

2021

**MACHINE LEARNING**

*Full Marks: 60*

Time: Two hours

*The figures in the margin indicate full marks for the questions.*

A. Multiple Choice Questions

1 x 20=20

1. Which of these is a reasonable definition of machine learning?
  - a. Machine learning is the science of programming computers.
  - b. Machine learning learns from labeled data
  - c. Machine learning is the field of allowing robots to act intelligently.
  - d. Machine learning is the field of study that gives computers the ability to learn without being explicitly programmed.
2. Let A and B be 3x3 (square) matrices. Which of the following must necessarily hold true? Check all that apply.
  - a.  $A*B*A = B*A*B$
  - b. If A is not a 3x3 identity matrix, then  $A*B = B*A$
  - c.  $A*B = B*A$
  - d.  $A+B = B+A$
3. Which among the following algorithms are used in Machine learning?
  - a. Naive Bayes
  - b. Linear Regression
  - c. Support Vector Machines
  - d. All of the Above
4. What is supervised learning?

- a. All data is unlabelled and the algorithms learn to inherent structure from the input data
  - b. All data is labelled and the algorithms learn to predict the output from the input data
  - c. It is a framework for learning where an agent interacts with an environment and receives a reward for each interaction
  - d. Some data is labelled but most of it is unlabelled and a mixture of supervised and unsupervised techniques can be used.
5. P1: Suppose you are working on weather prediction, and use a learning algorithm to predict tomorrow's temperature (in degrees Centigrade/Fahrenheit).
- P2: Suppose you are working on weather prediction, and your weather station makes one of three predictions for each day's weather: Sunny, Cloudy or Rainy. You'd like to use a learning algorithm to predict tomorrow's weather.
- a. P1 is a classification and P2 is a regression problem
  - b. P1 is a regression and P2 is a classification problem
  - c. Both P1 and P2 are classification problems
  - d. Both P1 and P2 are regression problems
6. What is unsupervised learning?
- a. a) All data is unlabelled and the algorithms learn to inherent structure from the input data
  - b. All data is labelled and the algorithms learn to predict the output from the input data
  - c. It is a framework for learning where an agent interacts with an environment and receives a reward for each interaction
  - d. Some data is labelled but most of it is unlabelled and a mixture of supervised and unsupervised techniques can be used.
7. What is Semi-Supervised learning?
- a. All data is unlabelled and the algorithms learn to inherent structure from the input data
  - b. All data is labelled and the algorithms learn to predict the output from the input data
  - c. It is a framework for learning where an agent interacts with an environment and receives a reward for each interaction
  - d. Some data is labelled but most of it is unlabelled and a mixture of supervised and unsupervised techniques can be used
8. What is Reinforcement learning?

- a. All data is unlabelled and the algorithms learn to inherent structure from the input data
  - b. All data is labelled and the algorithms learn to predict the output from the input data
  - c. It is a framework for learning where an agent interacts with an environment and receives a reward for each interaction
  - d. Some data is labelled but most of it is unlabelled and a mixture of supervised and unsupervised techniques can be used.
9. Choose the options that are correct regarding machine learning (ML). I) when the output variable is a category, such as “red” or “blue” or “disease” and “no disease”.
- II) When the output variable is a real value, such as “dollars” or “weight”.
- a. I) is classification & II) is regression
  - b. II) is classification & I) is regression
  - c. Both I) & II) are classification
  - d. Both I) & II) are regression
10. Which of the following sentence is FALSE regarding regression?
- a. It relates inputs to outputs.
  - b. It is used for prediction.
  - c. It may be used for interpretation.
  - d. It discovers causal relationships
11. In Regression modeling we develop a mathematical equation that describes how, (Predictor-Independent variable, Response-Dependent variable)
- a. one predictor and one or more response variables are related.
  - b. several predictors and several response variables response are related.
  - c. one response and one or more predictors are related.
  - d. All of these are correct.
12. In the mathematical Equation of Linear Regression  $Y = a + bX$  , (a, b) refers to \_\_\_\_\_
- a. (X-intercept, Slope)

- b. (Slope, X-Intercept)
  - c. (Y-Intercept, Slope)
  - d. (slope, Y-Intercept)
13. Which of the following is not numerical functions in the various function representation of machine learning?
- a. a) Case-based
  - b. Neural Network
  - c. Linear Regression
  - d. Support Vector Machines
14. RNNs stands for?
- a. Receives neural networks
  - b. Report neural networks
  - c. Recording neural networks
  - d. Recurrent neural networks
15. Which of the following is/are Common uses of RNNs?
- a. BusinessesHelp securities traders to generate analytic reports
  - b. Detect fraudulent credit-card transaction
  - c. Provide a caption for images
  - d. All of the above
16. Which of the following is well suited for perceptual tasks?
- a. Feed-forward neural networks
  - b. Recurrent neural networks
  - c. Convolutional neural networks
  - d. Reinforcement Learning
17. Which of the following activation function output is zero centered?
- a. Hyperbolic Tangent.
  - b. Sigmoid.
  - c. Softmax

- d. Rectified Linear unit(ReLU).
18. Which of the following is FALSE about Deep Learning and Machine Learning algorithms?
- Deep Learning algorithms work efficiently on a high amount of data.
  - Feature Extraction needs to be done manually in both ML and DL algorithms.
  - Deep Learning algorithms are best suited for unstructured data.
  - Deep Learning algorithms require high computational power.
19. Which of the following is FALSE for neural networks?
- Artificial neurons are similar in operation to biological neurons.
  - Training time for a neural network depends on network size.
  - Neural networks can be simulated on conventional computers.
  - The basic unit of neural networks are neurons.
20. What is an activation value?
- weighted sum of inputs
  - threshold value
  - main input to neuron
  - none of the mentioned

