

Data Science Bootcamp Capstone Project 2



Proje hakkında

Data: Ecmp_Crdt_Scr_BA.xls (01.03.2018 tarihinde oluşturulan 25 K abone bilgisini içermektedir)

İş Gereksinimi : 30, 60 ve 90 günde ödememe olasılığı olan aboneliklerin tespit edilmesi

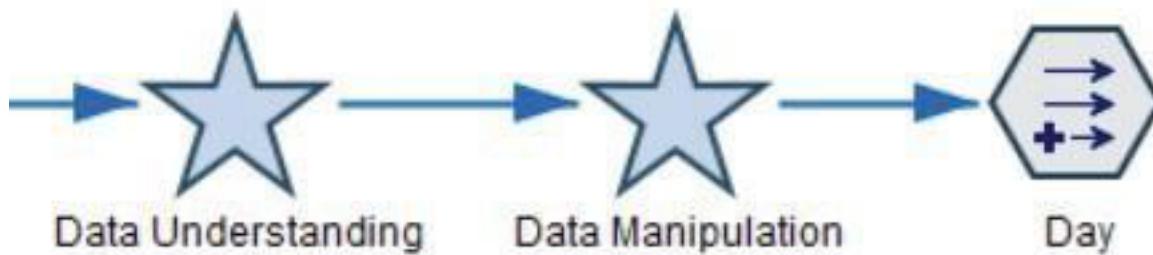
Kullanılacak Yöntem ve Araçlar: IBM SPSS Modeler (Clementine) kullanılarak en iki farklı (istatistiksel ve yapay zeka algoritmaları) kullanılmak suretiyle aşağıdaki işler talep edilmektedir.

- İstatistiksel modelin performansının tartışılması
- Yapay zeka model performansının değerlendirilmesi
- Her iki modelin birlikte kullanılması durumunun ele alınması

Projenin son teslimat tarihi: 26 Eylül, saat 23:59

Çalışmada izlenecek yöntem / varsayımlar:

1. Çalışmanın esası ve ilerleyisi CRISP-DM metodolojisine göre hiyerarşik olmalı
 - İş İhtiyacının Anlaşılması (Business Understanding)
 - Verinin Anlaşılması (Data Understanding)
 - Veri Hazırlama (Data Manipulation, Visulation, Summary Analytics)
 - Modelleme (Modeling)
 - Model Değerlendirme (Model Evaluation)
 - Karar Destek Sistemlerine Entegrasyon (Deployment)
2. Çalışmanız sonucunda elde edeceğiniz stream (ler) ayrı ayrı ya da bütünsel olarak tek bir stream olabilir.
Her bir adım supernode içine alınarak supernode adı verilmeli. Örneğin:



3. Data setine yeni alanlar (variable) eklenmeli; Örneğin; Abonenin yaşı,
Abone_Biyolojik_Yas = XXXXX - Birth_Date, Ortalamalar (Average), Oranlar (Rates), Flag tipte olabilir
4. Verinin Anlaşılması, Verinin görselleştirmisi, Modelleme, Model Değerlendirme adımlarının her birini içerecek şekilde tüm yapılanların özetlerinin anlatacak final bir rapor hazırlanmalı, sunum haline getirilmelidir.
5. Bu çalışma sonuçlarına göre size daha karmaşık iş problemlerinin çözümü ile ilgili yeni bir proje verilebilir.