

**(a) Eğitim örneklerinin genel koleksiyonu için Gini endeksini ve entropiyi hesaplayın.**

$$\text{Gini} = 1 - p(C0/C0+C1)^2 - p(C1/C0+C1)^2 = 1 - (10/20)^2 - (10/20)^2 = 1 - 0,25 - 0,25 = 0,5$$

$$\text{Entropi} = - (10/20)\log_2(10/20) - (10/20)\log_2(10/20) = 1$$

**(b) Customer ID özniteliği için Gini endeksini ve entropiyi hesaplayın.**

$$\text{Gini} = 1 - (1/1)^2 - (0/1)^2 = 0$$

$$\text{Entropi} = - (1/1)\log_2(1/1) - (0/1)\log_2(0/1) = 0$$

**(c) Gender özelliği için Gini endeksini ve entropiyi hesaplayın.**

$$\text{Gini}(M) = 1 - p(C0/M)^2 - p(C1/M)^2 = 1 - (6/10)^2 - (4/10)^2 = 0,48$$

$$\text{Gini}(F) = 1 - p(C0/F)^2 - p(C1/F)^2 = 1 - (4/10)^2 - (6/10)^2 = 0,48$$

$$\text{Gini} = [ (M) / (M+F) ] * \text{Gini}(M) + [ (F) / (M+F) ] * \text{Gini}(F)$$

$$= (10/20) * 0,48 + (10/20) * 0,48 = 0,48$$

$$\text{Entropi}(M) = - (6/10)\log_2(6/10) - (4/10)\log_2(4/10) = 0,971$$

$$\text{Entropi}(F) = - (4/10)\log_2(4/10) - (6/10)\log_2(6/10) = 0,971$$

$$\text{Entropi} = (10/20) * 0,971 + (10/20) * 0,971 = 0,971$$

**(d) Çok yollu bölme (multiway split) kullanarak Car Type özniteliği için Gini endeksini ve entropiyi hesaplayın.**

$$\text{Gini}(\text{Family}) = 1 - (1/4)^2 - (3/4)^2 = 0,375$$

$$\text{Gini}(\text{Sports}) = 1 - (8/8)^2 - (0/8)^2 = 0$$

$$\text{Gini}(\text{Luxury}) = 1 - (1/8)^2 - (7/8)^2 = 0,219$$

$$\text{Gini}(\text{Car Type}) = (4/20)*0,375 + (8/20)*0 + (8/20)*0,219 = 0,163$$

$$\text{Entropi}(\text{Family}) = - (1/4)\log_2(1/4) - (3/4)\log_2(3/4) = 0,811$$

$$\text{Entropi}(\text{Sports}) = - (8/8)\log_2(8/8) - (0/8)\log_2(0/8) = 0$$

$$\text{Entropi}(\text{Luxury}) = - (1/8)\log_2(1/8) - (7/8)\log_2(7/8) = 0,544$$

$$\text{Entropi} = (4/20)*0,811 + (8/20)*0 + (8/20)*0,544 = 0,379$$

**(e) Çok yollu bölmeyi kullanarak Shirt Size özelliği için Gini endeksini ve entropiyi hesaplayın.**

$$\text{Gini(Small)} = 1 - (3/5)^2 - (2/5)^2 = 0,48$$

$$\text{Gini(Medium)} = 1 - (3/7)^2 - (4/7)^2 = 0,4898$$

$$\text{Gini(Large)} = 1 - (2/4)^2 - (2/4)^2 = 0,5$$

$$\text{Gini(Extra Large)} = 1 - (2/4)^2 - (2/4)^2 = 0,5$$

$$\text{Gini} = (5/20)*0,48 + (7/20)*0,4898 + (4/20)*0,5 + (4/20)*0,5 = 0,491$$

$$\text{Entropi(Small)} = - (3/5)\log_2(3/5) - (2/5)\log_2(2/5) = 0,971$$

$$\text{Entropi(Medium)} = - (3/7)\log_2(3/7) - (4/7)\log_2(4/7) = 0,985$$

$$\text{Entropi(Large)} = - (2/4)\log_2(2/4) - (2/4)\log_2(2/4) = 1$$

$$\text{Entropi(Extra Large)} = - (8/8)\log_2(8/8) - (0/8)\log_2(0/8) = 0$$

$$\text{Entropi} = (5/20)*0,971 + (7/20)*0,985 + (4/20)*1 + (4/20)*0 = 0,788$$

**(f) Gender, Car Type veya Shirt Size hangi özellik daha iyidir?**

Hem Gini endeksi hem de entropisi en düşük değerlere sahip olan özellik daha iyidir. Car Type özelliği en iyi sonucu verir.

**(g) Customer ID en düşük Gini'ye sahip olmasına rağmen öznitelik testi koşulu olarak neden kullanılmaması gerektiğini açıklayın.**

Çünkü tüm ID'ler benzersizdir ve bu özniteliği kullanarak hiçbir bölme yapılamaz. Bu nedenle, öznitelik testi koşulu olarak kullanılmamalıdır.