

Your Name: \_\_\_\_\_

Your ID: \_\_\_\_\_

# of Questions: 15

Date and Time of Exam Creation: Fri, Feb 23, 2024 @ 20:54:43

Total Exam Points: 43.00

---

Question #: 1

We have a collection of 1,000 images from three classes: cat, bird, and dog. With these images, we would like to train an image classifier that categorizes an input image into one of the three classes.

To ensure the 1,000 images are of good quality for training the classifier, we ask a well-trained human inspector to go through all the images, to label the images and remove noisy ones. Eventually, we removed 200 images of low quality suggested by the inspector, and use the remaining 800 images to train the classifier; the 800 images comprise 200 cat images, 300 bird images, and 300 dog images.

Please select the correct option.

- ✓A. The human inspection process can be considered as a data cleaning step.
- B. If we are to use one-hot encoding for the labels of the three classes, we can set
  - Cat = [1 1 1]
  - Dog = [0 1 0]
  - Bird = [0 0 0].
- C. The image classification conducted here is an unsupervised-learning task.
- D. If we keep the 200 noisy images (suggested by the human inspector), we will end up having more training images and hence a better-performed image classifier.
- E. (a) and (b)
- F. (a), (b), and (d)
- G. None of others is correct.

Item ID: 241562 / 1

Item Description: Copy of Part1-Midterm-Trial-MCQ6

Item Weight: 3.0

---

**Question #: 2**

A discrete random variable takes a finite number of values, while a continuous random variable can only take infinite number of values. (2 marks)

- A. True
- ✓B. False

**Item ID:** 241561 / 1

**Item Description:** Copy of Part1-Midterm-Trial-TF1

**Item Weight:** 2.0

---

**Question #: 3**

This question is related to the understanding of modelling assumptions.  $f(x) = 5x - 3$  is a linear function. (2 marks)

- A. True
- ✓B. False

**Item ID:** 241560 / 1

**Item Description:** Copy of Part2-Midterm-Trial-TF1

**Item Weight:** 2.0

---

**Question #: 4**

One key step in Data Cleaning is to check the missing features of data samples. When we have insufficient number of training samples in our dataset, we may consider removing the examples with missing features.

- A. True
- ✓B. False

**Item ID:** 241559 / 1

**Item Description:** Copy of Part1-Midterm-Trial-TF3

**Item Weight:** 2.0

---

**Question #: 5**

Causality is a deterministic relationship; suppose we know A and B have causal relation, if A occurs, B is for sure to take place.

- A. True
- ✓B. False

**Item ID:** 241558 / 1

**Item Description:** Copy of Part1-Midterm-Trial-TF2

**Item Weight:** 2.0

---

**Question #:** 6

Three balls are drawn from three urns sequentially, one ball from each urn. The first urn contains 1 blue and 7 red balls, the second urn contains 2 blue and 6 red balls, and the third urn contains 3 red and 5 green balls. Find the probability that 2 red balls are chosen. (3 marks)

- A. 226/64
- B. 226/512
- ✓C. 270/512
- D. 270/1024
- E. None of the rest.

**Item ID:** 241557 / 1

**Item Description:** Copy of Part1-Midterm-Trial-MCQ1

**Item Weight:** 3.0

---

**Question #:** 7

A machine learning algorithm takes the temperature as one of its input features. The temperature is measured in Celsius. Please select the correct option.

- A. The temperature in Celsius is considered as interval data.
- B. We can calculate the mean and standard deviation of temperature.
- c. The temperature in Celsius is considered as ratio data.

- D. None of the rest.
- ✓E. (a), and (b)
- F. (a), and (c)

**Item ID:** 241556 / 1

**Item Description:** Copy of Part1-Midterm-Trial-MCQ2

**Item Weight:** 3.0

---

**Question #:** 8

A person draws 2 cards from a deck of 52 cards, one after another without replacing the previous card back. What is the probability of drawing two Queens in a row?

- A.  $2/52$
- B.  $4/52$
- ✓C.  $1/221$
- D.  $3/51$

**Item ID:** 241555 / 1

**Item Description:** Copy of Part1-Midterm-Trial-MCQ3

**Item Weight:** 3.0

---

**Question #:** 9

Which of the following task is likely to be achieved via supervised learning?

- ✓A. Using historical data for weather forecast.
- B. Grouping together users with similar viewing patterns in order to recommend similar content.
- c. Grouping a number of oranges by their size.
- None of the rest.

