Qn. Set Code-7

Semester: 5th
Programme: B.Tech
Branch: CSSE

AUTUMN END SEMESTER EXAMINATION-2024

5th Semester B.Tech

MACHINE LEARNING

CS31002

(For 2023 (L.E), 2022 & Previous Admitted Batches)

Time: 2 Hours 30 Minutes | Full Marks: 50

Answer any FIVE questions.
Question paper consists of two SECTIONS i.e. A and B.
Section A is compulsory.
Attempt any Four question from Sections B.
The figures in the margin indicate full marks.
Candidates are required to give their answers in their own words as far as practicable and all parts of a question should be answered at one place only.

SECTION-A

1. Answer the following questions. [1 × 10]

(a) A Machine Learning Engineer observed that some features in the dataset are highly correlated. He applied multiple linear regression for prediction purposes. What kind of result would be expected?

(b) If $\alpha > 0$ is a constant, the number of features is P, the number of training data is m, and $\beta = [\beta_0\ \beta_1\ \beta_2 \ldots \beta_P]^T$ then compute the regularization factor in LASSO regression.

(c) There are 60 data points in a binary classification problem, out of which 24 belong to class Positive and 36 belong to class Negative. Calculate the entropy of the data set.

(d) Which one is less sensitive to outliers between normalization and standardization, and why?

(e) In a regression task, if the true values are [3, 5, 2, 8] and the predicted values are [2.5, 5.5, 2, 7.5], calculate the Mean Absolute Error (MAE).

(f) The activation function plays a crucial role in a neural network. State (Yes/No). Justify either for Yes or for No.

(g) Given two clusters with centers C1(2, 3) and C2(5, 7), calculate the squared Euclidean distance between the two centers.

(h) Which of the following are real world applications of SVM?

i. Text & hypertext categorization

ii. Image Classification

iii. Clustering of new articles

iv. All of the above.

(i) Given a dataset with an accuracy of 84% and 200 samples, how many samples were incorrectly classified?

(j) Given a neural network with one hidden layer containing 5 neurons, if each neuron has 3 inputs, how many weights are there in the network, excluding biases?

SECTION-B

2. (a) State the multiple linear regression model, and write the algorithm for computing the coefficients in this model mentioning what the symbols and notations represent and explaining the steps. [5]

(b) Given data set : [5]

Value	15	18	21	24	30

(i) Calculate the mean and standard deviation of the data set.

(ii) Apply Z-Score normalization to the data set.

3. (a) Define and discuss the key differences between the training set and the test set. [5]

(b) Consider the following dataset about animals. Determine each attribute's information gain. Indicate which characteristic, depending on the information gain parameter, ought to be utilized as the first splitting node. [5]

Animal	Has Fur	Lays Eggs	Can Fly	Class
Dog	Yes	No	No	Mammal
Cat	Yes	No	No	Mammal
Eagle	No	No	Yes	Non-Mammal
Penguin	No	Yes	No	Non-Mammal
Platypus	Yes	Yes	No	Mammal
Bat	Yes	No	Yes	Mammal
Ostrich	No	No	No	Non-Mammal
Frog	No	Yes	No	Non-Mammal

4. (a) Discuss the concept of bias-variance tradeoff in the context of underfitting and overfitting. How does it relate to model performance? [5]

(b) Design individual single perceptron models that can be used to perform Boolean operations AND, OR with activation function uses step function and has bias too. [5]

5. (a) Given a binary classification problem data set with the following information where numbers indicates total samples [5]

Class / Feature 'A'	Feature 'A' = 1	Feature 'A' = 0
Class '1' (Positive)	**25**	**15**
Class '0' (Negative)	**30**	**30**

(i) Calculate the prior probabilities of P(Class=1), P(Class=0) and also calculate the likelihood of P(Feature A=1|Class=1) and P(Feature A=1|Class=0)

(ii) Using the same dataset as above, to predict the class of a new sample where Feature A = 1 calculate the posterior probabilities P(Class=1|FeatureA=1) and P(Class=0|FeatureA=1) using Bayes' theorem. Determine which class the model predicts.

(b) Match words and phrases in column A with the closest related meaning/word(s)/phrase(s) in column B. [5]

Column A		Column B	
1	It is one of an example of clustering techniques	i	Support Vector Machine
2	It is an iterative optimization algorithm to find the minimum value(local optima) of a function.	ii	Numpy
3	It is a technique that performs classification by finding the hyperplanes that maximizes the margin between the two classes.	iii	OpenCV
4	These are accessed via keys and not via their positions in Python.	iv	Dictionary
5	It is a Python data visualization library which provides a high level interface for drawing attractive and informative statistical graphics.	v	ReLU
6	This is an open source library which provides high performance, easy-to-use data structures and data analysis tools.	vi	Seaborn
7	It is an library design to solve computer vision problems.	vii	Tensorflow

8	It is an open source software library used to design, built and train deep learning models.	viii	Pandas Library
		ix	Gradient Decsent
		x	K-Means
		xi	Ensemble

6. (a) How does the regularization parameter/ hyper-parameter 'C' affect the SVM model's performance? Discuss its role in the trade-off between maximizing the margin and minimizing classification errors. [5]

(b) Perform KNN classification on the following dataset and predict the class for (brightness=20 and saturation=35) with K=5 using Euclidean distance. [5]

Brightness	Saturation	Class
40	20	Red
50	50	Blue
60	90	Blue
10	25	Red
70	70	Blue
60	10	Red
25	80	Blue

7. (a) Use Principal Component Analysis to decrease the dimensions of the provided data from 2(two) to 1(one). [5]

X	2	3	5	4	6
Y	3	4	7	6	8

(b) Given a neural network with: [5]

- **Input Layer**: 2 neurons
- **Hidden Layer**: 2 neurons with weights:

  $w_{11}= 0.5, w_{12}= 0.3$ (to Hidden Neuron 1)

  $w_{21}= 0.4, w_{22}= 0.6$ (to Hidden Neuron 2)
- **Output Layer**: 1 neuron with weights:

  $w_{h1}= 0.7, w_{h2}= 0.5$
- **Biases**: Hidden neurons have $b_1= 0.2, b_2= -0.1$; output neuron has $b_o= 0.1$
- **Input:** $x_1= 0.6, x_2= 0.9$

Calculate the output of the network using a **sigmoid** activation function.

*****