

## Veri Ambarlama ve Veri Madenciliği Dönem Projesi (Son Yükleme: 10 Ocak 2022)

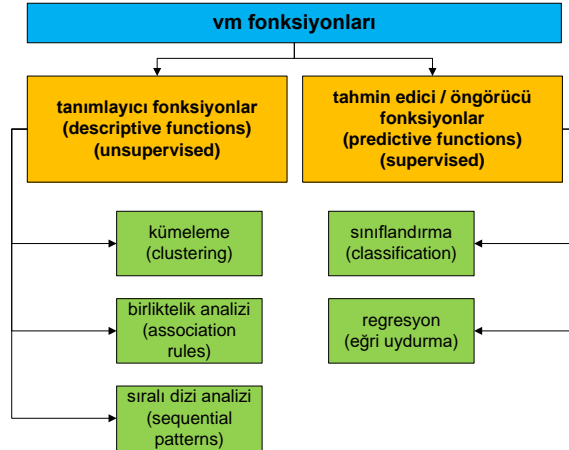
Veri Ambarlama ve Veri Madenciliği Dönem Projesi kapsamında, kendi seçeceğiniz bir konuda orta çapta bir veri madenciliği çalışması yapmanız beklenmektedir. Yapacağınız çalışma, veri madenciliği genel adımlarını içermelidir:

1. Problemin tanımı
2. Veri edinimi
3. Yöntemin belirlenmesi
4. Veri madenciliği fonksiyonunun çalıştırılması
5. Sonuçların elde edilmesi ve yorumlama

Problem tanımı ve veri edinimi aşamalarında ders izlencesinde yer alan veri kaynaklarından yararlanabilirsiniz. Ayrıca dersler sırasında bahsedilmiş kaynakları da kullanabilirsiniz. Bahsedilen sitelerden yararlanarak problem tanımlayabilirsiniz ya da veri setlerini elde edebilirsiniz. UCI sitesinde her veri seti için o veri seti üzerinde çalıştırılacak yöntemler de listelenmiştir, bu şekilde de ilerleyebilirsiniz. Örneğin “sınıflandırma” konusu üzerinde çalışma yapmak istiyorsanız, sınıflandırma yöntemi için uygun olarak işaretlenmiş veri setlerinden birisi üzerinde çalışmaya başlayabilirsiniz. Ya da “kanser araştırmaları” konusu ile ilgileniyorsanız, o konuda bu sitelerde yer alan veri setlerini araştırarak veri setine erişebilirsiniz ve veri seti üzerinde bir problem tanımı yapabilirsiniz.

Uymanız gereken ana kurallar ve genel gereksinimler:

- Dönem sürecinde gördüğümüz veri madenciliği fonksiyonlarından en az birisinin uygulanması zorunludur. Alternative olarak iki yöntem uygulayıp sonuçlarının karşılaştırılması şeklinde bir çalışma da kabul edilebilir niteliktedir.



- Kullanacağınız veri setinde ilk halinde en az 500 adet veri bulunmalıdır. Ön işleme yaparsanız ön işlemler sonucunda bu veri miktarı azalabilir, bu durum kabul edilir niteliktedir. Ancak ön işleme sonrasında veri setinizde ortalama 300 veri bileşeninden daha az sayıda veri kalıyor ise daha hacimli bir veri seti aramanız iyi olacaktır. Veri bileşeni sayısı azaldıkça veriden anlamlı sonuçlar çıkabilme olasılığı da düşmektedir.
- Elde edeceğiniz sonuçları, müşteriye sunulabilecek nitelikte cümlelere çevirerek yorumlayınız.

Yüklemeniz gereken dosyalar:

- Veri seti
- Proje raporu (ekte verilen taslağı kullanmak zorundasınız)
- Proje sunumu (ekte verilen taslağı kullanmak zorundasınız)

Puanlandırma:

- Proje genel kalitesi: 30
- Proje raporu: 40
- Sunum dosyası ve sunumun yapılması: 30

Ödev dosyası ile birlikte örnek projeler de sunulmuştur. Bu projeler geçtiğimiz dönemlerde öğrenciler tarafından yürütölmüş projelerdir. Örnek olarak sadece projelerin sunumları verilmiştir, diğör dokümanlar paylaşılmamıştır. Örnek projelerin sunumlarını inceleyerek sizden bekleneni daha iyi anlayabilirsiniz.

**Belirtilen tarihten sonra yapılan yüklemeler kabul edilmeyecektir. Teslim tarihi ya da sunum tarihinde raporlu olacaksanız raporunuzu ileterek tarihler üzerinde değışiklik yapabiliriz. Rapor belirtmeksizin değışiklik yapılmasını önermeyiniz.**

**Proje sonucunuzu Blackboard'daki Proje ödevine yüklemeniz gerekmektedir. Mail ile sonuç iletmeyiniz.**

**Sunum günü ve saati: 13 Ocak 2022 ders saatlerinde**  
**Sunum ortamı ve şekli: Blackboard / canlı oturum**