

Big Data on AWS

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

Duration: **3 Gün**

Overview

AWS üzerinde Büyük Veri eğitimi size, Amazon EMR, Amazon Redshift, Amazon Kinesis ve AWS büyük veri platformunun geri kalan diğer üyeleri gibi bulut tabanlı büyük veri çözümlerini tanıtır. Bu eğitimde Hive ve Hue gibi geniş kapsamlı Hadoop araçlarını kullanarak verileri işlemek için Amazon EMR'yi nasıl kullanacağınızı göstereceğiz. Ayrıca büyük veri ortamlarını nasıl oluşturacağınızı, Amazon DynamoDB, Amazon Redshift, Amazon Quicksight, Amazon Athena ve Amazon Kinesis ile nasıl çalışacağınızı ve güvenlik ve düşük maliyet için büyük veri ortamlarını tasarlarken en iyi uygulamalardan nasıl faydalanacağınızı da öğreteceğiz.

Prerequisites

Herhangi bir ön koşul yoktur.

Who Should Attend

- Büyük veri çözümlerinin tasarlanması ve uygulanmasından sorumlu olan, Çözüm Mimarları gibi kişiler
- AWS üzerindeki büyük veri çözümlerinin ardındaki servisleri ve mimari düzenleri öğrenmek isteyen Veri Bilimadamları ve Veri Analistleri

What You Will Learn

- AWS çözümlerinin büyük veri ekosistemine uyarlanması
- Amazon EMR bağlamında Apache Hadoop'dan faydalanma
- Amazon EMR kümesi bileşenlerini tanımlama
- Amazon EMR kümesini başlatma ve yapılandırma
- Hive, Pig ve Streaming dahil olmak üzere Amazon EMR için kullanılabilir ortak programlama çalışma çerçevelerinden faydalanma
- Amazon EMR'nin kullanım kolaylığını daha da geliştirmek için Hue'dan faydalanma
- Amazon EMR üzerinde Spark ile bellek içi analitikleri kullanma
- Uygun AWS veri depolama çözümlerini seçme
- Büyük verilerin neredeyse gerçek zamanlı olarak işlenebilmesi için Amazon Kinesis'i kullanmanın faydalarını belirleme
- Verileri etkin şekilde saklamak ve analiz etmek için Amazon Redshift'ten faydalanma
- Büyük veri çözümü için maliyetleri ve güvenliği kavrama ve yönetme
- Büyük Veri çözümünü güvence altına alma
- Verilerin özümsemesi, aktarılması ve sıkıştırılması için seçenekleri belirleme
- Özel sorgu analitiği için Amazon Athena'dan faydalanma
- Amazon QuickSight kullanarak verileri ve sorguları betimlemek amacıyla görselleştirme yazılımlarını kullanma
- AWS Data Pipeline kullanarak büyük veri iş akışlarını yönetme

Outline

Module 1

- Overview of Big Data
- Big Data Ingestion and Transfer
- Big Data Streaming and Amazon Kinesis
- Lab 1: Using Amazon Kinesis to Stream and Analyze Apache Server Log Data
- Big Data Storage Solutions
- Big Data Processing and Analytics
- Lab 2: Using Amazon Athena to Query Log Data From Amazon S3

Module 2

- Apache Hadoop and Amazon EMR
- Lab 3: Storing and Querying Data on Amazon DynamoDB
- Using Amazon EMR
- Hadoop Programming Frameworks
- Lab 4: Processing Server Logs With Hive on Amazon EMR
- Web Interfaces on Amazon EMR
- Lab 5: Running Pig Scripts in Hue on Amazon EMR
- Apache Spark on Amazon EMR
- Lab 6: Processing NY Taxi data using Spark on Amazon EMR

Module 3

- Amazon Redshift and Big Data
- Visualizing and Orchestrating Big Data
- Lab 7: Using TIBCO Spotfire to Visualize Data
- Managing Big Data Costs
- Securing Your Amazon Deployments
- Big Data Design Patterns