

B.E. Information Technology Fourth Year First Semester

Session: 2021-22, ODD Semester

Subject: Machine Learning

Full Marks - 70

Time - 3 Hours

CO1 (10 Marks)

1. State the different types of machine learning algorithms. [2]
2. When to use regression. [2]
3. Mention the application each that is best solved using: (a) K-nearest neighbor (KNN) algorithm and (b) Support Vector Machine (SVM). [2]
4. How does ensemble of classifiers improve the accuracy of a classification system? [2]
5. Justify the necessity for dimensionality reduction in the context of machine learning. [2]

CO2 (10 Marks)

6. Use the K-Means algorithm to cluster the following 8 points into 3 clusters: $A_1 = (2, 10)$, $A_2 = (2, 5)$, $A_3 = (8, 4)$, $A_4 = (5, 8)$, $A_5 = (7, 5)$, $A_6 = (6, 4)$, $A_7 = (1, 2)$, $A_8 = (4, 9)$. Suppose that the initial seeds (centers of each cluster) are A_1 , A_4 and A_7 . Run the K-Means algorithm for 2 epochs. At the end of each epoch show:
 - i. The new clusters (i.e. the points belonging to each cluster); [3]
 - ii. The centers of the new clusters; [3]
 - iii. Draw a 10 by 10 space with all the 8 points and show the clusters after the second epoch and the new centroids. [4]

CO3 (10 Marks)

7. You are asked to solve the problem of identifying fake bloggers. A fake blogger is one who has posted at least 10% of fake blogs. As you are aware, blogs can be in multiple languages and individual blogs can be in a mix of language. You have been given the identifiers of nearly 1000 bloggers and you have been told that at least

20% of the bloggers post fake blogs. You decided to solve the problem using a machine learning approach.

Answer the following questions in this context:

- i. What are the features that you would consider? Name them and justify their relevance. [2]
- ii. What approach do you use to extract these features? Discuss how you store them. [2]
- iii. Determine the appropriate machine learning algorithm and construct a flow diagram depicting how you would determine the fake bloggers. [4]
- iv. Discuss the metrics used to verify your algorithm. [2]

CO4 (20 Marks)

8. State 2 pros and cons of using deep learning. [2]
9. Draw and explain the architecture of convolutional network. [8]
10. What Is the Difference Between a Feed-forward Neural Network and Recurrent Neural Network? [4]
11. What is an Auto-encoder? [2]
12. What do you understand by Restricted Boltzmann Machine (RBM)? [2]
13. What are the issues faced while training in Recurrent Neural Networks? [2]

CO5 (20 Marks)

14. Elaborate on the various issues like control learning, control policies, Q-learning and convergence in reinforcement learning. Give a suitable application. [10]
15. Write the difference between supervised learning and reinforcement learning. [3]
16. What are the types of reinforcement learning? [2]
17. What is a Markov Decision Process? [2]
18. How does Monte Carlo algorithms work? [3]

-----X-----

CO1: Explain concepts of Machine Learning paradigm.

CO2: Comprehend mathematical analysis and various theories of machine learning approaches.

CO3: Choose and **experiment** different types of ANN and to get an insight of when to apply a particular approach for solving problems.

CO4: Express and **illustrate** Deep Learning frameworks and its applications.

CO5: Describe and **illustrate** methods of Reinforcement Learning.
