

X

<https://swayam.gov.in>https://swayam.gov.in/nc_details/NPTEL

2111cs010077@mallareddyuniversity.ac.in ✓

NPTEL (<https://swayam.gov.in/explorer?ncCode=NPTEL>) » Deep Learning - IIT Ropar (course)Course
outlineAbout
NPTEL ()How does an
NPTEL
online
course
work? ()

Week 1 ()

Week 2 ()

Week 3 ()

week 4 ()

Week 5 ()

Week 6 ()

Week 7 ()

Week 8 ()

- ☐ A quick recap
of training
deep neural
networks
(unit?)

Week 8 : Assignment 8

The due date for submitting this assignment has passed.

Due on 2024-09-18, 23:59 IST.

Assignment submitted on 2024-09-18, 19:04 IST

1) Which of the following activation functions is not zero-centered?

1 point

- ☒ Sigmoid
☐ Tanh
☒ ReLU
☒ Softmax

Yes, the answer is correct.

Score: 1

Accepted Answers:

Sigmoid

ReLU

Softmax

2) What is the gradient of the sigmoid function at saturation?

Yes, the answer is correct.

Score: 1

Accepted Answers:

(Type: Numeric) 0

3) What are the challenges associated with using the Tanh(x) activation function?

1 point

1 point

- ☐ It is not zero centered
☒ Computationally expensive

unit=107&less
on=108)

☐ Unsupervised
pre-training
(unit?
unit=107&less
on=109)

☐ Better
activation
functions
(unit?
unit=107&less
on=110)

☐ Better
initialization
strategies
(unit?
unit=107&less
on=111)

☐ Batch
Normalization
(unit?
unit=107&less
on=112)

☐ Lecture
Material for
Week 8 (unit?
unit=107&less
on=113)

☒ Week 8
Feedback
Form: Deep
Learning - IIT
Ropar (unit?
unit=107&less
on=191)

☒ **Quiz: Week 8
: Assignment
8
(assessment?
name=296)**

Week 9 ()

week 10 ()

Week 11 ()

Week 12 ()

- ☐ Non-differentiable at 0
☒ Saturation

Yes, the answer is correct.

Score: 1

Accepted Answers:

Computationally expensive

Saturation

4) Which of the following activation functions is preferred to avoid the vanishing gradient problem? **1 point**

- ☐ Sigmoid
☐ Tanh
☒ ReLU
☐ None of these

Yes, the answer is correct.

Score: 1

Accepted Answers:

ReLU

5) Which of the following best describes the concept of saturation in deep learning? **1 point**

- ☒ When the activation function output approaches either 0 or 1 and the gradient is close to zero.
☐ When the activation function output is very small and the gradient is close to zero.
☐ When the activation function output is very large and the gradient is close to zero.
☐ None of the above.

Partially Correct.

Score: 0.33

Accepted Answers:

When the activation function output approaches either 0 or 1 and the gradient is close to zero.

When the activation function output is very small and the gradient is close to zero.

When the activation function output is very large and the gradient is close to zero.

6) Which of the following methods can help to avoid saturation in deep learning? **1 point**

- ☒ Using a different activation function.
☐ Increasing the learning rate.
☐ Increasing the model complexity
☐ All of the above.

Yes, the answer is correct.

Score: 1

Accepted Answers:

Using a different activation function.

7) Which of the following is true about the role of unsupervised pre-training in deep learning? **1 point**

- ☐ It is used to replace the need for labeled data
☒ It is used to initialize the weights of a deep neural network

**Download
Videos ()**

Books ()

**Text
Transcripts
()**

**Problem
Solving
Session -
July 2024 ()**

- ☐ It is used to fine-tune a pre-trained model
- ☐ It is only useful for small datasets

Yes, the answer is correct.

Score: 1

Accepted Answers:

It is used to initialize the weights of a deep neural network

8) What is the main cause of the Dead ReLU problem in deep learning?

1 point

- ☐ High variance
- ☒ High negative bias
- ☐ Overfitting
- ☐ Underfitting

Yes, the answer is correct.

Score: 1

Accepted Answers:

High negative bias

9) What is the purpose of Batch Normalization in Deep Learning?

1 point

- ☐ To improve the generalization of the model
- ☐ To reduce overfitting
- ☐ To reduce bias in the model
- ☒ To ensure that the distribution of the inputs at different layers doesn't change

Yes, the answer is correct.

Score: 1

Accepted Answers:

To ensure that the distribution of the inputs at different layers doesn't change

10) In Batch Normalization, which parameter is learned during training?

1 point

- ☐ Mean
- ☐ Variance
- ☒ γ
- ☐ ϵ

Yes, the answer is correct.

Score: 1

Accepted Answers:

γ