

Lesson 4.1 : Machine Learning Applications of Spark

DISTRIBUTED COMPUTING WITH SPARK SQL

Machine Learning Applications of Spark



Brooke Wenig/Conor Murphy
Machine Learning Practice Lead/
Data Scientist, Databricks

UC DAVIS
Continuing and Professional Education

Slide 2: Welcome Back!



Welcome Back!

Focused on using Spark with SQL

This module will focus on how you can
use machine learning with Spark SQL

Slide 3: Module Preview



Module Preview

Fundamentals of machine learning

Difference between regression and classification problems

Discuss common evaluation metrics

Show how it ties into the Spark ecosystem

Slide 4: Module Preview



Module Preview

Demonstrate a data scientist workflow

Build a linear regression model

Apply the model by making predictions using Spark SQL

Demonstrate how to extend built-in SQL functions with your user-defined functions

Slide 5: Module Preview



Module Preview

Contextualize machine learning with the rest of the course material

Show how to develop a machine learning model

Tour a data science problem and connect it back to business outcomes

Slide 6: Learning Objectives



Learning Objectives

Demonstrate how a machine learning model is trained

Apply machine learning in SQL using Spark user-defined functions