Databricks - Scalable Deep Learning with TensorFlow and Apache Spark

Code: DB401
Length: 2 days
URL: View Online

This course starts with the basics of the tf.keras API including defining model architectures, optimizers, and saving/loading models. You then learn advanced concepts such as callbacks, regularization, TensorBoard, and activation functions. After training your models, you build integrations with the MLflow tracking API to reproduce and version your experiments. You will apply model interpretability libraries such as LIME and SHAP to understand how the network generates predictions. You will also gain familiarity with Convolutional Neural Networks (CNNs) and how to perform transfer learning to reduce model training time. Substantial class time is spent on scaling your deep learning applications, from distributed inference with pandas UDFs to distributed hyperparameter search with Hyperopt to distributed model training with Horovod. This course is taught fully in Python

Skills Gained

After taking this class, students will be able to:

- Build deep learning models using Keras/TensorFlow
- Tune hyperparameters at scale with Hyperopt
- Track experiments using MLflow
- Apply models at scale using pandas UDFs
- Scale & train distributed models using Horovod
- Apply model interpretability libraries to understand & visualize model predictions
- Use CNNs (convolutional neural networks) and perform transfer learning to reduce model training time
- Implement Generative Adversarial Networks

Who Can Benefit

Data scientist Machine learning engineer

Prerequisites

- Intermediate experience with Python/pandas
- Familiarity with machine learning concepts
- Experience with Spark is helpful, but not required

Course Details

Additional Notes

Additional Notes

- The appropriate, web-based programming environment will be provided to students
- This class is taught in Python only

Topics

Intro to Neural Networks with Keras

- Neural network architectures
- Activation functions
- Evaluation metrics
- Batch sizes, epochs, etc.

MLflow

- Reproducible ML/DL

Convolutional Neural Networks

- Convolutions
- Batch Normalization
- Max Pooling
- ImageNet Architectures

Deep Learning Pipelines

- Model inference at scale

Horovod

- Distributed Tensorflow training
- Ring-All Reduce

Schedule (as of 1)

Date	Location