End-Term Exam
CSN-382
Machine Learning
Indian Institute of Technology Roorkee

Date: 03-May-2024

Name: _____

Roll Number: ____ ____ ____ ____ ____ ____ ____ ____

Note:

1. **Any case of copying will be given ZERO marks without checking.**
2. **Answer all questions in the provided answer-space, in the paper sheet only. Use a rough sheet for a solution.**
3. **No partial marking, only the clearly written right answer will get the full marks.**

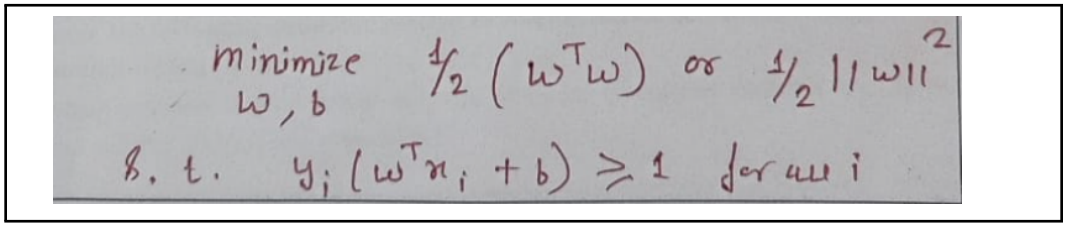
Total Marks - 50 Marks

Que-1. Answer the following questions related to SVM, answer should be to the point and limited within the box. [1+2+2+3+2]

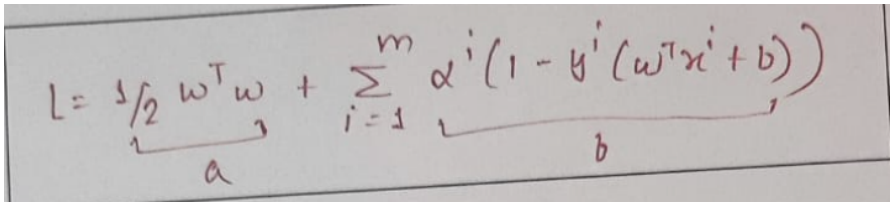
a. Who are the deciding data points among all data points for the final hyper plane?

Support Vector

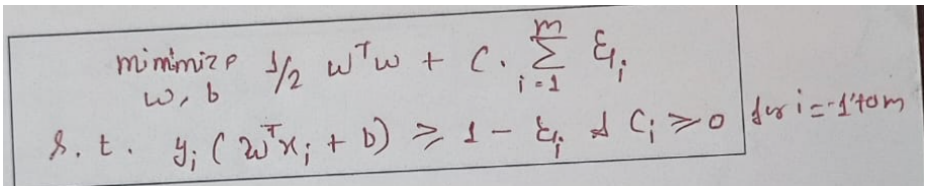
b. Write the equation (L) of optimal margin (m) between the two supporting hyper-planes for the weight vector 'w'.



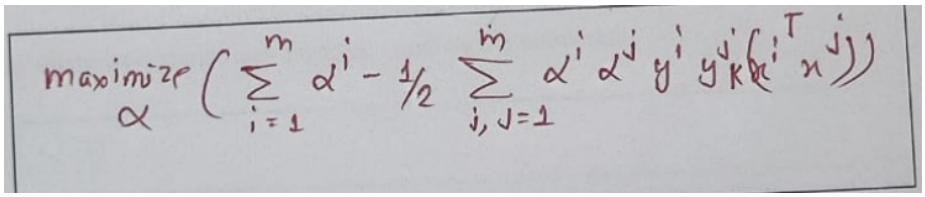
c. Expand 1b such that it represents the equation of Dual optimization problem for hard margin SVM (before applying partial differentiation wrt. W and b).



d. The penalty parameter C in SVM (Support Vector Machine) is a hyperparameter that controls the trade-off between maximizing the margin and minimizing the classification error. Write the optimization function of soft-margin SVM with C, and point out a situation where C can be tweaked to get the hard-margin optimization function.



e. Write the final hard-margin SVM's dual optimization equation in terms of dot product of data points with kernel trick.

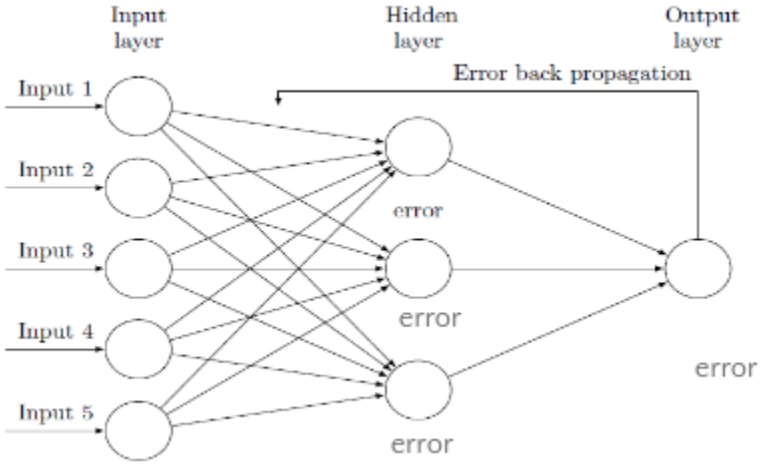


Que-2. Answer the following questions related to Neural Networks, answer should be to the point and limited within the box. [1+2+2+2+1+1+1]

a. In which scenario a deep neural network of linear layers can outperform a traditional linear classifier like logistic regression?

