

**Açıklama ve uyarılar:** Ödev *el yazısı ile hazırlanacak* ve yanıtlar pdf dosyasına dönüştürüldükten sonra en geç 27 Kasım 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. Türkiye’ye ait il bazlı bir sınıflandırma çözümlemesinde  $X$ = ildeki firma sayısı (bin adet) verilerini kullanarak  $Y$ = ilde sanayi odası olup olmayacağını (yok ise 0, var ise 1) kestirmek istiyoruz. Buna ilişkin lojistik regresyon tahminleri aşağıda verilmiştir:

Call:

```
glm(formula = oda ~ firma, family = binomial, data = veriseti)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-3.2235	-0.3602	-0.3197	-0.2794	2.0263

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-3.3083	0.6350	-5.210	1.89e-07 ***
firma	1.8571	0.6218	2.987	0.00282 **

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 63.766 on 78 degrees of freedom  
Residual deviance: 41.518 on 77 degrees of freedom  
AIC: 45.518

- (a) Tahmin çıktıları kullanarak 1737 adet firmaya sahip bir ilde sanayi odası bulunma olasılığını hesaplayınız. (*Dikkat:* Katsayı tahminlerini ondalık virgülden sonra 1. basamağa yuvarlayınız ve  $e = 3$  kabul ediniz.)

**Yanıt:**

$$\begin{aligned}\Pr(\text{oda} = 1 | \text{firma} = 1737) &= \frac{e^{-3,3+1,9 \times 1,737}}{1 + e^{-3,3+1,9 \times 1,737}} \\ &= \frac{e^0}{1 + e^0} = \frac{1}{1 + 1} = 0,5\end{aligned}$$

- (b) Model tahminlerine göre 1500 adet firmaya sahip bir ilin sınıfı ne olur?

**Yanıt:** Önceki soruda firma sayısı 1737 olduğu zaman sanayi odası bulunma olasılığının 0,5 (yüzde 50) olarak hesaplamıştık. 1500 firmaya sahip bir ilde bu olasılık daha düşük olacaktır. Dolayısıyla bu ilin sınıfı oda = “yok” şeklinde belirlenir.

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

2. Birinci soruda tahmin edilen sınıflandırma modeline ait karışıklık matrisi aşağıdaki gibidir:

		<i>Kestirilen Durum</i>		
		Var	Yok	Toplam
<i>Gerçek Durum</i>	Var	5	8	13
	Yok	1	67	68
Toplam		6	75	81

(a) Sınıflandırmaya ait *doğru pozitif* oranı nedir?

**Yanıt:** Bu soruda pozitif sınıf, sanayi odası *var* olan illerdir. Sanayi odası olan toplam 13 ilden 5'i doğru şekilde sınıflandırılmıştır. Dolayısıyla doğru pozitif oranı:  $5/13 = 0,38$  olarak hesaplanır.

(b) Sınıflandırmaya ait *yanlış pozitif* oranı nedir?

**Yanıt:** Sanayi odası *olmayan* negatif sınıftaki 68 ilden 1'i yanlış olarak *var* (pozitif) şeklinde sınıflandırılmıştır. Dolayısıyla yanlış pozitif oranı:  $1/68 = 0,01$ 'dir.

(c) Sınıflandırmaya ait *doğru negatif* oranı nedir?

**Yanıt:** Burada negatif sınıf sanayi odası *yok* olan iller olduğu için doğru negatif oranı:  $67/68 = 0,99$ 'dur.

3. Birinci soruda kullanılan “veriseti” adlı R veri çerçevesi objesine ait değişken adları aşağıda verilmiştir:

