Week 1 PYQs
Machine Learning Foundations
May 22 - Sep 25

Note

1. If no questions are listed in that term Quiz, it indicates that the corresponding quiz did not include questions from that week.

2. The answer key may contain minor inaccuracies, as answers were taken directly from the original quiz/exam papers. Some of these might have been corrected later.

3. For any issues or suggestions regarding this document, feel free to contact:
   `23f3004032@ds.study.iitm.ac.in`
   or reach out on Discourse: `@23f3004032`

Quiz 1 May 22

1. For the data sets $(x^i, y^i) = [(1,1),(2,3),(3,4),(4,4),(5,7)]$, $i = 1$ to 5, consider the regression model $f(x) = x + 2$. What is the mean squared loss?

   [Ans 2]

2. Identify which of the following require classification:

   (a) Predicting the amount of rainfall in May 2022 in North India based on precipitation data of the year 2021

   (b) Predicting the price of a land based on its area and distance from the market

   (c) Predicting whether an email is spam or not.

   (d) Predicting whether the patient is having cancer or not.

   [Ans: C,D]

Quiz 2 May 22

ET May 22

Quiz 1 Sep 22

1. Identify which of the following requires use of classification technique.

   (a) Credit card fraud detection

   (b) Sentiment analysis

   (c) Predicting whether an email is spam or not

   (d) Predicting whether the patient is having cancer or not

   [Ans: a, b, c, d]

2. For the data sets $(x^i, y^i) = [(1,1),(2,2),(3,4),(4,5),(5,5)]$, consider $f(x) = x + 2$. What is the mean squared loss? (one decimal place)

   [Ans: 2.8]

Quiz 2 Sep 22

ET Sep 22

1. For the data sets $(x^i, y^i) = [(1,0),(2,2),(3,4),(4,5),(5,5)]$, $i = 1$ to 5, Consider the regression model $f(x) = x$. What is the mean squared loss of $f(x)$ (Enter answer correct to one decimal place)? [Ans: 0.6]

Quiz 1 Jan 23

1. Which of the following is/are true about loss after applying Encoder and Decoder functions to a given data?

   (a) The loss value can be equal to any real number.

   (b) The loss value cannot be negative.

   (c) A perfect encoder and decoder function will not exist for any data, that is loss value can never be equal to zero.

   (d) For some data, there can exist a perfect encoder and decoder functions, that is loss value can be equal to zero.

   [Ans: b, d]

2. Which of the following statements is/are true?

   (a) Email spam detection falls under classification problem.

   (b) Predicting the price of a house based on previous datasets falls under classification problem.

   (c) Classifying whether handwriting belongs to a particular person or not based on previous data falls under the classification problem.

   (d) Finding out the probability that a particular poem is written by Shakespeare falls under unsupervised learning.

   [Ans: a, c]

3. For the data set $\{(x^i, y^i)\} = \{(1,1), (2,6), (3,8), (4,15), (5,26)\}$, consider $f(x) = x^2+1$. What is mean squared loss? (nearest integer)

   [Ans: 2]

4. What will be the average misclassification error when $g(X) = \text{sign}((x_1 - x_2)^2 - 8)$ is used?

X	y
$(5,2)$	1
$(5,6)$	-1
$(6,3)$	1
$(6,4)$	-1
$(4,8)$	-1

   [Ans: 0.2]

Quiz 2 Jan 23

ET Jan 23

1. Which of the following statements regarding classification and regression models are correct?

   (a) The task of the regression algorithm is to map the input value (x) with the continuous output variable (y) but the task of the classification algorithm is to map the input value(x) with the discrete output variable(y).

   (b) In Classification, we try to find the decision boundary, which can divide the dataset into different classes but In Regression, we try to find the best fit line, which can predict the output more accurately.

   (c) The Classification Algorithm can be further divided into Linear and Non-linear classifier and the regression algorithms can be divided into Binary regression and Multi-class regression.

   (d) The task of the regression algorithm is to map the input value (x) with the discrete output variable (y) but the task of the classification algorithm is to map the input value(x) with the continuous output variable(y).

   [Ans: 1, 2]

Quiz 1 May 23

1. Consider dataset:

$$(2,3) \to 5,\ (-1,1) \to 2,\ (4,2) \to 7,\ (0,-2) \to 1,\ (-3,5) \to 4$$

Model $f(X) = w^T x$, $w = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Compute

$$L = \max_i |f(x^i) - y^i|$$

[Ans: 5]

2. Which of the following statements are correct?

   (a) $\mathbb{I}(44\%3 = 0 \text{ and } \pi \in \mathbb{R}) = 1$

   (b) Mean Squared Error (MSE) is a suitable loss function for classification problems.

