

# BLM0463 Veri Madenciliğine Giriş

## Ödev 1

**İkili sınıflandırma problemi için aşağıdaki tabloda gösterilen eğitim örneklerini göz önüne alınız.**

- (a) Eğitim örneklerinin genel koleksiyonu için Gini endeksini ve entropiyi hesaplayın.
- (b) Customer ID özniteliği için Gini endeksini ve entropiyi hesaplayın.
- (c) Gender özelliği için Gini endeksini ve entropiyi hesaplayın.
- (d) Çok yollu bölme (multiway split) kullanarak Car Type özniteliği için Gini endeksini ve entropiyi hesaplayın.
- (e) Çok yollu bölmeyi kullanarak Shirt Size özelliği için Gini endeksini ve entropiyi hesaplayın.
- (f) Gender, Car Type veya Shirt Size hangi özellik daha iyidir?
- (g) Customer ID en düşük Gini'ye sahip olmasına rağmen öznitelik testi koşulu olarak neden kullanılmaması gerektiğini açıklayın.

Customer ID	Gender	Car Type	Shirt Size	Class
1	M	Family	Small	C0
2	M	Sports	Medium	C0
3	M	Sports	Medium	C0
4	M	Sports	Large	C0
5	M	Sports	Extra Large	C0
6	M	Sports	Extra Large	C0
7	F	Sports	Small	C0
8	F	Sports	Small	C0
9	F	Sports	Medium	C0
10	F	Luxury	Large	C0
11	M	Family	Large	C1
12	M	Family	Extra Large	C1
13	M	Family	Medium	C1
14	M	Luxury	Extra Large	C1
15	F	Luxury	Small	C1
16	F	Luxury	Small	C1
17	F	Luxury	Medium	C1
18	F	Luxury	Medium	C1
19	F	Luxury	Medium	C1
20	F	Luxury	Large	C1

### Ödev teslim şekli:

BTU Moodle sistemi üzerinden bu platformda belirtilen ve bildirilen son tarihten önce pdf olarak teslim edin. Ödev dosya adı için sistematik bir dosya adı kullanın. Örneğin, [BLM463\\_Odev1\\_AdSoyad\\_OgrenciNo.pdf](#), BLM463 dersinin 1. Ödevi için iyi bir örnek dosya adıdır.