```
> names(veriseti)
[1] "oda"   "firma"
>
```

(a) “oda” adlı nitel değişkeni “firma” değişkenini kullanarak *lojistik* model ile sınıflandıracak R komut(lar)ını yazınız.

**Yanıt:**  
`glm.fit <- glm(oda~firma, data=veriseti, family=binomial)`  
(Dikkat: “glm.fit” nesnesine farklı ad verilebilir.)

(b) “oda” adlı nitel değişkeni “firma” değişkenini kullanarak *doğrusal DÇ* modeli ile sınıflandıracak R komut(lar)ını yazınız.

**Yanıt:**  
`lda.fit <- lda(oda~firma, data=veriseti)`  
(Dikkat: “lda.fit” nesnesine farklı ad verilebilir.)

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

- (c) “oda” adlı nitel değişkeni “firma” değişkenini kullanarak  $K-EK$  ( $k = 3$ ) modeli ile sınıflandıracak R komut(lar)ını yazınız.

**Yanıt:**

```
train <- as.matrix(verisetişfirma)
test <- as.matrix(verisetişfirma)
class <- as.matrix(verisetişoda)
kek.fit<-knn(train, test, class, k=3)
```

(Dikkat: “train”, “test”, “class”, “kek.fit” nesnelere farklı adlar verilebilir.)

4. Doğrusal ve ikinci derece diskriminant çözümlemesine yönelik olarak aşağıdaki soruları yanıtlayınız:

- (a) Bayes karar sınırı eğer doğrusal ise *eğitim veri setinde* doğrusal DÇ'nin mi yoksa ikinci derece DÇ'nin mi daha iyi sonuç vermesini beklersiniz? Kısaca açıklayınız.

**Yanıt:** Bayes karar sınırı (gerçek sınır) doğrusal olsa bile daha esnek bir yöntem olan ikinci derece diskriminant çözümlemesi eğitim veri setinde her zaman daha iyi sonuç verir. Ancak test veri setinde doğrusal DÇ daha iyi sonuç verecektir.

- (b) Bayes karar sınırı eğer doğrusal-dışı ise *test veri setinde* doğrusal DÇ'nin mi yoksa ikinci derece DÇ'nin mi daha iyi sonuç vermesini beklersiniz? Kısaca açıklayınız.

**Yanıt:** Bayes karar sınırının doğrusal-dışı olması durumunda daha esnek bir yöntem olan ikinci derece diskriminant çözümlemesi hem eğitim hem de test veri setlerinde daha iyi sonuç verir.

- (c)  $X$  değişkenleri sayısı arttıkça ikinci derece DÇ çözümlemesinin doğrusal DÇ'ye göre giderek daha iyi sonuç vermesini mi yoksa sonuçların kötüleşmesini mi beklersiniz? Kısaca açıklayınız.

**Yanıt:**  $X$  değişkenlerinin sayısı arttıkça ikinci derece DÇ çözümlemesinde tahmin edilecek parametre sayısı geometrik şekilde artar. Bu durum (veriseti büyüklüğü sabitken) yakıştırmanın giderek kötüleşmesine yol açar.

- (d) Gözlem sayısı arttıkça ikinci derece DÇ çözümlemesinin sade Bayes'e göre giderek daha iyi sonuç vermesini mi yoksa sonuçların kötüleşmesini mi beklersiniz? Kısaca açıklayınız.

**Yanıt:** İkinci derece DÇ çözümlemesi daha esnek bir yöntemdir. Dolayısıyla gözlem sayısı arttıkça bu yöntemin sade Bayes'ten daha iyi sonuçlar vermesini bekleriz.

5. Bahis oranı kavramına ilişkin olarak aşağıdaki soruları yanıtlayınız:

- (a) Bir grup firmaya yönelik iflas bahis oranı eğer 1'e 7 ise bu firmaların ortalamada yüzde kaçının iflas etmesini beklersiniz?

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

**Yanıt:** İflas bahis oranının 1'e 7 olması 8 firmadan 1'inin iflas etmesinin beklendiği anlamına gelir. Bu da yüzde 12,5 demektir.

- (b) Bir grup firmaya yönelik iflas bahis oranı eğer 7'ye 1 ise bu firmaların ortalamada yüzde kaçının iflas etmesini beklersiniz?

**Yanıt:** İflas bahis oranı eğer 7'ye 1 olursa 8 firmadan 7'si iflas edecek demektir. Bu durumda firmaların ortalama yüzde 87,5'inin iflas etmesini bekleriz?