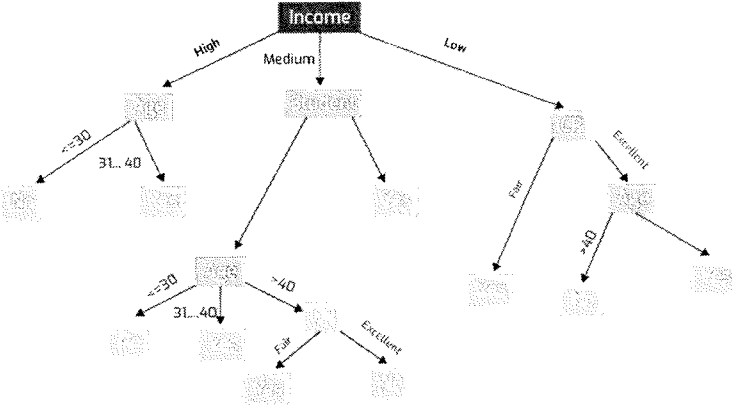
Layers involved  
3M

Weights  
1M

Bias  
1M

Part C

(3Q x 10M = 30Marks)

Q No	Solution	Scheme of Marking	Max. Time required for each Question
6	<p>Which classification technique is appropriate to decide 'LOAN is to be sanctioned or not?'. Explain the related classification technique with appropriate figure.</p> 	<p>Block Dig 6M</p> <p>Explanation 4M</p>	<p>20 Mins</p>
7	<p>Malignant Tumour = MT, Positive Lab Test = PT, Negative Lab Test = NT</p> <p><math>h_1</math> = the particular tumour is of malignant type = MT in our example  <math>h_2</math> = the particular tumour is of not malignant type = !MT in our example</p> <p><math>P(MT) = 0.005</math>                      <math>P(!MT) = 0.995</math>  <math>P(PT MT) = 0.98</math>                      <math>P(PT !MT) = 0.02</math>  <math>P(NT !MT) = 0.97</math>                      <math>P(NT MT) = 0.03</math></p> <p>So, for the new patient, if the laboratory test report shows positive result, let us see if we should declare this as the malignancy case or not.</p> <ul style="list-style-type: none"> <li>• <math>P(h_1   PT) = \frac{P(PT h_1) \cdot P(h_1)}{P(PT)}</math>  <math>= \frac{P(PT MT) \cdot P(MT)}{P(PT)}</math> ..... from (1) – slide 51  <math>= 0.98 * 0.005 = 0.0049 = 0.49\%</math></li> <li>• <math>P(h_2   PT) = \frac{P(PT h_2) \cdot P(h_2)}{P(PT)}</math>  <math>= \frac{P(PT !MT) \cdot P(!MT)}{P(PT)}</math>  <math>= 0.02 * 0.995 = 0.0199 = 1.99\%</math></li> </ul> <p>• As <math>P(h_2 PT)</math> is higher than <math>P(h_1 PT)</math>, it is clear that the hypothesis <math>h_2</math> has more probability of being true.</p> <p>• So, <math>h_{MAP} = h_2 = !MT</math>.</p>	<p>Total Marks: 10M</p> <p>Writing the required numerical in proper order</p> <p>5M</p> <p>Applying Bayes theorem and correct answer</p> <p>5M</p>	<p>20 Mins</p>

8

$$Entropy = \sum_{i=1}^c -p_i * \log_2(p_i)$$

- Pi values for Yes= 8/18 = 0.44 & No= 10/18= 0.56
- Entropy(S) = -0.44 log2(0.44) - 0.56 log2(0.56) = 0.99
- Information gain(S, A) = Entropy(Sbs) - Entropy(Sas)
- Entropy(Sas) =  $\sum_{i=1}^n w_i Entropy(p_i)$

**a) Original data set**

	Yes	No	Total
<b>Count</b>	8	10	18
<b>pi</b>	0.44	0.56	
<b>-pi*LOG(pi)</b>	0.52	0.47	0.99

Total Entropy = 0.99

**b) Splitted data set(based on the CGPA)**

CGPA = High			CGPA = Medium			CGPA = Low			Total
Yes	No	Total	Yes	No	Total	Yes	No	Total	
Count 4	2	6	Count 4	3	7	Count 0	5	5	
pi 0.67	0.33		pi 0.57	0.43		pi 0	1		
-pi*LOG(pi) 0.39	0.53	0.92	-pi*LOG(pi) 0.46	0.52	0.99	-pi*LOG(pi) 0	0	0	

Total Entropy = 0.69

Information Gain = 0.30

Splitted data set(based on 'Communication')

Communication = 'Good'

Communication = 'Bad'

Total Entropy = 0.63

Information Gain = 0.36

Calculating entropy:

5M

20 Mins

Calculating total entropy and information gain:

5M

