

# Week 2 Introduction

---



## Topic Introduction

### Week 2 Lesson Video

(18:14)

Watch this video for a comprehensive introduction to this week's learning content. You can follow along in Blackboard as you watch, with the video organized into the following chapters:

- Lesson 2.1 Multi-Class Classification with Neural Networks
- Lesson 2.2 Training Using Backpropagation and Gradient Descent (08:28)



## Required Resources

The following resources are required for your learning this week. Make sure you review everything linked below, as you may be tested on the concepts in an upcoming Knowledge Check or Homework.

**Reading** | Géron, A. (2023). [Hands-on machine learning with scikit-learn, Keras and TensorFlow: Concepts, tools, and techniques to build intelligent systems](https://bu.primo.exlibrisgroup.com/permalink/01BOSU_INST/u1o13f/alma99209313482001161) ([https://bu.primo.exlibrisgroup.com/permalink/01BOSU\\_INST/u1o13f/alma99209313482001161](https://bu.primo.exlibrisgroup.com/permalink/01BOSU_INST/u1o13f/alma99209313482001161)). O'Reilly Media, Inc.

- Read pages 309 to 316 which introduce backpropagation on multi-level perceptrons (= neural networks).
- Read pages 349 to 353 for a brief discussion of some of the hyperparameters and optimization techniques (including number of layers and number of neurons in each layer).

**Video** | Sanderson, G. (2017, October 6). [Gradient descent, how neural networks learn | DL Chapter 2 \[Video\]](https://www.youtube.com/watch?v=IHZwWFHwaw&list=PLZHQObOWTQDNU6R1_67000Dx_ZCJB-3pi&index=2). ([https://www.youtube.com/watch?v=IHZwWFHwaw&list=PLZHQObOWTQDNU6R1\\_67000Dx\\_ZCJB-3pi&index=2](https://www.youtube.com/watch?v=IHZwWFHwaw&list=PLZHQObOWTQDNU6R1_67000Dx_ZCJB-3pi&index=2)) YouTube.

**Video** | Sanderson, G. (2017, November 3). [Backpropagation, intuitively | DL Chapter 3 \[Video\]](https://www.youtube.com/watch?v=llg3gGewQ5U&list=PLZHQObOWTQDNU6R1_67000Dx_ZCJB-3pi&index=3). ([https://www.youtube.com/watch?v=llg3gGewQ5U&list=PLZHQObOWTQDNU6R1\\_67000Dx\\_ZCJB-3pi&index=3](https://www.youtube.com/watch?v=llg3gGewQ5U&list=PLZHQObOWTQDNU6R1_67000Dx_ZCJB-3pi&index=3)) YouTube.

These videos follow up on the 3Blue1Brown video, But what is a neural network? | DL Chapter 1. These are absolutely essential resources! Definitely watch DL2 and DL3. Then watch them again.

## Optional Resources

The following are optional resources for you to review if you wish to learn more about the topics we covered.

**Reading** | Pages 379-392 in Géron, A. (2023).

- Offers a deep dive into optimizers and learning-rate scheduling. It may be “too much information” for now, but you can skim through the main ideas.

**Video** | Sanderson, G. (2017, November 3). [Backpropagation calculus | DL Chapter 4 \[Video\]](https://www.youtube.com/watch?v=tIeHLnjs5U8&list=PLZHQObOWTQDNU6R1_67000Dx_ZCJB-3pi&index=4). ([https://www.youtube.com/watch?v=tIeHLnjs5U8&list=PLZHQObOWTQDNU6R1\\_67000Dx\\_ZCJB-3pi&index=4](https://www.youtube.com/watch?v=tIeHLnjs5U8&list=PLZHQObOWTQDNU6R1_67000Dx_ZCJB-3pi&index=4)) YouTube.



## Key Terms

- **Multi-class Classification (in Neural Networks):** A task where each input is assigned to exactly one of several possible classes, using a softmax function in the output layer.
- **Softmax Function:** An activation function that converts a vector of raw scores into probabilities that sum to 1, used in multi-class classification.
- **Forward Pass:** The step in training where input is passed through the network to compute predictions and loss.
- **Backpropagation:** An algorithm that computes the gradient of the loss function with respect to each weight using the chain rule.
- **Gradient:** The vector of partial derivatives of the loss with respect to each weight; it points in the direction of steepest increase.
- **Optimizer:** An algorithm that adjusts a neural network's weights during training to minimize the loss function and improve model performance (e.g., SGD, RMSProp, Adam).
- **Learning Rate:** A hyperparameter controlling the step size in gradient descent updates.
- **Learning Rate Scheduling:** Strategies to adjust the learning rate during training to improve convergence and performance.
- **Momentum:** A technique that accumulates a velocity vector to speed up learning and smooth out noisy updates.