**Item ID:** 241554 / 1

**Item Description:** Copy of Part1-Midterm-Trial-MCQ4

**Item Weight:** 3.0

---

**Question #:** 10

A machine learning algorithm takes the letter grade of students as one of its input features. The letter grade can take any element from {A+, A, A?, B+, B, B?, C+, C, D+, D, F}, subject to some distribution curving. For example, A+ corresponds to 95%, A corresponds to 85%, and A- corresponds to 80%. Which of the following statements is/are true?

- A. The letter grade is an example of nominal variable.
- B. The letter grade is an example of ordinal variable.
- C. The letter grade is an example of interval variable.
- D. The letter grade is an example of discrete variable.
- E. (a), and (c)
- F. (b), and (c)
- ✓G. (b), and (d)

**Item ID:** 241553 / 1

**Item Description:** Copy of Part1-Midterm-Trial-MCQ5

**Item Weight:** 3.0

---

**Question #:** 11

The values of feature x and their corresponding values of target y are shown in the table below.

x	3	4	5	6	7
y	5	4	3	2	1

Find the least square regression line  $y = a x + b$  and then estimate the value of y when  $x = 8$ .

- A.  $y = 8$
- B.  $y = +1$
- ✓C.  $y = 0$
- D.  $y = ?1$
- E. None of the above

Item ID: 241552 / 1

Item Description: Copy of Part2-Midterm-Trial-MC1

Item Weight: 3.0

---

Question #: 12

A set of linear equations is written as

$$\mathbf{w}^T \mathbf{X} = \mathbf{y}^T$$

where

$$\mathbf{X} \in \mathcal{R}^{3 \times 2}$$

and

$$\mathbf{y} \in \mathcal{R}^{2 \times 1}$$

. How many simultaneous equations are there in this set of equations?

- A. 1
- ✓B. 2
- C. 3
- D. 4
- E. 5

Item ID: 241551 / 1

Item Description: Copy of Part2-Midterm-Trial-MC2

Item Weight: 3.0

---

Question #: 13

This question is related to the understanding of linear systems and partial derivatives. Which of the following statements below is correct?

- A. In over-determined linear systems, the number of parameters is greater than the number of unknown equations.
- ✓B. The system

$$\begin{bmatrix} 1 & 4 \\ 2 & 7 \\ -3 & 11 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -2.5 \\ 4 \end{bmatrix}$$

has no *exact* solution but an approximated solution is available using the left inverse.

- C. If  $\mathbf{f}(\mathbf{x})$  is a vector-valued function of size  $p \times 1$  and  $\mathbf{x}$  is an  $m \times 1$  vector, then differentiation of  $\mathbf{f}(\mathbf{x})$  with respect to  $\mathbf{x}$  is an  $m \times p$  matrix.
- D. A linear function needs to satisfy the properties of homogeneity only.
- E. None of the other options.

**Item ID:** 241550 / 1

**Item Description:** Copy of Part2-Midterm-Trial-MC3

**Item Weight:** 3.0

**Question #:** 14

Suppose the random variable X has a probability mass function (pmf) given in the table below.

X	1	2	3	4	5
Pr[X]	0.1	(BLANK1)	0.2	0.4	(BLANK2)

We also know that the expected value of X is 3.5.

- 1) What is the probability of  $\Pr[X=5]$ ? 1 (2 Marks)
- 2) What is the probability of  $\Pr[X \leq 2]$ ? 2 (2 Marks)

1. Range - Min:0.19999 Max:0.20001

2. Range - Min:0.19999 Max:0.20001

**Item ID:** 241549 / 3

**Item Description:** Copy of Part1-Midterm-Trial-FIB1

**Item Weight:** 4.0

**Question #:** 15

You are given a collection of 5 training data points of two features  $(x_1, x_2)$  and their target output  $(y)$  which are packed as follows:

$$\text{Feature matrix: } \mathbf{X} = \begin{bmatrix} 1 & 2 \\ 0 & 6 \\ 1 & 0 \\ 0 & 5 \\ 1 & 7 \end{bmatrix}, \text{ Target output: } \mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}.$$

Predict the output (up to 4 decimal places) of  $(x_1, x_2) = (1, 3)$  using the linear regression model. (4 marks)

1) What is the mean of squared error of the estimated model?   1   (up to 4 decimal places, 2 mark)

2) The prediction for y is   2   (up to 4 decimal places, 2 marks).

1. Range - Min:1.3886 Max:1.3888

2. Range - Min:2.9999 Max:3.0001

**Item ID:** 241548 / 3

**Item Description:** Copy of Part2-Midterm-Trial-FIB1

**Item Weight:** 4.0