   (c) The loss obtained after applying encoder and decoder functions to a given data can be negative.

   (d) Gaussian mixture model is a family of distributions for modeling multimodal data.

[Ans: a, d]

Quiz 2 May 23

ET May 23

1. For the data set $(x_i, y_i) = [(1,1),(2,3),(3,9),(4,15),(5,27)]$, $i = 1$ to 5, consider the regression model $f(x) = x^2$. What is the mean squared loss of f(x). (Enter answer correct to one decimal place). [Ans: 1.1 to 1.3]

Quiz 1 Sep 23

1. Which of the following is/are supervised learning task(s)?

   (a) Predicting future weather conditions based on the past weather observations

   (b) Reducing the dimensionality of images while retaining essential features,enabling efficient storage and transmission.

   (c) Identifying objects in images with multiple categories, such as cats, dogs, cars, and trees

   (d) Categorizing news articles into topics like politics, sports, entertainment, and technology.

   [Ans: a, c, d]

2. Consider the following input data points. Which model has least misclassification error?

X	Y
$(0, 1)$	$+1$
$(0, 2)$	$+1$
$(1, 3)$	$+1$
$(3, 4)$	-1
$(4, 5)$	-1
$(7, 6)$	-1

   (a) $\text{sign}(x_1 - 2)$

   (b) $\text{sign}(x_2 + 2)$

   (c) $\text{sign}(x_1 + x_2)$

   (d) $\text{sign}(x_2 - x_1)$

   [Ans: d]

3. Consider a regression model $h(x_1, x_2) = \frac{x_1+x_2}{2}$. For the below data points, find the mean squared loss.

X	Y
$(1, 2)$	2
$(2, 3)$	3
$(3, 4)$	4
$(4, 5)$	5

   [Ans: 0.25]

4. Find best fit line using least squares.

x	1	3	4	5	7
y	-3	7	1	0	4

(a) $\hat{y} = 0.7x - 1$

(b) $\hat{y} = x + 0.5$

(c) $\hat{y} = x + 0.2$

(d) $\hat{y} = 0.2x + 1$

[Ans: a]

Quiz 2 Sep 23

ET Sep 23

1. Consider an encoder-decoder pair used for dimensionality reduction, where the encoder function is denoted as $f(x_1, x_2, x_3) = (x_1 + x_2 + x_3)/3$ and the decoder function is denoted as $g(u) = [u \quad u^2 \quad u]$. Compute the reconstruction error ($R(f, g) = \frac{1}{n}\sum_{i=1}^{n} ||X_i - g(f(X_i))||^2$) for the following dataset: $X_1 = [1, 2, 0], X_2 = [-1, 1, 0], X_3 = [1, 2, 3]$. Enter the answer correct to two decimal places. [Ans: 3.31 to 3.35]

Quiz 1 Jan 24

1. Which of the following is/are supervised learning task(s)?

   (a) Predicting future sales of a product based on the past sales data.

   (b) Reducing the dimensionality of images while retaining essential features, enabling efficient storage and transmission.

   (c) Recognizing and interpreting handwritten characters or words, often used in digit recognition or reading handwritten text.

   (d) Categorizing text documents into predefined classes, such as spam or non-spam emails.

   [Ans: a, c, d]

2. Consider an encoder-decoder pair used for dimensionality reduction, where the encoder function is denoted as

   $$f(x_1, x_2, x_3) = (x_1 + x_2 - x_3,\ x_1 + x_3)$$

   and the decoder function is denoted as

   $$g(u) = (u_1, u_2, u_1 + u_2).$$

   Compute the reconstruction error

   $$R(f, g) = \frac{1}{n} \sum_{i=1}^{n} \|X_i - g(f(X_i))\|^2$$

   for the following dataset:

   $$X_1 = [1, 2, 3], \quad X_2 = [0, 2, 2], \quad X_3 = [1, 0, -1].$$

   Enter the answer correct to two decimal places.

