

Databricks - Introduction to Data Science and Machine Learning

Code: INTRO-DS-ML-SELF

URL: [View Online](#)

In this course data analysts and data scientists practice the full data science workflow by exploring data, building features, training regression and classification models, and tuning and selecting the best model. By the end of this course, you will have built end-to-end machine learning models ready to be launched into production.

Skills Gained

- Contextualize the role of machine learning in the broader technology and business landscape
- Introduce the main topics of supervised machine learning and build a machine learning pipeline in Spark
- Train and evaluate models in a distributed environment
- Perform and interpret exploratory data analysis including statistics and plotting
- Featurize a dataset
- Train linear regression models
- Train logistic regression models
- Tune hyperparameters using grid search and cross-validation

Prerequisites

Getting Started with Apache Spark™ DataFrames self-paced course (optional, but strongly encouraged)

Course Details

Course Outline

- Course Overview and Setup
- What is ML?
- ML Workflows
- Exploratory Analysis
- Featurization
- Regression Modeling
- Classification
- Model Selection
- Capstone Project

Platforms

Supported platforms include Azure Databricks, Databricks Community Edition, and non-Azure Databricks.

- If you're planning to use the course on Azure Databricks, select the "Azure Databricks" Platform option.
- If you're planning to use the course on Databricks Community Edition or on a non-Azure version of Databricks, select the

"Other Databricks" Platform option.

Format

The course is a series of eight self-paced lessons available in both Scala and Python. A final capstone project involves writing an end-to-end machine learning pipeline including exploratory analysis, featurizing the data, training a machine learning model, and tuning model hyperparameters using grid search and cross-validation.

ExitCertified® Corporation and iMVP® are registered trademarks of ExitCertified ULC and ExitCertified Corporation and Tech Data Corporation, respectively
Copyright ©2019 Tech Data Corporation and ExitCertified ULC & ExitCertified Corporation.
All Rights Reserved.

Generated 10