Question Paper

Exam Date & Time: 22-Jun-2024 (02:30 PM - 05:30 PM)

MANIPAL ACADEMY OF HIGHER EDUCATION

MIT MPL - BTech VI Semester - End Semester Makeup Examination - July 2024

MACHINE LEARNING [ICT 4032]

Marks: 50 **Duration: 180 mins.**

Descriptive

Answer all the questions.

* Answer all questions.
* Assume the missing data suitably.
* Write neatly and legibly.
* Give suitable examples wherever necessary.

1) How does the Parzen window estimator or kernel density function differ from histogram-based methods in estimating the density function of a dataset? Explain the basic concept of using Parzen window estimator and how it provides a smoother density estimate. Utilize the given dataset [1, 1.5, 3, 5] with a bin width of h=2 and bin ranges B1=[1:3], B2=[3:5], B=[5:7] to illustrate the application of Parzen window estimator. (5)

2) Compare and contrast supervised, unsupervised, and reinforcement learning techniques in the context of machine learning. Provide detailed examples showcasing the applications, advantages, and limitations of each method. (3)

3) Illustrate the distinction between univariate linear regression and multiple linear regression. Provide a concise explanation highlighting the number of predictor variables used in each regression model and their impact on prediction accuracy. (2)

4) Consider a Bayesian network B with boolean variables given in**Fig Q2A**. (5)

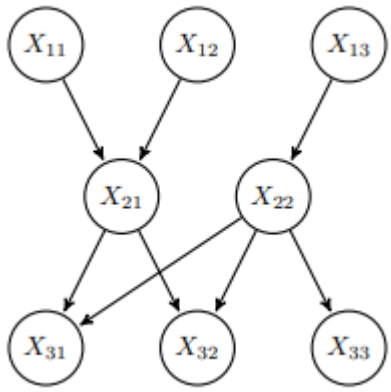

Figure Q2A: Bayesian network.

(a) Are there any variable(s) conditionally independent of X33 given X11 and X12? If so, list all.

(b) Are there any variable(s) conditionally independent of X33 given X22? If so, list all.

(c) Write the joint probability (P(X11, X12, X13, X21, X22, X31, X32, X33) factored according to the Bayes net. How many parameters are necessary to define the conditional probability

distributions for this Bayesian network?

(d) Write an expression for ($P(X_{13} = 0, X_{22} = 1, X_{33} = 0)$ in terms of the conditional probability distributions given in your answer to part (c).

5) Given the table **Q2B** representing the relationship between the mass (y in grams) of a chemical and the time (x in seconds) for which the chemical reaction has been occurring, determine the equation of the regression line to model this relationship. (3)

Table Q2B: Dataset

Time, x(second)	5	7	12	16	20
Mass, y(grams)	40	120	180	210	240

6) Illustrate the working of decision tree induction algorithms used in machine learning. Outline the step-by-step process for constructing decision trees, including the selection of splitting criteria and handling of nodes. (2)

7) Given dataset in **Table Q3A** representing 1000 fruits with binary features (Long, Sweet, Yellow) and three possible classes (Banana, Orange, Other), apply the Naive Bayes classifier to predict the class of a new fruit with features (not long, sweet, yellow). (5)

Table Q3A: Data values

Fruit	Long(x1)	Sweet(x2)	Yellow (x3)
Orange	0	1	0
Banana	1	0	1
Banana	1	1	1
Other	1	1	0
---	---	---	---

Type	**Long**	**Not Long**	**sweet**	**Not Sweet**	**Yellow**	**Not Yellow**	**Total**
Banana	400	100	350	150	450	50	500
Orange	0	300	150	150	300	0	300
Other	100	100	150	50	50	150	200
Total	500	500	650	350	800	200	1000

8) State the main differences between agglomerative and divisive clustering algorithms. (3)

9) Describe the concept of competitive learning in neural networks. Explain how neurons compete to become activated in competitive learning algorithms and provide an example illustrating its application in machine learning. (2)

10) Given **figure Q.4A** containing initial weights, biases, and training inputs/outputs, explain the forward pass technique in the context of a neural network comprising two inputs, two hidden neurons, and two output neurons, each inclusive of a bias. Conduct a forward pass utilizing the given inputs of 0.05 and 0.10 to determine the network's predictions. Following the forward pass, discuss the pivotal role of the backward pass in the training process of the neural network, emphasizing its significance in adjusting weights to minimize prediction errors and enhance overall model performance. Assume sigmoid activation function and loss function as mean squared error. (5)

