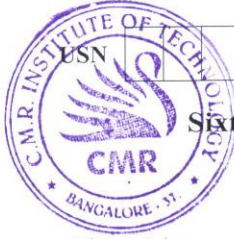


CBCS SCHEME



21AI63

Sixth Semester B.E. Degree Examination, June/July 2024

Machine Learning

Time: 3 hrs.

Max. Marks: 100

Note: Answer any FIVE full questions, choosing ONE full question from each module.

Module-1

- 1 a. Illustrate the basic design issues and approaches to machine learning. (08 Marks)
- b. What is machine learning? Explain with an example why to use machine learning. (04 Marks)
- c. Explain the main challenges of machine learning that can arise when selecting a learning algorithm and training it on data. (08 Marks)

OR

- 2 a. Discuss the broad categories of machine learning system. (10 Marks)
- b. Write Find-S algorithm and apply the same for the given instance.

Location	Price	Size	Condition	Buy House
Urban	Expensive	Large	New	Yes
Sub urban	Moderate	Medium	Old	No
Urban	Moderate	Small	New	Yes
Urban	Moderate	Large	Old	Yes

(10 Marks)

Module-2

- 3 a. Illustrate various methods to prepare the data for machine learning algorithms. (10 Marks)
- b. Explain multilabel classification and multi output classification with code snippet. (10 Marks)

OR

- 4 a. Explain the various performance measure to evaluate a classifier with an example:
 - (i) Measuring accuracy using cross-validation.
 - (ii) Confusion matrix
 - (iii) Precision
 - (iv) Recall
 - (v) The ROC curve. (10 Marks)
- b. Explain how do you frame the problem and choose an appropriate performance measures for a dataset in a machine learning project. (06 Marks)
- c. Explain Grid search method to fine-tune the model. (04 Marks)

Module-3

- 5 a. What is gradient descent. Explain various types of gradient descent with necessary diagrams. (10 Marks)
- b. Show that how SVM make predictions using Quadratic programming and Kernelized SVM. (10 Marks)

Important Note : 1. On completing your answers, compulsorily draw diagonal cross lines on the remaining blank pages.
2. Any revealing of identification, appeal to evaluator and/or equations written eg, 42+8 = 50, will be treated as malpractice.

OR

- 6 a. Explain the following with respect to logistic regression :
(i) Estimating probabilities. (10 Marks)
(ii) Training and cost functions. (10 Marks)
- b. Discuss non-linear SVM classification. How can you use Polynomial Kernel, Gaussian and RBF Kernel? (10 Marks)

Module-4

- 7 a. Explain how decision trees are trained, visualized and used in making predictions. (10 Marks)
b. Explain Bagging and Pasting with an example. (10 Marks)

OR

- 8 a. Explain CART algorithm. Discuss regularization hyper parameters in Decision trees. (10 Marks)
b. What is Boosting? Explain AdaBOOSE and gradient Boosting. (10 Marks)

Module-5

- 9 a. What is Bayes theorem. Describe Brute-force Map learning algorithm. (08 Marks)
b. Discuss the minimum description length algorithm. (08 Marks)
c. Explain the steps of Gibbs algorithm. (04 Marks)

OR

- 10 a. Write EM algorithm and explain in details. (10 Marks)
b. Explain Naïve Bayes classifier with an example. (10 Marks)

Q 2 b)

Find-S algorithm

	Location	Price	size	condition	Buy how
P ₁	Urban	Expensive	Large	New	Yes
P ₂	Suburban	moderate	medium	old	No
P ₃	Urban	moderate	small	New	Yes
P ₄	Urban	moderate	Large	old	Yes

∴ There are 4 attributes so each attribute we initially fill ϕ

Step 1 $\langle \phi \ \phi \ \phi \ \phi \rangle$

Step 2 Generalize first +ve instance

P₁ U E L New ~~Yes~~

h₁ = $\langle U \ E \ L \ N \ \phi \rangle$

Step 3 scan next P₂. It is -ve so exclude it & hypothesis remains same without any change

P₃ - $\langle U \ m \ S \ N \rangle$

h₃ = $\langle U \ ? \ ? \ N \rangle$

Q4 = $\langle U, m, L, O \rangle$

Q4 = $\langle U, ? , ? , ? \rangle$

Q 1 a) Design issues - 04 m
approaches to machine Learning 04 m

b) what is machine learning - definition 02 m
Example with explanation - 02 m

c) main challenges of machine Learning
eight challenges list with 2-3
line explanation ^{for} each challenge
1 challenge - 1 m.

Q 2 a) Broad categories

Q 3 a) Methods to prepare data for machine learning algo

5 methods - 2 m each

b) multilabel classification with code - 5m

multi output classification with code - 5m

Q 4 a) measuring accuracy using

cross-validation - 2m

confusion matrix - 2m

Precision - 2m

Recall - 2m

Roc curve - 2m

b) Framing a problem - 3m

~~selecting~~ appropriate performance

measure for dataset - 3m

c) Grid-search method with code - 4m

Q 59) what is gradient descent - 2m
Types of gradient descent - 8m

b) SVM with Quadratic Programming - 5m
-1- Kernelized SVM - 5m

Q 69) Estimating probabilities - 5m
Training cost function - 5m

b) non-linear SVM classification 4m
polynomial - 2m
(Kernel)
Gaussian - 2m
RBF - 2m

Q 79) Decision tree training - 4m
Visualization - 3m
making prediction - 3m

b) Bagging with example - 5m
Pasting with example - 5m

Q 8 a) CART algorithm - 4m
Hyper parameter regularization in
decision tree - 6m

b) what is Boosting - 2m
Ada Boost - 4m
Gradient Boosting - 4m

Q 9 a) Bayes theorem - 3m
Brute-force map algorithm - 5m

b) minimum description length algo - 6m
example - 2m

c) Steps of Gibbs algo - 4m

Q 10 a) EM algo steps - 5m
explanation of each step - 5m

b) Naive Bayes classifier - 5m
example - 5m