B.Tech. DEGREE EXAMINATION, MAY 2023
Fourth Semester

18AIC205J – NEURAL NETWORKS AND MACHINE LEARNING
(For the candidates admitted during the academic year 2018-2019 to 2021-2022)

Note:

(i)   **Part - A** should be answered in OMR sheet within first 40 minutes and OMR sheet should be handed over to hall invigilator at the end of 40th minute.

(ii) **Part - B & Part - C** should be answered in answer booklet.

Time: 3 hours   Max. Marks: 100

PART – A (20 × 1 = 20 Marks)
Answer ALL Questions

		Marks	BL	CO	PO

1. Maria wants to generate new images that are similar to a given set of images. Which type of neural network architecture is suitable for this?   *(Marks 1, BL 1, CO 1, PO 2)*
   (A) MLP   (B) CNN
   (C) RNN   (D) Auto encoder

2. _____ occurs when a model performs well on the training data but poorly on new data, due to memorizing noise or learning the wrong under lying patterns   *(Marks 1, BL 1, CO 1, PO 2)*
   (A) Under fitting   (B) Testing error
   (C) Training error   (D) Over fitting

3. A regularizing techniques commonly used in machine learning to prevent over fitting of neural networks is termed as to be   *(Marks 1, BL 2, CO 1, PO 2)*
   (A) Early stopping   (B) Dropout
   (C) Data augmentation   (D) Best fit

4. One of the commonly used activation functions in perception is   *(Marks 1, BL 1, CO 1, PO 2)*
   (A) Tanh   (B) Rely
   (C) Sigmoid   (D) Softmax

5. A company wants to predict the sales of its product based on the previous year's sales, the number of marketing campaigns, and the price of the product. Which of the following is an advantage of using an MLP in the scenario?   *(Marks 1, BL 2, CO 2, PO 3)*
   (A) The ability to handle non linear relationship between the input variables and the output   (B) The ability to perform clustering analysis
   (C) The ability to handle missing values   (D) The ability to explain the relationship between input and output variables

6. In the back propagation algorithms, what is the purpose of the activation function?   *(Marks 1, BL 2, CO 2, PO 2)*
   (A) To transform the input data into a format   (B) To calculate the output of the neuron
   (C) To update the weights in the network   (D) To introduce non-linearity into the network

7. What is the purpose of the hidden layers in a BPNN?   1 1 2 5
   - (A) To transform the input data into a format
   - (B) To calculate the output of the network
   - (C) To introduce non-linearity into the network
   - (D) To update the weights in the network

8. Two main components of a HMM?   1 1 2 2
   - (A) Hidden state and visible state
   - (B) Hidden state and transition probabilities
   - (C) Visible state and observation probabilities
   - (D) Transition probabilities and observation probabilities

9. Derive the equation for a simple linear regression model   1 2 3 3
   - (A) $y = mx + b$
   - (B) $y = mx$
   - (C) $y = b$
   - (D) $y = x$

10. What is the logistic function in the logistic regression equation?   1 1 3 2
    - (A) $f(x) = e \wedge x$
    - (B) $f(x) = \log(x)$
    - (C) $f(x) = 1 / (1 + e \wedge (-x))$
    - (D) $f(x) = x \wedge 2$

11. The probability formula used in the Naive Bayes algorithm?   1 1 3 4
    - (A) $p(x \mid y) = p(y \mid x)^* p(x) / p(y)$
    - (B) $p(y / x) = p(x \mid y)^* p(y) / p(x)$
    - (C) $p(x \mid y) = p(y)^* p(x)$
    - (D) $p(y / x) = p(y)^* p(x \mid y)$

12. You are using KNN to identify floral species based on petal length and width. Your nearest neighbour count is K = 3. What will the new flower be if its 3 nearest neighbours are all "SETOSA".   1 2 3 5
    - (A) Setosa
    - (B) Versicolor
    - (C) Virginiga
    - (D) Cannot be determined

13. LDA categories hand written digits (0-9). How many decision boundaries are needed to classify data with more than two classes?   1 2 4 3
    - (A) 1
    - (B) 2
    - (C) 3
    - (D) It depends on the number of features

14. PCA reduces the dimensionality of a dataset with 20 features. How many main components can you get from dataset?   1 1 4 3
    - (A) 10
    - (B) 20
    - (C) 5
    - (D) 15

