

Go to forum





Week 2

Grades

Notes

Discussion Forums

Messages

Course Info

# Week 2

Introduction to Machine Learning in Production

#### Week 2

Discuss the topic here.

#### Week 2: Select and Train a Model





This week is about model strategies and key challenges in model development. It covers error analysis and strategies to work with different data types. It also addresses how to cope with class imbalance and highly skewed data sets

# Learning Objectives

- · Identify the key challenges in model development.
- Describe how performance on a small set of disproportionately important examples may be more crucial than performance on the majority of examples.
- Explain how rare classes in your training data can affect performance.
  Define three ways of establishing a baseline for your performance.
- . Define structured vs. unstructured data.
- Identify when to consider deployment constraints when choosing a model.
  List the steps involved in getting started with ML modeling.
- . Describe the iterative process for error analysis.

- Describe the Iterative process for error analysis.
  Identify the key factors in deciding what to prioritize when working to improve model accuracy.
  Describe methods you might use for data augmentation given audio data vs. image data.
  Explain the problems you can have training on a highly skewed dataset.
  Identify a use case in which adding more data to your training dataset could actually hurt performance.
  Describe the key components of experiment tracking.

#### Selecting and Training a Model

- ▶ Video: Modeling overview 2 min
- ▶ Video: Key challenges 5 min
- Video: Why low average error isn't good enough 10 min
- ▶ Video: Establish a baseline 7 min
- ▶ Video: Tips for getting started 6 min
- (II) Quiz: Selecting and Training a Model 4 questions Due Dec 27, 1:59 AM CST

# Error analysis and performance auditing

- ▶ Video: Error analysis example 8 min
- **▶ Video:** Prioritizing what to work on 5 min
- ▶ Video: Skewed datasets 12 min
- ▶ Video: Performance auditing 7 min

# Data iteration

- ▶ Video: Data-centric Al development 2 min
- ▶ Video: A useful picture of data augmentation 5 min
- ▶ Video: Data augmentation 8 min
- ▶ Video: Can adding data hurt? 6 min
- ▶ Video: Adding features 8 min
- ▶ Video: Experiment tracking 4 min
- ▶ Video: From big data to good data 3 min
- Reading: Week 2 Optional References 3 min

### Lecture Notes (Optional)

(A) Hagradad External Tool: Lacture Notes W2 5 min



Substance external room sectors moter may some
Graded assessment
(B) Quiz: Modeling challenges 5 questions Due Dec 27, 1:59 AM CST
Ungraded Lab
Lab: A journey through Data 1h