

Açıklama ve uyarılar: Ödev *el yazısı ile hazırlanacak* ve yanıtlar pdf dosyasına dönüştürüldükten sonra en geç 16 Ekim tarihinde uzak.etu.edu.tr sistemi üzerinden teslim edilecektir. E-posta vb farklı yollarla gönderilen ya da geç teslim edilen ödevler dikkate alınmayacaktır.

Sorular

1. Yapay zeka, makine öğrenmesi, istatistik ve istatistiksel öğrenme kavramları arasındaki farkı Açıklayınız.

Yanıt: Yapay zeka; insan gibi düşünebilen, davranan, bilinç sahibi olabilen insan yapımı makine ya da bilgisayarlara denir. Makine öğrenmesi, yapay zekayı uygulamaya yönelik bilgisayar programları ve algoritmaları konu alan bilim dalıdır. İstatistik; verileri toplamayı, çözümlenmeyi, yorumlamayı ve anlatmayı konu alan bilim dalıdır. Son olarak, istatistiksel öğrenme ise veriler içindeki ilişki ve örüntüleri anlamaya yönelik geniş kuram ve yöntemler bütünüdür. Dolayısıyla istatistiksel öğrenme, istatistik bilimi ile makine öğrenmesi arasında bir köprü görevi üstlenmektedir.

2. İnternet üzerinde satış yapan bir firma için veri analizi yapmanız istendiğini düşünelim. Firma, stok planlaması yapmak için daha önce satmış olduğu çeşitli ürünlerin iade edilip edilmeyeceğini tahmin etmek istiyor. Bunun için elimizde son 3 ay içinde satılmış olan 60.000 adet giyim ürününe ait fiyat, model, renk, beden, kumaş gibi toplam 20 adet özelliğe ilişkin veriler var. Satılan bu ürünlerin 30 gün içinde iade edilip edilmediği bilgisi de elimizde mevcut.

- (a) Yapacağınız veri çözümlemesinde n ve p değerleri nedir? Açıklayınız.

Yanıt: Bu çözümleme için $n = 60.000$ ve $p = 20$ 'dir.

- (b) Bu çözümleme regresyon mu yoksa sınıflandırma şeklinde mi olacaktır? Açıklayınız.

Yanıt: Ürünlerin iade edilip edilmeyeceğini kestirmek bir *sınıflandırma* çözümlemesi olur. Sınıflandırma yöntemleri çoğu zaman regresyon yaklaşımını kullanmaz ancak regresyona dayalı sınıflandırma yöntemleri de vardır.

- (c) Çözümlemenizde kestirim mi yoksa çıkarsama mı daha önceliklidir? Açıklayınız.

Yanıt: Burada firma için asıl önemli olan insanların iade davranışına neden olan etmenleri incelemek değildir. Firma, maliyetlerini düşürmek için etkili bir *kestirim* yapmak istemektedir.

TOBB - Ekonomi ve Teknoloji Üniversitesi
İKT457 – Ekonomi ve Finans için Yapay Zeka I, Birinci Ödevi

- (d) Kullanacağınız öğrenme yöntemi denetimli mi yoksa denetimsiz mi olacak? Açıklayınız.

Yanıt: Elimizde daha önce satılmış olan ürünlerin iade edilip edilmediği bilgisi bulunmaktadır. Çözümlemede bu tepki değişkenini kullanabileceğimiz için *denetimli* bir öğrenme söz konusudur.

- (e) Kullanacağınız öğrenme yöntemi parametrik mi yoksa parametrik-dışı mı olacak? Açıklayınız.

Yanıt: Sınıflandırmaya yönelik çok sayıda istatistiksel öğrenme yöntemi bulunmaktadır. Dolayısıyla parametrik ya da parametrik-dışı bir yöntem kullanabiliriz.

3. Ekonomi ve finans alanlarında küme çözümlemesinin yararlı olacağı bir uygulama örneği tasarlayınız. Örneğinizdeki girdi ve olası çıktı değişken(ler)ini açıklayınız ve amacınızın kestirim mi yoksa çıkarsama mı olduğunu tartışınız.

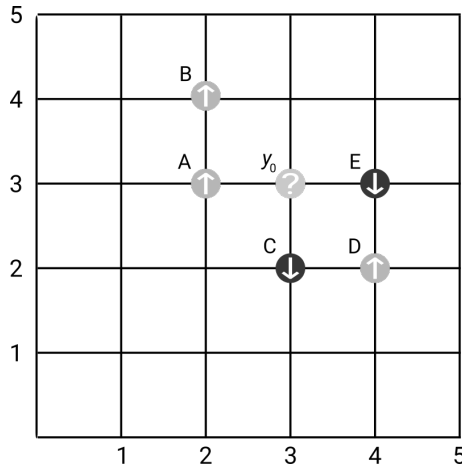
Yanıt: (Burada kitap ya da ders notlarındakilere benzer bir örnek düşünülebilir.)

4. Aşağıdaki çizelgede X_1 ve X_2 şeklinde iki adet nicel değişken ve 1 adet nitel Y değişkenine ait gözlemler verilmiştir. K-enyakın komşu yöntemini kullanarak $x_{01}, x_{02} = (3,3)$ noktasındaki y_0 için sınıf kestirimi yapmak istiyorsunuz.

Gözlem	X_1	X_2	Y
A	2	3	Artış
B	2	4	Artış
C	3	2	Azalış
D	4	2	Artış
E	4	3	Azalış

- (a) Çizelgede verilen A, B, ... değerlerini ve y_0 noktasını grafik üzerinde gösteriniz.

Yanıt:



TOBB - Ekonomi ve Teknoloji Üniversitesi
İKT457 – Ekonomi ve Finans için Yapay Zeka I, Birinci Ödevi

(b) $K = 3$ için “artış” ve “azalış” olasılıklarını hesaplayarak \hat{y}_0 'ı bulunuz.

Yanıt: Şekilde de görüldüğü gibi, y_0 'a en yakın üç nokta A, C ve E'dir. Bu durumda $K = 3$ için artış olasılığı $1/3$, azalış olasılığı ise $2/3$ olur. Dolayısıyla $\hat{y}_0 = azalış$ olarak sınıflandırma yapılır.

(c) $K = 5$ için “artış” ve “azalış” olasılıklarını hesaplayarak \hat{y}_0 'ı bulunuz.

Yanıt: Şekilde de görüldüğü gibi, y_0 'a en yakın beş noktanın tümü dikkate alındığı zaman artış olasılığı $3/5$, azalış olasılığı ise $2/5$ olur. Dolayısıyla bu durumda $\hat{y}_0 = artış$ olarak sınıflandırma yapılır.