Multiple layers brings non-linearity in the classification boundary.

b. Frame the following and answer the related questions for the given neural network.



B1. How many weight matrices are to be tuned by the network?

2

B2. What is the dimension of each matrix to be tuned by the network?

3*5 and 1*3

B3. If the output at the output layer has a sigmoid activation function, and the actual output value is Y, what will be the loss function?

Y, what will be the loss function? $J(\theta) = 1/m \left[\sum_{i=1}^{m} -y^i \log h_\theta(x^i) + (1-y^i) \log(1-h_\theta(x^i)) \right]$

c. Which one is preferable between ReLu and Sigmoid function at output layer?

Sigmoid

d. In which scenario softmax is preferable over sigmoid at the output layer?

Multi-class classification

e. Which one is preferable between ReLu and sigmoid at hidden layers?

Relu

Que3. Which of the following statements is/are correct? [2+2]

3a.

A. Random Forest is known for its high accuracy because it combines the predictions of multiple decision trees, reducing the risk of overfitting and increasing the model's robustness.
B. Random Forest can be used for both regression and classification tasks, and it can handle a wide range of data types, including numerical, categorical, and binary data.
C. Random Forests are less sensitive to noise and outliers in the data compared to other algorithms.

A, B, C all are correct.

3b. Which of the following statements is/are correct?

A. Random Forests can be computationally expensive, especially for large datasets or many decision trees.
B. Random Forest can be challenging to interpret because it combines the predictions of multiple decision trees.

A and B both are correct.

Que4. For the following dataset answer the following questions. [2+2+2]

Temperature	Humidity	Wind Speed	Cloud cover	Has Rained (Y)
Low	Low	Slow	No	No
Low	High	Fast	Yes	yes
High	Low	Slow	No	No
High	High	Slow	Yes	No
Low	High	Fast	Yes	Yes
High	High	Slow	No	No
Low	High	Fast	Yes	Yes
Low	Low	Fast	Yes	No
High	Low	Slow	No	No
High	High	Slow	No	No

E(X) represents the value of entropy of feature X.

4a. Which of the following statements is correct? Write ‘Correct’ behind the option.

1. E(Temperature) < E(Humidity) [ ]
2. E(Temperature) > E(Humidity) [Correct]
3. E(Temperature) = E(Humidity) [ ]

4b. Which of the following statements is correct? Write 'Correct' behind the option.

1. E(Wind Speed) < E(Cloud cover) [Correct]
2. E(Wind Speed) > E(Cloud cover) [ ]
3. E(Wind Speed) = E(Cloud cover) [ ]

4c. Which of the following features would be chosen as the top node or root node in a decision tree. Write 'Correct' behind the option.

1. Temperature [ ]
2. Humidity [ ]
3. Wind Speed [Correct]
4. Cloud cover [ ]

Que5. Consider the following loss table for a model while training and answer the following questions. [2+2 Marks]

Epoch	Training Loss	Validation Loss
1	0.36	0.66
2	0.31	0.54
3	0.23	0.33
4	0.21	0.50
5	0.19	0.55
6	0.17	0.60
7	0.13	0.77

5a. Which of the following correctly explain this.

A) Model is trained perfectly.

B) Model incur underfitting.

C) Model incur overfitting after 3rd epoch. Correct

D) Model incur overfitting after 6th epoch.

5b. Use of validation data while training a model can be useful for. Write 'correct' after the options.

A) Detecting underfitting.

B) Detecting overfitting. Correct

C) Both A and B.

D) None of A and B.

Que6. Consider the following table with column A and B, provide the matching pair for the values of A and B. [2 Marks]

A	B
1) Linear Regression	P) Lasso
2) L-1 Regularization	Q) Ridge
3) Clustering	R) Unsupervised Learning
4) L-2 regularization	S) Supervised Learning

1 → S | 2 → P | 3 → R | 4 → Q

Que7. Consider the following two statements. [2 Marks]

P: Type 1 error, which is also called false negative, where the model incorrectly predicts the positive class.

Q: Type 2 error, which is also called false positive, where the model incorrectly predicts the negative class.

Which of the above statements is/are true?

A) Both P and Q

B) Only P

C) Only Q

D) None of P and Q Correct

Que8. Given below are two statements, one is labelled as Assertion A and the other is labelled as Reason R. [2 Marks]

Assertion A: Machine Learning requires good quality and sufficient data to train and test the algorithm

Reason R: For correct classification, good quality data which is free from noise, and sufficient data is required for training and testing of the algorithm.

In light of the above statements, choose the most appropriate answer from the options given below:

A) Both A and R are correct and R is the correct explanation of A Correct

B) Both A and R are correct but R is NOT the correct explanation of A

C) A is correct but R is not correct.

D) A is not correct but R is correct.

Que9. Consider the following confusion matrix. Answers are expected up to three decimal points. We have four classes called P, Q, R, and S. Answer the following questions using the confusion matrix. [2+2+2+2+2 Marks]

9a. f-score for class P is 0.613-0.623

9b. f-score for class Q is 0.523-0.533

9c. f-score for class R is 0.512-0.522

9d. f-score for class S is 0.562-0.572

9e. Micro f-score for this model is 0.553-0.563