

Trainings

Apache Spark



Zielgruppe: Analysten, Software-Entwickler,
-Architekten

Dauer: 2 Tage

Abstract

Egal ob Batch- oder Stream-Processing – Apache Spark hat sich dank seiner Performance als verteilte In-Memory-Technologie innerhalb von kurzer Zeit einen Stammplatz im Ökosystem der Big Data Tools erarbeitet. Dieses Training führt in den Umgang mit Spark zur Analyse großer Datenmengen ein. Dabei werden sowohl Batch- als auch Streaming-Verfahren diskutiert. Ein Schwerpunkt des Trainings ist die Formulierung von analytischen Anfragen und die Nutzung maschineller Lernverfahren. Ausgehend von konkreten Business-Anforderungen lernen die Teilnehmer geeignete Architekturen, Techniken und Tools kennen, um Lösungen zu implementieren, welche die Business-Anforderungen befriedigen. In diesem Training steht immer die Praxis im Vordergrund: Grundlage des Trainings ist eine komplexe Datenbasis, an welcher Methoden, Tools & Techniken von den Teilnehmern geübt werden.

Agenda:

- Spark Grundlagen und Architektur
- Spark APIs und die RDD Datenstruktur
- Abfragen formulieren mit Spark SQL
- Transformationen und Aktionen im Spark Kontext

- Zeppelin als Spark-Frontend
- Machine Learning mittels der Spark MLlib
- Überblick über das Apache Spark Ökosystem
- Design von Spark-Architekturen zur Umsetzung konkreter Usecases

Hinweis:

- Die Kursgebühr beinhaltet Schulungsunterlagen, Mittagessen, Getränke und Snacks.
- Die Teilnehmer müssen ein eigenes Notebook zum Training mitbringen.

Sollten Sie Fragen zum Training haben, eine Hotelempfehlung benötigen oder ein anderes Anliegen rund um unsere Trainings haben, kontaktieren Sie uns bitte unter trainings@inovex.de oder unseren Academy-Leiter Collin Rogowski unter Tel. +49 (0)172 5673497. Vielen Dank! Wir freuen uns auf Sie!