15. You have a dataset of mixed signals from multiple sources, and you use ICA to separate the sources. What is the size of the transformed datasets?   1 2 4 4
    - (A) Same as the original dataset
    - (B) Smaller than the original dataset
    - (C) Larger than the original dataset
    - (D) It depends on the number of sources

16. Imagine a 1000 observation, 20-feature dataset. The dataset if trained using SVM with a radial basis function (RBF) kernel. What is the cause of SVM using a RBF kernel?   1 1 5 4
    - (A) It is computationally expensive
    - (B) It is not suitable for high-dimensional data
    - (C) It may over fit the data
    - (D) It may under fit the data

17. Bagging in ensemble learning means _____.   1   1   5   2
    (A) Training many models on   (B) Averaging predictions from
        subsets of data and combining   different models
        their predictions by the
        majority vote
    (C) Training many models on   (D) Taking the maximum prediction
        distinct data features   from many models

18. How does a random forest handle missing values?   1   1   5   2
    (A) It drops the rows with missing   (B) It fills in the missing values
        values   with the median value of the
                                                 features
    (C) It fills in the missing values   (D) It uses a separate category for
        with the mode values of the   missing values
        feature

19. What is the meaning of the K-value in K-means clustering?   1   1   5   2
    (A) The number of features in the   (B) The number of data points in
        dataset   the dataset
    (C) The number of clusters to be   (D) The number of iterations of the
        formed   algorithm

20. Flower measures, including petal and sepal length and width. Based on   1   2   5   4
    appearance, you wish to classify the flowers. Which is the suitable
    algorithm?
    (A) K-means clustering   (B) Hierarchical clustering
    (C) PCA   (D) LDA

PART – B (5 × 4 = 20 Marks)   Marks   BL   CO   PO
Answer ANY FIVE Questions

21. Consider project X, which predicts subscription-based consumer turnover.   2   1   3
    Your customer dataset includes age, gender, usage history and subscription
    plan. You choose a neural network for the predictive model.
    (i)   Explain how you would preprocess the data before feeding it into   2
          the neural network.
    (ii) Outline the architecture of the neural network you would use for   2
          this project.

22. Consider a weather model that uses an HMM with two hidden states,   4   3   2   4
    "Rainy" and "Sunny" and two possible observations, "Umbrella" and "No
   umbrella". The transition probabilities between the hidden states are given
    by:

	Rainy	Sunny
Rainy	0.7	0.3
Sunny	0.4	0.6

    The observation probabilities for each hidden state are given by

	Umbrella	No-umbrella
Umbrella	0.9	0.1
No-umbrella	0.2	0.8

If the observed sequence is "Umbrella, no umbrella, umbrella", what is the probability that the sequence was generated by the model?

	Marks	BL	CO	PO

23. A company wants to analyze the factors that influence employee satisfaction. They have collected data on employee tenure, salary, job title, and other factors. Use logistic regression to identity the most significance predictors of employee satisfaction.   `4  2  3  4`

24. A university is considering adding an AI course. They choose options A and B. The training can be offered online or in person. The institutions seeks the choice that maximizes the course enrollment. Option A: Accommodates 500 students for $50,000, if the course is offered online, the university expects to enroll 300 students, 400 students or 500 students. Option B: If the course is offered in person for cost of $70,000 for 700 pupils. Courses had a 50%, chance of enrolling 400 students, 30 % chance of enrolling 500 students, and 20% chance of enrolling 700 students. Construct a decision tree to help the university to decide which option to choose.   `4  3  3  4`

25. Consider the scatterplot where the quality of a certain food is studied with varying temperature. Discuss on the risks involved while adding a new dimension viz. Time to the model by making it as a 3D scatterplot.   `4  3  1  4`

Scatter plot of Quality vs Temp
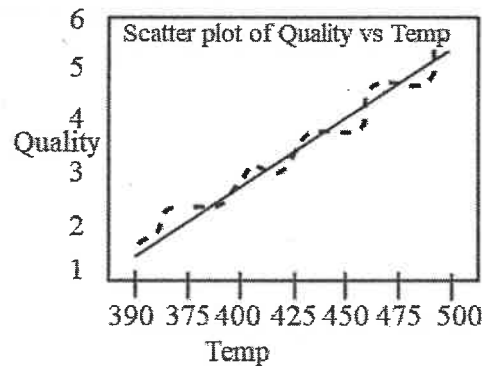
26. Identity the machine learning algorithm making use of the bagging technique. Also, discuss the various steps involved in the algorithm with a neat illustration.   `4  2  5  4`

