



Week 3

Machine Learning Modeling Pipelines in Production

Overview

- Week 1
- Week 2
- Week 3
- Week 4
- Week 5

Grades

Notes

Discussion Forums

Messages

Course Info

Week 3

Discuss the topic here.

Go to forum

15 threads · Last post 15 days ago

Week 3: High-Performance Modeling

Robert Crowe

Implement distributed processing and parallelism techniques to make the most of your computational resources for training your models efficiently.

Learning Objectives

- Identify the rise in computational requirements with current network architectures
- Select techniques to optimize the usage of computational resources to best serve your model needs
- Carry out high-performance data ingestion to reduce hardware accelerators idling time
- Distinguish between data and model parallelism to train your models in the most efficient way
- Implement knowledge distillation to reduce models that capture complex relationships among features so that they fit in constrained deployment environments

Show Less

High-Performance Modeling

Video: Distributed Training 10 min

Resume

Lab: Distributed Strategies with TF and Keras 1h

Video: High-Performance Ingestion 11 min

Video: Training Large Models - The Rise of Giant Neural Nets and Parallelism 13 min

Reading: High-Performance Modeling 5 min

Practice Quiz: High-Performance Modeling 6 questions

Graded External Tool: Distributed Multi-worker TensorFlow Training on Kubernetes 1h 30m Due Jan 3, 2:59 AM EST

Knowledge Distillation

Video: Teacher and Student Networks 3 min

Video: Knowledge Distillation Techniques 9 min

Lab: Knowledge Distillation 1h

Video: Case Study - How to Distill Knowledge for a Q&A Task 8 min

Reading: Knowledge Distillation 2 min

Practice Quiz: Knowledge Distillation 7 questions

Lecture Notes (Optional)

Ungraded External Tool: Lecture Notes W3 5 min