

Week 3 Introduction



Topic Introduction

Week 3 Lesson Video

(15:39)

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 3.1 Monitoring Training Using Learning Curves (6:14)
- Lesson 3.2 Advanced Training Techniques (09:25)



Required Resource

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). O'Reilly Media, Inc.

- Read pages 392-396 in Chapter 11: Training Deep Neural Networks, a useful discussion of various forms of regularization, including dropout.



Key Terms

- **Learning Curve:** A plot showing how a model's performance (typically training and validation loss or accuracy) changes over time, usually as a function of training epochs.
- **Early Stopping:** A regularization technique that halts training when validation performance stops improving, to prevent overfitting and save computation.
- **Exploding Gradients:** A problem during training where gradients become excessively large, causing unstable updates and possibly leading the model weights to diverge.
- **Vanishing Gradients:** A training issue where gradients become very small, especially in deep networks, slowing learning or preventing it altogether.
- **Gradient Clipping:** A technique to prevent exploding gradients by capping the magnitude of gradients during backpropagation.
- **Dropout:** A regularization method that randomly sets a fraction of input or hidden units to zero during training to reduce overfitting.
- **Weight Initialization:** The process of setting initial values for the model's weights before training, which can significantly affect convergence and learning dynamics.