Practical Machine Learning for Cloudera Platform

Learn via: **Classroom / Virtual Classroom / Online**

Duration: **1 Day**

https://bilginc.com/en/training/practical-machine-learning-for-cloudera-platform-960-training/

Overview

Cloudera University's one-day Introduction to Machine Learning with Spark ML and MLlib will teach you the key language concepts to machine learning, Spark MLlib, and Spark ML. The course includes coverage of collaborative filtering, clustering, classification, algorithms, and data volume.

Prerequisites

There are no prerequisites for this course.

Who Should Attend

This course is intended for software engineers who have basic Linux experience in addition to experience with either the Scala or Python programming languages (code examples and exercises are presented in both languages, so students can choose whichever language they prefer).

What You Will Learn

Through instructor-led discussion, as well as hands-on exercises, participants will learn topics including:

- Data types, statistics support, feature extraction, transforming vectors, using the StandardScaler class
- An overview of dimensionality reduction
- Machine learning models, regression, linear regression support, and regularization.
- Finally, the course discusses machine learning with Spark ML topics such as using data frames, transformers and estimators, an introduction to pipelines, using pipelines to generate models, and regularization.

Outline

1. Machine Learning Overview

- Introduction
- Collaborative Filtering
- Clustering
- Classification
- Relationship of Algorithms and Data Volume

2. Machine Learning with Spark MLlib

- Introduction
- Data Types
- Basic Statistics
- Feature Extraction
- Dimensionality Reduction
- Models
- Regression

3. Machine Learning with Spark ML

- Overview of Spark ML
- DataFrames
- Transformers and Estimators
- Pipelines

- Decision Tree Classifiers
- k-Means Clustering