B. Group B

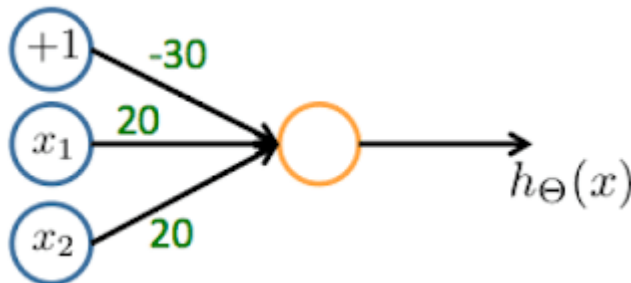
2\*6=12

- Suppose you train a logistic classifier  $h_{\theta}(x) = g(\theta_0 + \theta_1 x_1 + \theta_2 x_2)$ .  
Suppose  $\theta_0 = 6, \theta_1 = -1, \theta_2 = 0$  Draw a figures to represent the decision boundary found by your classifier?
- Imagine you are working on a project which is a binary classification problem. You trained a model on training dataset and get the below confusion matrix on validation dataset.

		<b>Predicted:</b> <b>NO</b>	<b>Predicted:</b> <b>YES</b>
n=165			
<b>Actual:</b> <b>NO</b>		50	10
<b>Actual:</b> <b>YES</b>		5	100

Based on the above confusion matrix, calculate the accuracy.

3. Consider the following neural network which takes two binary-valued inputs  $x_1, x_2 \in \{0, 1\}$  and outputs  $h_{\theta}(x)$ . Which logical function does it (approximately) compute?



4. P1: Suppose you are working on stock market prediction, and you would like to predict the price of a particular stock tomorrow (measured in dollars). You want to use a learning algorithm for this.
- P2: Suppose you are working on stock market prediction. You would like to predict whether or not a certain company will declare bankruptcy within the next 7 days (by training on data of similar companies that had previously been at risk of bankruptcy).
- P3: Suppose you are working on stock market prediction, typically tens of millions of shares of Microsoft stock are traded (i.e., bought/sold) each day. You would like to predict the number of Microsoft shares that will be traded tomorrow.

P1 is a \_\_\_\_\_ problem, P2 is a \_\_\_\_\_ problem and P3 is a \_\_\_\_\_ problem

5. Explain the different layers in CNN.
6. What Are Vanishing and Exploding Gradients?

1. Explain with figure the hypothesis function:  $h_{\theta}(x) = \theta_0 + \theta_1 x$  for
  - (i)  $\theta_0 = 0, \theta_1 = 0.5$
  - (ii)  $\theta_0 = 1, \theta_1 = 0.5$  and
  - (iii)  $\theta_0 = 2, \theta_1 = 0$
2. Explain cost function. Accompany your explanation with a diagram
3. Explain the principle of the gradient descent algorithm. Accompany your explanation with a diagram
4. Differentiate between LMS training rule (normal) and gradient descent for n features.
5. Consider the following training data set with 7 tuple and 2 attributes:

Observation	X1	X2	Class
1	3	4	◆
2.	2	2	◆
3	4	4	◆
4	1	4	◆
5	2	1	●
6	4	3	●
7	4	1	●

- (a) Plot all the seven training points. Are the classes { ◆, ● } linearly separable?
  - (b) Sketch the optimal separating hyperplane and provide the equation for this hyperplane
  - (c) Describe the classification rule for the maximal margin classifier. It should be something along the lines of “Classify to Red if  $\beta_0 + \beta_1 X_1 + \beta_2 X_2 > 0$ , and classify to Blue otherwise.” Provide the values for  $\beta_0$ ,  $\beta_1$  and  $\beta_2$ .
  - (d) Indicate the support vectors for the maximal margin classifier.
6. You are given the neural networks as shown fig 1(a) which take two binary valued inputs  $x_1, x_2 \in \{0, 1\}$  and the activation function is shown in figure 1(b). Which of the logical functions does it compute? Assume  $b=+1$

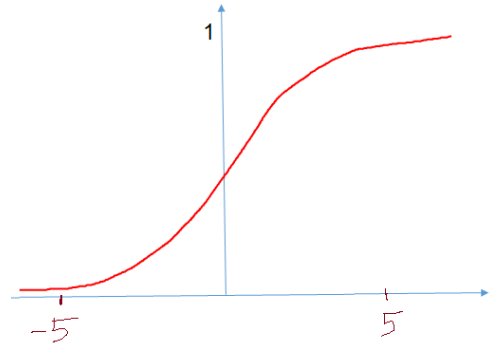
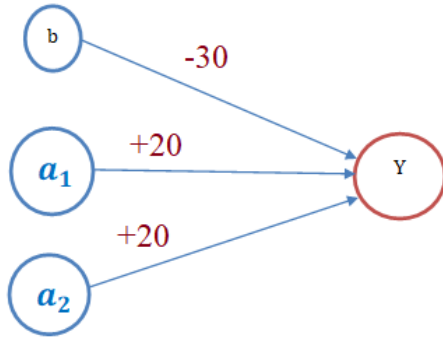


Figure 1. (a)

(b)

7. Design a neural network to realize XNOR function

\*\*\*\*\*