



(https://swayam.gov.in)



(https://swayam.gov.in/nc_details/NPTEL)

vijay.mahawar@hotmail.com

NPTEL (https://swayam.gov.in/explorer?ncCode=NPTEL) » Introduction to Large Language Models (LLMs)
(course)

Click to register
for Certification
exam

(https://examform.nptel.ac.in/2025_01/exam_form/dashboard)

If already
registered, click
to check your
payment status

Course
outline

About NPTEL
()

How does an
NPTEL online
course work?
()

Week 1 ()

Week 2 ()

Week 3 ()

Lec 05 :
Introduction to
Deep Learning
(unit?
unit=30&lesson
=31)

Lec 06 :
Introduction to
PyTorch (unit?
unit=30&lesson
=32)

Week 3 : Assignment 3

The due date for submitting this assignment has passed.

Due on 2025-02-12, 23:59 IST.

Assignment submitted on 2025-02-10, 18:33 IST

1) State whether the following statement is True/False. 1 point
The Perceptron learning algorithm can solve problems with non-linearly separable data.

- ☐ True
- ☒ False

Yes, the answer is correct.
Score: 1
Accepted Answers:
False

2) In backpropagation, which method is used to compute the gradients? 1 point

- ☐ Gradient descent
- ☒ Chain rule of derivatives
- ☐ Matrix factorization
- ☐ Linear regression

Yes, the answer is correct.
Score: 1
Accepted Answers:
Chain rule of derivatives

3) Which activation function outputs values in the range [-1,1]? 1 point

- ☐ ReLU
- ☒ Tanh
- ☐ Sigmoid
- ☐ Linear

Yes, the answer is correct.

● Lecture Material
(unit?
unit=30&lesson
=33)

● Feedback Form
(unit?
unit=30&lesson
=34)

● Quiz: Week 3 :
Assignment 3
(assessment?
name=35)

Week 4 ()

Week 5 ()

Week 6 ()

Week 7 ()

Week 8 ()

Week 9 ()

Week 10 ()

Week 11 ()

Week 12 ()

Year 2025
Solutions ()

Score: 1

Accepted Answers:

Tanh

4) What is the primary goal of regularization in machine learning?

1 point

- ☐ To improve the computational efficiency of the model
- ☒ To reduce overfitting
- ☐ To increase the number of layers in a network
- ☐ To minimize the loss function directly

Yes, the answer is correct.

Score: 1

Accepted Answers:

To reduce overfitting

5) Which of the following is a regularization technique where we randomly deactivate neurons during training?

1 point

- ☐ Early stopping
- ☐ L1 regularization
- ☒ Dropout
- ☐ Weight decay

Yes, the answer is correct.

Score: 1

Accepted Answers:

Dropout

6) Which activation function has the vanishing gradient problem for large positive or negative inputs?

1 point

- ☐ ReLU
- ☒ Sigmoid
- ☐ GELU
- ☐ Swish

Yes, the answer is correct.

Score: 1

Accepted Answers:

Sigmoid

7) Which activation function is defined as: $f(x)=x \cdot \sigma(x)$, where $\sigma(x)$ is the sigmoid function?

1 point

- ☒ Swish
- ☐ ReLU
- ☐ GELU
- ☐ SwiGLU

Yes, the answer is correct.

Score: 1

Accepted Answers:

Swish

8) What does the backpropagation algorithm compute in a neural network?

1 point

- ☐ Loss function value at each epoch
- ☒ Gradients of the loss function with respect to weights of the network
- ☐ Activation values of the output layer
- ☐ Output of each neuron

Yes, the answer is correct.

Score: 1

Accepted Answers:

Gradients of the loss function with respect to weights of the network

9) Which type of regularization encourages sparsity in the weights?

1 point

- ☐ L1 regularization
- ☒ L2 regularization
- ☐ Dropout
- ☐ Early stopping

No, the answer is incorrect.

Score: 0

Accepted Answers:

L1 regularization

10) What is the main purpose of using hidden layers in an MLP?

1 point

- ☐ Helps to the network bigger
- ☐ Enables us to handle linearly separable data
- ☒ Learn complex and nonlinear relationships in the data
- ☐ Minimize the computational complexity

Yes, the answer is correct.

Score: 1

Accepted Answers:

Learn complex and nonlinear relationships in the data