
Google Cloud Platform'da Veri Mühendisliği

Google Cloud Platform üzerinde Veri Mühendisliği eğitimi, katılımcılara veri işleme sistemleri tasarlamayı, uçtan uca veri ardışık düzenleri oluşturmayı, verileri analiz etmeyi ve makine öğrenimi gerçekleştirmeyi öğret

Hedefler:

- Google Cloud'da veri işleme sistemleri tasarlamayı ve oluşturmayı
- Cloud Dataflow'da otomatik ölçeklendirme veri ardışık düzenlerini uygulayarak toplu ve akış verilerini işlemeyi
- Google BigQuery kullanarak son derece büyük veri kümelerinden işletme içgörülerini elde etmeyi
- Tensorflow ve Cloud ML kullanarak makine öğrenimi modellerini kullanarak eğitmeyi, değerlendirmeyi ve tahmin etmeyi
- Cloud Dataproc'ta Spark ve ML API'lerini kullanarak yapılandırılmamış verilerden yararlanmayı
- Akış verilerinden anında içgörüler etkinleştirmeyi

Topics:

- Giriş
Google Cloud Dataproc'a Genel Bakış
 - Küme oluşturma ve yönetme.
 - Özel makine türlerinden ve öncelikli çalışan düğümlerinden yararlanma.
 - Kümeleri ölçekleme ve silme.
- Dataproc İşlerini Çalıştırma
 - Pig and Hive işlerini yürütmek.

- Depolama ve hesaplamanın ayrılması.
- Dataproc'u Google Cloud Platform ile entegre etme
 - Kümeyi başlatma eylemleriyle özelleştirin
 - BigQuery Desteği.
- Google'ın Makine Öğrenimi API'leri ile Yapılandırılmamış Verileri Anlamlandırma
 - Google'ın Makine Öğrenimi API'leri
 - Yaygın ML Kullanım Durumları
 - ML API'leri çağırma
- Google BigQuery ve Cloud Dataflow ile Sunucusuz Veri Analizi
 - BigQuery ile Sunucusuz Veri Analizi
 - BigQuery nedir
 - Sorgular ve Fonksiyonlar
 - BigQuery'ye veri yükleme
 - BigQuery'den verileri dışa aktarma
 - İç içe geçmiş ve tekrarlanan alanlar
 - Birden çok tabloyu sorgulama
 - Performans ve fiyatlandırma
- Dataflow ile Sunucusuz, Otomatik Ölçeklendirilen Veri Ardışık Düzenleri
 - Beam programlama modeli
 - Beam Python'da veri ardışık düzenleri
 - Beam Java'da veri ardışık düzenleri

- Beam kullanarak ölçeklenebilir Büyük Veri işleme
- Ek verileri dahil etmek
- Akış verilerini işleme
- GCP Referans mimarisi
- Google Cloud Platform'da TensorFlow ile Sunucusuz Makine Öğrenimi
- Makine Öğrenimine Başlarken
 - Makine öğrenimi (ML) nedir
 - Etkili ML: kavramlar, türler
 - ML veri kümeleri: genelleme
- Tensorflow ile ML Modelleri Oluşturma
 - TensorFlow'a başlarken
 - TensorFlow grafikleri ve döngüleri + lab
 - ML eğitimini izleme
- ML Modellerini CloudML ile Ölçeklendirme
 - Neden Cloud ML?
 - TensorFlow modelini paketleme
 - Uçtan uca eğitim
- Özellik Mühendisliği
 - İyi özellikler yaratmak
 - Girişleri dönüştürme
 - Sentetik özellikler

- Cloud ML ile ön işleme
- Google Cloud Platform'da Esnek Akış Sistemleri Oluşturma
- Akış Analizi Ardışık Düzenlerinin Mimarisi
 - Akış verisi işleme: Zorluklar
 - Değişken veri hacimlerini işleme
 - Sıralanmamış / geç verilerle başa çıkma
- Değişken Hacimleri Besleme
 - Cloud Pub / Sub nedir?
 - Nasıl çalışır: Konular ve Abonelikler
- Akış Ardışık Düzenlerini Uygulama
 - Akış işlemedeki zorluklar.
 - Geç verileri işleyin: filigranlar, tetikleyiciler, biriktirme.
- Akış Analizi ve Gösterge Tabloları
 - Akış analizi: verilerden kararlara
 - BigQuery ile akış verilerini sorgulama
 - Google Data Studio nedir?
- Bigtable ile Yüksek Verimlilik ve Düşük Gecikme
 - Cloud Spanner nedir?
 - Bigtable şeması tasarlama
 - Bigtable'a alma