27. Discuss briefly the steps involved in Viterbi algorithm with relevant examples.   `4  2  4  3`

PART – C (5 × 12 = 60 Marks)
Answer ALL Questions

	Marks	BL	CO	PO

28. a. Design the single layer perceptron with two iteration. For AND Boolean function, the perceptron has starting weights $w1 = 0.2$, $w2 = 0$, learning rate $\alpha = 0.2$, and bias $= 0.4$. Step function $f(x)$ produces 0 or 1, as the activation function, $f(x)$ outputs 1 (or) 0 depending on its values.   `12  3  1  3`
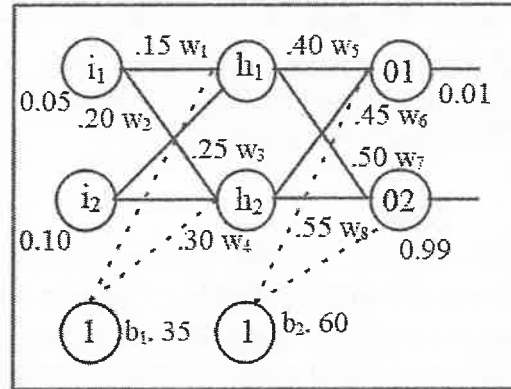
(OR)

b. Discuss the problems related to regeneralization in machine learning and how they can be addressed. Illustrate with suitable examples and diagrams. 12 2 1 4

29. a. Design and develop the back propagation algorithm using the given multi-layer perceptron. 12 3 2 3

(OR)

b. Discuss the Hidden Markov Models (HMMs) and their real-world applications. Give a dataset to demonstrate HMMs in your chosen scenario. 12 2 2 4

30. a. What is meant by linear regression? And illustrate how linear regression can be used for curve fitting when we have the following data points: 12 2 3 4

X	1	2	3	4	5
Y	3	5	7	9	11

(OR)

b. Consider the following table - it consists of the heights, age, weights (target) value for 10 people. As you can see, the weight value of ID11 is missing. We need to predict the weight of this person based on their height and age using Naïve Bayes algorithm. Explain the steps of Naïve Bayes algorithm to predict the missing weight of value of IDII. 12 2 3 3

ID	Heights	Age	Weight
1	5	45	77
2	5.11	26	47
3	5.6	30	55
4	5.9	34	59
5	4.8	40	72
6	5.8	36	60
7	5.3	19	40
8	5.8	28	60
9	5.5	23	45
10	5.6	32	58
11	5.5	38	?

31. a. Compute the principal component using PCA algorithm where the given data = {2, 3, 4, 5, 6, 7; 1, 5, 3, 6, 7, 8} along with its steps. 12 2 4 4

(OR)

b. Explain the working principle of Support Vector Machines (SVMs). How does the maximum margin classifier work in SVMs, and what are the various types of regression in SVMs? 12 2 4 5

32. a. Cluster the following eight points (with (x, y) representing locations), Into three clusters: $A_1(2, 10)$, $A_2(2, 5)$, $A_3(8, 4)$, $A_4(5, 8)$, $A_5(7, 5)$, $A_6(6, 4)$, $A_7(1, 2)$, $A_8(4, 9)$. Initial cluster centers are; $A_1(2, 10)$, $A_4(5, 8)$ and $A_7(1, 2)$. Use K-means algorithm to find the three cluster centers after the seconds iteration. Explain the steps to be followed. 12 3 5 4

(OR)

b. Consider the below sample data set. In this dataset, four predictors are used. These characteristics predict cardiac disease. Create a random forest algorithm that predicts heart the diseases using the below data 12 2 5 3

Blood flow	Blocked arteries	Chest pain	Weight	Cardiac disease
Abnormal	No	No	130	No
Normal	Yes	Yes	195	Yes
Normal	No	Yes	218	No
Abnormal	Yea	Yes	180	Yes

* * * * *