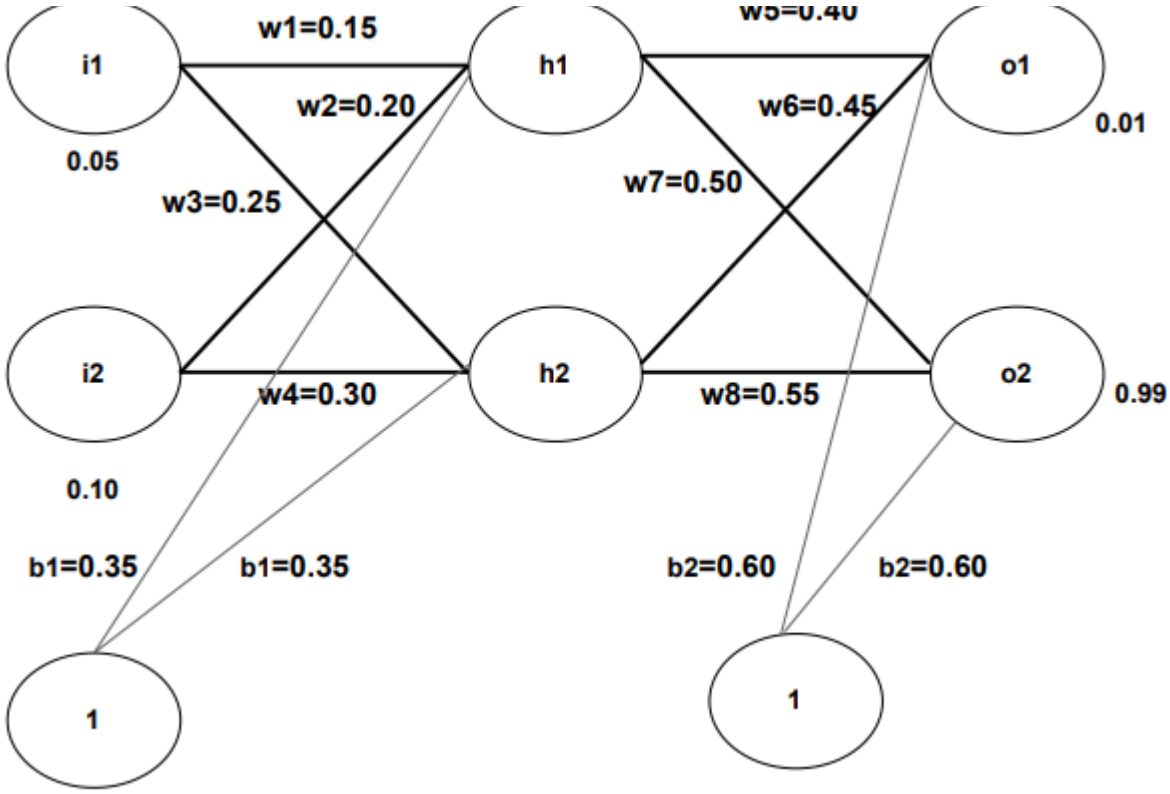

Figure Q4A: Neural network.

11) Compare and contrast hard voting and soft voting in ensemble learning. Provide a detailed explanation of how each method combines the predictions of individual classifiers, along with a numerical example illustrating their differences. (3)

12) How does a Support Vector Machine (SVM) algorithm classify data points in machine learning? Provide an explanation of the key concepts behind SVM and discuss its advantages and limitations in real-world applications. (2)

13) Perform the K-means clustering for K=3, given the data {(2,10), (2,5), (8,4), (5,8), (7,5), (6,4), (1,2), (4,9).}. Initial cluster centers are (2,10), (5,8) and (1,7). Show the convergence for the above data. Justify your answer. (5)

14) Outline the key steps involved in the machine learning pipeline, from collecting data to making predictions. Describe each step in detail and explain its significance in the overall process of developing machine learning models. (3)

15) How does gradient descent contribute to optimizing machine learning models? Explain its role in minimizing the cost function during model training and how this process leads to improved model performance. (2)

-----End-----