

İÇERİK

Data Scientist Ne Yapar?
Data Scientist Çalışma Alanları
Data Scientist Yardımcı Araçlar
Yapay Öğrenme
Yapay Öğrenme Yardımcı Araçlar
Örnek Uygulama Geliştirme

AMAÇ

Oran Teknoloji Uygulamalı Data Scientist eğitimleri ile Veri bilimcinin kullanacağı araçların ve yöntemlerin öğrenilmesi hedeflenmektedir.

Eğitim sonrası katılımcıların;

- Kaynak kod yönetim sistemlerinin, GitHub'ın öğrenilmesi.
- Veri bilimcinin R programlama dilini öğrenmesi.
- R dili ortamının kurulması.
- R dili ile veri analizini yapılması.
- İstatistik yöntemlerinin öğrenilmesi.
- Verinin R ortamına aktarılması, R paketlerinin yüklenmesi, R dilinde hata ayıklamanın yapılması.
- Çeşitli kaynaklardan verinin elde edilme yöntemlerinin öğrenilmesi.
- Veri tabanlarından, web sitelerinden ve API'lerden veriyi elde etme yöntemlerinin öğrenilmesi.
- Verinin temizlenmesi için geliştirilen yöntemlerin öğrenilmesi.
- Verinin edinilmesi, temizlenmesi ve paylaşımı süreçleri öğrenilecektir.
- Veri keşif tekniklerinin öğrenilmesi.
- Görsel teknikler kullanarak veri keşif tekniklerinin öğrenilmesi.
- Çok boyutlu veriler üzerinde görsel tekniklerin uygulamalı olarak öğrenilmesi.
- Tekrarlanabilir araştırma tekniklerinin öğrenilmesi.
- Veri ve üzerinde çalışan R programının paketlenmesi ve dağıtılması.
- Veri üzerinde yapılan analiz sonuçlarının ve R programının paylaşılabilir bir doküman içinde paketlenmesi.
- Yapay Öğrenme Yöntemlerinin Öğrenilmesi.
- Linear Modeller
- Regression, Classification, Naive Bayes konularında detaylı bilgi sahibi olması hedeflenmektedir.

1- Data Scientist Ne Yapar

- Soru sormak
- İdeal veri kümesinin tanımlanması.
- Ulaşılabilen verilerin belirlenmesi.
- Veriyi edinme.
- Veriyi temizleme.
- Keşfedici veri analizi.
- İstatistiksel tahmin ve modelleme.
- Sonuçları yorumlamak.
- Sonuçları doğrulamak.
- Raporlama.
- Tekrarlanabilir metodolojinin hazırlanması.
- Sonuçların kişilerle paylaşılması ve dağıtılması.

2- Data Scientist Çalışma Alanları

- Finans
- Sağlık
- Seçimler
- IoT/M2M
- Siber Güvenlik

3- Data Scientist Yardımcı Araçlar

- R dili
 - R studio
- Python dili
 - NumPy, SciPy, SymPy
 - IPython
 - Pandas

4- Yapay Öğrenme

- Clustering
- Classification
- Regression
- Outlier Detection

5- Yapay Öğrenme Yardımcı Araçlar

- Mahout <https://mahout.apache.org/>
- Spark MLlib <http://spark.apache.org/mllib/>
- Scikit-learn <http://scikit-learn.org/stable/>
- NLTK <http://www.nltk.org/>
- PyBrain <http://pybrain.org/>
- mlpy <http://mlpy.sourceforge.net/>
- PyML <http://pyml.sourceforge.net/>
- Orange <http://orange.biolab.si/>
- PyMVPA <http://www.pymvpa.org/>

6- Örnek Uygulama Geliştirme