ALTINBAŞ ÜNİVERSİTESİ

VA520 VERI ANALITIĞI İÇIN MAKINE ÖĞRENMESINDE GELIŞMELER

FİNAL

**Son Teslim tarihi: 1 HAZİRAN**

|  |  |  |  |
| --- | --- | --- | --- |
| Öğrenci No: | 223739001 | Öğrenci Adı-Soyadı | EMRE ÖZYÜREK |
| Proje Adı: | MÜŞTERİ KAYBI ÖNGÖRÜSÜ İÇİN KNIME:  C4.5 DECISION TREE ALGORİTMASI İLE TELCO CHURN ANALİZİ | | |

1. Veri:
2. Veri Seti seçiniz. Seçtiğiniz veri setini anlatınız. (Kategorik, sayısal vb gibi)

Telco Customer Churn veri seti bu proje için seçilmiştir. CSV dosyası olan bu veri seti, 7043 satır 21 sütundan oluşmaktadır. 21 sütunun yapılan analizler sonucunda 17 kategorik, 3 nümerik, 1 kategorik ama kardinal ve 2 nümerik ama kategorik tipte olduğu tespit edilmiştir.

1. Ön İşleme:
2. Elinizdeki veri setinde anlamsız, gereksiz ve fazlalık olan nitelikleri temizleyiniz.

örn: Row Filter , column Filter gibi

1. Dönüştürme:
2. Elinizdeki veri setindeki sayısal niteliklerin bazılarında normalizasyon ve bazılarında ise sepetleme yapmalısınız. Hangi niteliklerde bu işlemi yaptığınız açıklanmalıdır.

Örn: Normalizer, binner (Numeric Binner ya da Auto Binner)

1. Sınıflama algoritmalarından en az birini kullanmalısınız.

Örn: k-means, CART, C4.5

**ÖNEMLİ**: KNIME dosyasınız export ediniz. Kullandığınız düğümleri raporda anlatınız.

Rapor dosyanız pdf halinde olmalıdır. Rapor + Knime dosyasını zip halinde TEAMS grubuna vaktinde yükleyiniz.