   [Ans: 3.12–3.20]

Quiz 2 Jan 24

ET Jan 24

Quiz 1 May 24

1. In a sports analytics company, the team is interested in predicting whether a basketball player will score above 20 points in a game using classification models. They collected data on various factors $x = [x_1, x_2, x_3]$ from past games of different players. The factors include minutes played, shots attempted, and free throw percentage. The data and corresponding labels are shown below:

x	y
[1, 5, 0.8]	1
[0, 3, 0.9]	0
[1, 2, 0.7]	0
[1, 8, 0.6]	1
[0, 7, 0.5]	0

Compute the misclassification rate if they use the following step function:

$$u(z) = \begin{cases} 1, & \text{if } z \geq 3 \\ 0, & \text{otherwise} \end{cases}$$

and the linear combination z is given by $z = 0.6x_1 + 0.3x_2 + 0.1x_3$. Enter the answer correct to one decimal place. [Ans:0.2]

Quiz 2 May 24

ET May 24

Quiz 1 Sep 24

1. Consider the following input data points:

x	y
$[0, 1, 1]$	-1
$[0, 1, 0]$	-1
$[1, 1, 3]$	$+1$
$[1, 0, 5]$	$+1$
$[0, 0, 6]$	$+1$
$[1, 0, 4]$	$+1$

Which among the following models has the least misclassification error? (Hint: Consider $\text{Sign}(0) = 1$)

(a) $\text{Sign}(x_1 + x_2 + x_3 - 5)$

(b) $\text{Sign}(x_1 + x_2 - 2)$

(c) $\text{Sign}(x_1 + x_2 + x_3)$

(d) $\text{Sign}(x_3 - x_2 - x_1)$

[Ans: a]

Quiz 2 Sep 24

ET Sep 24

Quiz 1 Jan 25

1. A bank wants to predict whether a customer will default on a loan using a classification model. The bank collects data from past customers, which includes the following features:

   - x_1: Credit score (scaled between 0 and 1)
   - x_2: Monthly income (in \$1000s)
   - x_3: Number of late payments in the past year

   The target variable y represents whether the customer defaulted (1 for default, 0 for no default). The data is provided in the table below:

X	y
[0.9, 5, 1]	0
[0.4, 3, 5]	1
[0.8, 4, 2]	0
[0.2, 2, 7]	1
[0.7, 6, 0]	0

   The bank uses the following linear combination to compute z:

   $$z = 0.5x_1 + 0.3x_2 - 0.2x_3$$

   The step function $u(z)$ is used to classify the outcome:

   $$u(z) = \begin{cases} 1, & \text{if } z \geq 1.5 \\ 0, & \text{otherwise} \end{cases}$$

   Based on the above data, answer the given subquestions.

   (a) For the input $x = [0.9, 5, 1]$, what will be the predicted output $u(z)$?

   i. 0
   ii. 1
   iii. 1.75
   iv. Cannot be determined

   [Ans: b]

   (b) If the threshold in the step function is changed to $z \geq 1.0$, what will happen to the misclassification rate?

   i. Increase
   ii. Decrease

iii. Remain same
iv. Cannot be determined

[Ans: a]

Quiz 2 Jan 25

ET Jan 25

Quiz 1 May 25

1. The table below shows original and predicted values. Compute square error loss and 0-1 loss.

Original	-1	0	1	2	3	1	0	3
Predicted	1	1	1	2	1	0	0	3

   Which of the following options is/are correct?

   (a) Square error loss $\leq$ 0-1 loss

   (b) Square error loss $\geq$ 0-1 loss

   (c) The 0-1 loss is 0.25

   (d) The square error loss is 1.25

   [Ans: a, d]

2. A hospital collects patient data and wants to find patterns without labels. What type of task is clustering most suitable for?

   (a) Predicting the recovery time for a new patient.

   (b) Diagnosing diseases based on labeled medical records.

   (c) Grouping patients with similar symptoms and test results without prior labels.

   (d) Determining the most important variable for predicting disease.

   [Ans: c]

Quiz 2 May 25

ET May 25

Quiz 1 Sep 25

1. A bank wants to decide whether a loan applicant should be given a loan or not, based on features like income, credit score and past records. Which of the following machine learning techniques is applicable for this scenario?

   (a) Regression

   (b) Classification

   (c) Clustering

   (d) Dimensionality Reduction

   [Ans: b]

2. Consider the dataset $(2,4,3),(5,5,6),(2,2,1),(4,4,6)$ with

   $$f(x_1,x_2,x_3) = 0.5x_1 + 0.5x_2 - x_3, \quad g(u) = [u,u,u].$$

   Find the reconstruction loss.

   [Ans: 72]

3. Consider a regression dataset with 4 points (2,0),(3,3),(5,5),(6,8)

   (a) Which model has the smallest squared loss?

   i. $f(x) = 2x - 4$
   ii. $f(x) = 4$
   iii. $f(x) = 8$
   iv. $f(x) = 4 - 2x$

   [Ans: a]

   (b) Is there a model $f(x) = wx + b$ with zero squared loss?

   i. No
   ii. Yes
   iii. Insufficient information

   [Ans: a]

Quiz 2 Sep 25

ET Sep 25