

Python ile Kapsamlı Veri Bilimi Eğitimi

Eğitim Hakkında

Python ile Veri Bilimi eğitim kursu, mühendislere, veri bilimcilere, istatistikçilere ve diğer kantitatif uzmanlara verileri analiz etmek ve grafiklendirmek için ihtiyaç duydukları Python programlama becerilerini öğretir.

Neler Öğreneceksiniz

- Python temel veri türleri arasındaki farkı anlamayı,
- Farklı python koleksiyonlarını ne zaman kullanacağınızı bilmeyi,
- Python işlevlerini uygulamayı,
- Python'da kontrol akışı yapılarını anlamayı,
- İstisna işleme yapıları aracılığıyla hataları işlemeyi,
- Cevaplanabilir, eyleme geçirilebilir bir soruyu nicel olarak tanımlayabilmeyi,
- Hem yapılandırılmış hem de yapılandırılmamış verileri Python'a aktarmayı,
- Yapılandırılmamış verileri yapılandırılmış biçimlere ayrıştırmayı,
- NumPy dizileri ve pandas veri çerçeveleri arasındaki farkları anlamayı,
- Python'un Python / Hadoop / Spark ekosisteminde nerede olduğunu anlamayı.

Eğitim İçeriği

Temel Python Giriş

Geçmiş ve mevcut kullanım

Yazılımın Kurulması

Python Dağılımları

Dize Değişmezleri ve sayısal nesneler

Koleksiyonlar (listeler, tuplelar, dicts)

Python'daki Datetime sınıfları

Python'da Bellek Yönetimi

Kontrol akışı

Fonksiyonlar

İstisna işleme

Eyleme Geçirilebilir, Analitik Soruları Tanımlama

Soru üzerinde çıkarım yapmak için nicel yapıyı tanımlama

Yapıları desteklemek için gereken verilerin belirlenmesi

Veri ve analitik yaklaşımla ilgili sınırlamaları belirleme

Duyarlılık analizlerinin oluşturulması

Verileri Getirme

Yapılandırılmış Veriler
Yapılandırılmış Metin Dosyaları
Excel çalışma kitapları
SQL veritabanları
Yapılandırılmamış Metin Verileriyle Çalışma
Yapılandırılmamış Metni Okuma
Python ile Doğal Dil İşlemeye Giriş

NumPy: Matrix Dili

Ndarray'e giriş
NumPy işlemleri
Yayın
NumPy'de (maskeli dizi) eksik veri
NumPy Yapılandırılmış diziler
Rastgele sayı üretimi

Pandas ile Veri Hazırlama

Filtreleme
Değişkenler oluşturma ve silme
Sürekli Verilerin Ayırıştırılması
Verileri ölçekleme ve standartlaştırma
Yinelemeleri Tanımlama
Sahte Kodlama
Veri Kümelerini Birleştirme
Verileri Değiştirme
Uzun ve geniş arası

Pandas ile Keşfedici Veri Analizi

Tek Değişkenli İstatistik Özetler ve Aykırı Değerlerin Tespiti
Çok Değişkenli İstatistik Özetler ve Aykırı Değer Tespiti
Pandas kullanarak grup bazında hesaplamalar
Özet Tablolar

Verileri Grafik Olarak İnceleme

Histogram
Kutu ve bıyık arsa
Dağılım grafikleri
Orman arazileri
Grafığe göre gruplama
Matplotlib, Pandas ve Seaborn ile Gelişmiş Grafikler

Python, Hadoop ve Spark

Python, Hadoop ve Spark'taki farka giriş
Spark ve Hadoop'tan Python'a veri aktarma
Spark veya Hadoop'tan yararlanan paralel yürütme

Kayıp veri

Eksik verilerdeki kalıpları keşfetmek ve anlamak

Rastgele Eksik
Rastgele Değil Eksik
Tamamen Rastgele Eksik
Veri isnat yöntemleri

Geleneksel Çıkarımsal İstatistikler

Grupları Karşılaştırma
P-Değerleri, özet istatistikler, yeterli istatistikler, çıkarımsal hedefler
T Testleri (eşit ve eşit olmayan varyanslar)
ANOVA
Ki-Kare Testleri
Korelasyon

Çok Değişkenli İstatistiklere Sık Yaklaşımlar

Doğrusal Regresyon
Çok değişkenli doğrusal regresyon
Doğrusal Olmayan İlişkileri Yakalamak
Model Uyuşlarını Karşılaştırma
Yeni verileri puanlama
Poisson Regresyon Uzantısı

Lojistik regresyon

Lojistik Regresyon Örneği
Sınıflandırma Metrikleri

Çok Değişkenli İstatistiklere Makine Öğrenimi Yaklaşımları

Makine Öğrenimi Teorisi
Veri ön işleme
Kayıp veri
Sahte Kodlama
Standardizasyon
Eğitim / Test verileri

Denetimli ve Denetimsiz Öğrenmeye Karşı

Denetimsiz Öğrenme: Kümeleme

Kümeleme Algoritmaları
Küme Performansını Değerlendirme

Boyutsal küçülme

Önsel
Temel bileşenler Analizi
Cezalandırılmış Regresyon

Denetimli Öğrenme: Regresyon

Doğrusal Regresyon
Cezalandırılmış Doğrusal Regresyon
Stokastik Gradyan İniş
Yeni Veri Kümelerini Puanlama
Çapraz doğrulama
Varyans Yanlılığı-Ticaret

Özellik Önemi

Denetimli Öğrenme: Sınıflandırma

Lojistik regresyon

KEMENT

Rastgele Orman

Topluluk Yöntemleri

Özellik Önemi

Yeni Veri Kümelerini Puanlama

Çapraz doğrulama