

Apache Kafka Analitik Eğitimi

Açıklama

Apache Kafka Analitik Eğitimi, Apache Kafka'nın temel özelliklerini, gerçek zamanlı veri akışlarını işleme ve analitik sistemlere entegrasyonunu öğrenmeyi hedeflemektedir. Bu 3 günlük ileri seviye eğitim, Kafka'nın veri toplama, işleme, depolama ve analiz süreçlerinde nasıl kullanılacağını anlamak isteyen analitik uzmanları, veri mühendisleri ve yazılım geliştiriciler için tasarlanmıştır.

Kimler Katılmalı?

- Kafka'ya dayalı projelere başlayan veya Kafka projeleri üzerinde çalışan ekipler
- Veri Bilimciler
- Veri Mühendisleri
- Veri Analistleri

Eğitim İçeriği

1. Gün: Apache Kafka ve Veri Entegrasyonu

Apache Kafka'ya Giriş

- Kafka'nın Veri Analitiğindeki Rolü
- Gerçek Zamanlı Veri Analitiği için Kafka'nın Avantajları
- Kafka'nın Temel Bileşenleri (Producer, Consumer, Topic, Partition, Broker)

Kafka Connect ile Veri Entegrasyonu

- Kafka Connect Yapısı ve Kullanım Alanları
- Source ve Sink Connector'ler
- JDBC Connector ile Veri Alma ve Gönderme
- HDFS ve Object Storage Sink Connector Kullanımı
- Custom Connector Yazma ve Konfigürasyonu

Kafka SQL ile Veri Sorgulama

- KSQLDB Nedir?
- KSQL ile Stream ve Table Kavramları
- KSQL ile Veri Filtreleme ve Transformasyon
- Stream-Stream ve Stream-Table Join İşlemleri
- KSQL ile Gerçek Zamanlı Dashboard Oluşturma

2. Gün: Apache Flink ve Spark ile Gerçek Zamanlı Veri İşleme

Apache Flink ile Kafka Verisini İşleme

- Flink Yapısına Genel Bakış
- Flink ve Kafka Entegrasyonu
- Flink ile Stream Processing
- Flink Windowing Kavramı ve Uygulama Örnekleri
- Flink ile Stateful Processing

Apache Spark ile Kafka Verisini İşleme

- Spark Streaming Yapısına Giriş
- Spark Structured Streaming ile Kafka'dan Veri Okuma
- Spark Streaming ile Window Aggregation ve Filtering
- Spark ile Stream ve Batch İşlemleri Birleştirme
- Checkpointing ve Fault Tolerance Yönetimi

3. Gün: İleri Düzey Uygulamalar ve Performans Optimizasyonu

İleri Düzey Kafka SQL ve Stream Processing

- KSQL ile Aggregation ve Window Functions
- KSQL Persistent Queries
- Performans Optimizasyonu için Partitioning ve Parallelism

Flink ile Gelişmiş Uygulamalar

- Flink State Backend Yönetimi
- Event Time ve Watermark Kullanımı
- Flink ile Veri Anomali Tespiti ve Alerting

Spark ile İleri Düzey Uygulamalar

- Spark Streaming ile ML Pipeline Entegrasyonu
- Spark ve Kafka ile Gerçek Zamanlı Analitik Dashboard
- Spark Streaming Performans Optimizasyonu

Veri Analitiđi için Kafka Ekosistem Araçları

- Kafka Monitoring Araçları (Confluent Control Center, Prometheus, Grafana)
- Schema Registry Kullanımı ve Avro/Protobuf Serileştirme
- Log Compaction ve Cleanup Policy Yönetimi

Ön Koşullar

Java programlama konusunda giriş düzeyi bilgi sahibi olunması gereklidir