

Roll No.....

Dr B R Ambedkar National Institute of Technology, Jalandhar

B Tech 4th Semester (CSE)

CSPC-204, Machine Learning

End Semester Examination, May 2024

Duration: 03 Hours

Max. Marks: 50

Date: 18 May 2023

Marks Distribution & Mapping of Questions with Course Outcomes (COs)										
Question No.	<u>1</u>	<u>2a</u>	<u>2b</u>	<u>3a</u>	<u>3b</u>	<u>4a</u>	<u>4b</u>	<u>5a</u>	<u>5b</u>	
Marks	<u>16</u>	<u>4</u>	<u>4</u>	<u>4</u>	<u>4</u>	<u>4</u>	<u>4</u>	<u>4</u>	<u>4</u>	
CO No.	<u>All</u>	<u>2</u>	<u>3</u>	<u>2</u>	<u>2</u>	<u>2</u>	<u>2</u>	<u>3</u>	<u>4</u>	
Learning Level	<u>1,2</u>	<u>2</u>	<u>3</u>	<u>3</u>	<u>1</u>	<u>3</u>	<u>2</u>	<u>3</u>	<u>2</u>	

1. Attempt all parts, each carry *two* marks

- Differentiate the Hypothesis Search space for FINDS, and candidate elimination algorithm. What are the problems associated with candidate elimination algorithm.
- Compare information gain and Gini index with respect to decision tree.
- What is inductive learning hypothesis? Explain general to specific ordering of hypothesis with an example.
- Create a DT for Boolean function $f(A,B,C) = (A \text{ XOR } B) \wedge (B \text{ XOR } C)$.
- A 4-input neuron has weights 1, 2, 3 and 4. The transfer function is linear with the constant of proportionality being equal to 2. The inputs are 4, 10, 5 and 20 respectively. What will be the output?
- In a neighborhood, 90% children were falling sick due flu and 10% due to measles and no other disease. The probability of observing rashes for measles is 0.95 and for flu is 0.08. If a child develops rashes, find the child's probability of having flu.
- Differentiate margin error and classification error in SVM.
- Define the use of precision, recall and F1-Score.

2. (a) List the benefits of SVM over Logistic Regression. Also, derive the cost function for hard marginal SVM.

(b) Use K-Means Algorithm to find the three cluster centers after the second iteration for eight points (with (x, y) representing locations): A1(2, 10), A2(2, 5), A3(8, 4), A4(5, 8), A5(7, 5), A6(6, 4), A7(1, 2), A8(4, 9). Initial cluster centers are: A1(2, 10), A4(5, 8) and A7(1, 2). Use Manhattan distance as similarity measure.

3. (a) Suppose you have a dataset (as given below) of emails labeled as either "spam" or "not spam" based on the presence of certain keywords. Now your task is to create a Naive Bayes classifier to classify new emails as spam or not spam. Find the label of following email contents "*free offer click here urgent message*".

Email(x)	Label (y)
"get free money"	spam
"get rich quick"	spam
"urgent click here"	spam
"urgent message"	Not-spam
"important information"	Not-spam
"verify your account"	Not-spam

(b) Show the working of single perceptron ANN. Also, show how chain rule of differentiation is used for updating the weights in multilayer perceptron model. Consider a suitable network for illustration.

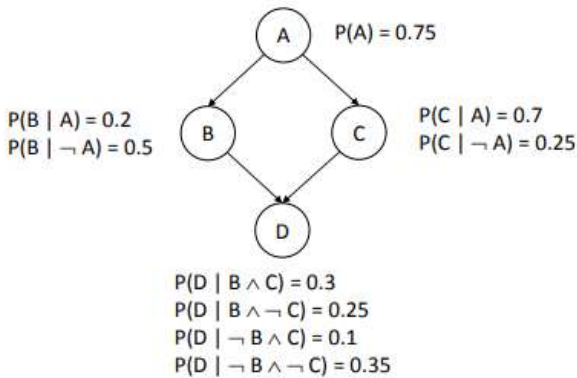
4. (a) How to handle attributes with differing costs in decision tree. Consider the table below and find the information gain of the attributes Weather, Temperature, Humidity and Wind. Find the root node of the tree and give suitable reasons of selecting the attribute with higher information gain as a root node.

Day	Weather	Temperature	Humidity	Wind	Play?
1	Sunny	Hot	High	Weak	No
2	Cloudy	Hot	High	Weak	Yes
3	Sunny	Mild	Normal	Strong	Yes
4	Cloudy	Mild	High	Strong	Yes
5	Rainy	Mild	High	Strong	No
6	Rainy	Cool	Normal	Strong	No
7	Rainy	Mild	High	Weak	Yes
8	Sunny	Hot	High	Strong	No
9	Cloudy	Hot	Normal	Weak	Yes
10	Rainy	Mild	High	Strong	No

(b) What criteria can be used for selecting an activation function in a deep neural network? Justify your answer by providing some suitable

examples. Compare the ReLU, Sigmoid, and tanh activation functions.

5. (a) What is a Bayesian belief network? Explain the concept of conditional independence by considering an example Bayesian belief network. Find the joint probability of $P(A, \neg B, \neg C, D)$ by considering the following network.



- (b) Evaluate the limitations of linear regression for classification tasks. Additionally, demonstrate how logistic regression improves classification performance compared to linear regression. Provide the cost function used in logistic regression.