

# FIB-1

In the following single neuron, the neuron has three inputs  $x = (x_1, x_2, x_3)$  that receives only binary signals (either 0 or 1). The number of all possible input patterns this node can receive is \_\_\_\_\_

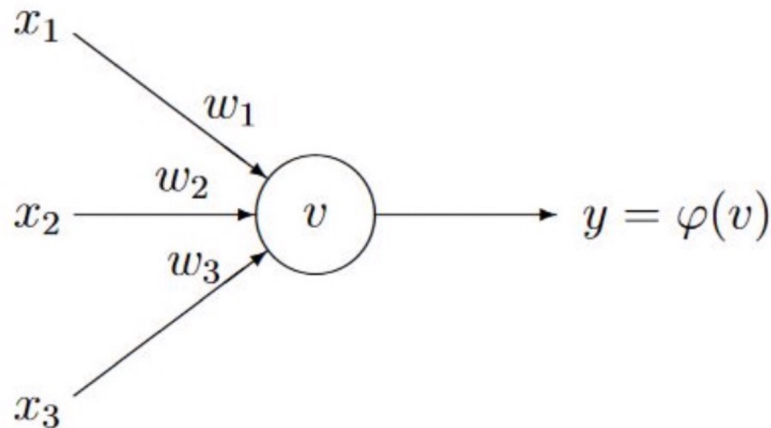


Figure 1: Single unit with three inputs.

Please fill in your answer here

Number    Correct Answer

1

Range - Min: 7.9999 Max: 8.00001

# FIB-2

Suppose we are minimizing  $f(x) = x^4$  with respect to  $x$ . We initialize  $x$  to be 2. We perform gradient descent with learning rate 0.1. After the first iteration,  $x$  will become \_\_\_\_\_ (please specify your answer with a number of up to 2 decimal places)

Please provide your answer here

Number    Correct Answer

1

Range - Min: -1.21 Max: -1.19

# FIB-3

You are given a collection of 6 training data points of three features  $(x_1, x_2, x_3)$  and their one-hot encoding of class labels  $(y_1, y_2, y_3)$  which are packed as follows:

$$\mathbf{X} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 0 & 6 \\ 1 & 1 & 0 \\ 0 & 1 & 2 \\ 5 & 7 & -2 \\ -1 & 4 & 0 \end{bmatrix}, \mathbf{Y} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \text{class1} \\ \text{class1} \\ \text{class2} \\ \text{class3} \\ \text{class2} \\ \text{class3} \end{bmatrix}.$$

(i) Predict the class label of  $(x_1, x_2, x_3) = (1, -2, 3)$  using the linear regression model with a bias term.

My prediction for the class label is class 1 (enter 1, 2, or 3?).

(ii) Predict the class label of  $(x_1, x_2, x_3) = (1, -2, 3)$  using the full polynomial model of third order.

My prediction for the class label is class 2 (enter 1, 2, or 3?).

Please provide your answer below

1)

2)

Number	Correct Answer
1	Choice of: 1   <u>2</u>   3
2	Choice of: 1   2   <u>3</u>

# MCQ-1

We partition a data set into training, validation and test data. Which of the following statement is correct?

Seq	Answer Choice	Correct
a)	Training data and test data may overlap in one experiment.	<input type="checkbox"/>
b)	Test data and validation data may overlap in one experiment.	<input type="checkbox"/>
c)	We use validation data to select parameters.	<input type="checkbox"/>
d)	Test data are unseen during training.	<input type="checkbox"/>
e)	a) and b)	<input type="checkbox"/>
f)	b) and c)	<input type="checkbox"/>
g)	c) and d)	<input checked="" type="checkbox"/>

# MCQ-2

What are valid evaluation metrics for machine learning?

Seq	Answer Choice	Correct
a)	Equal error rate	<input type="checkbox"/>
b)	False positive rate	<input type="checkbox"/>
c)	False negative rate	<input type="checkbox"/>
d)	Validation dataset	<input type="checkbox"/>
e)	b) and c)	<input type="checkbox"/>
f)	a), b), c), and d)	<input type="checkbox"/>
g)	a), b), c)	<input checked="" type="checkbox"/>

# MCQ-3

To build a classifier to classify people by gender from their images, which of the following dataset is most suitable?

Seq	Answer Choice	Correct
a)	Images with labels of people's nationality.	<input type="checkbox"/>
b)	Images with no labels.	<input type="checkbox"/>
c)	Images with labels of people's age.	<input type="checkbox"/>
d)	Images of people labeled as male or female.	<input checked="" type="checkbox"/>

# MCQ-4

A set of linear equations is written as  $\mathbf{X}\mathbf{w} = \mathbf{y}$ , where  $\mathbf{X} \in \mathcal{R}^{5 \times 3}$ ,  $\mathbf{w} \in \mathcal{R}^{3 \times 1}$ , and  $\mathbf{y} \in \mathcal{R}^{5 \times 1}$ .  
How many simultaneous equations are there in this set of equations?

Seq	Answer Choice	Correct
a)	1	<input type="checkbox"/>
b)	2	<input type="checkbox"/>
c)	3	<input type="checkbox"/>
d)	5	<input checked="" type="checkbox"/>
e)	6	<input type="checkbox"/>

# TF-1

If we run K-means algorithms multiple times with different initializations, we always get the same result.

**Correct Answer Choice**

- ☐ True
- ☒ False



# TF-2

A vector function has the mapping of  $\mathbf{f}(\mathbf{x}): \mathbb{R}^3 \rightarrow \mathbb{R}^2$ . The gradient of this function with respect to  $\mathbf{x}$  is a matrix of partial derivatives, which is a  $2 \times 3$  matrix.

Correct Answer Choice

- ☒ True
- ☐ False

# TF-3

The ReLU function is given as  $f(a) = \max(0, a)$ . The output of ReLU is discrete data.

**Correct Answer Choice**

- ☐ True
- ☒ False

# TF-4

Increasing the maximum depth of a decision tree will in general increase the bias of the resulting trained decision tree.

**Correct Answer Choice**

- ☐ True
- ☒ False

# TF-5

We build a classifier to classify emails into regular mails and spam. As such, the classifier is making a deductive reasoning.

**Correct Answer Choice**

- ☐ True
- ☒ False

# TF-6

To perform k-means clustering on a set of unlabeled data, one possible way is to randomly group the set of unlabeled data into k clusters as an initialization.

**Correct Answer Choice**

- ☒ True
- ☐ False