

CEVAPLAR

1.

Tid	Nesne
1	A, B, C, D
2	A, B, C, E
3	A, B, E, F, H
4	A, C, H

Minimum Destek = %50  
Minimum Güven = %60

a. Eşik destek sayısı =  $0,5 * 4 = 2$

Nesne	Destek değeri
A	4
B	3
C	3
D	1
E	2
F	1
H	2

Nesne kümesi	Destek değeri
A,B	3
A,C	3
A,E	2
A,H	2
B,C	2
B,E	2
B,H	1
C,E	1
C,H	1
E,H	1

Nesne kümesi	Destek değeri
A,B,C	2
A,B,E	2

Yoğun nesne kümesi =  $\{\{A, B, C\}, \{A, B, E\}\}$

b.

Birliktelik	Destek	Güven	Geçerlilik
$AB \rightarrow C$	$\text{sayı}(A,B,C) / N = 2 / 4 = 0.5$	$\text{sayı}(A,B,C) / \text{sayı}(A,B) = 2 / 3 = 0.66$	✓
$B \rightarrow H$	$\text{sayı}(B,H) / N = 1 / 4 = 0.25$	$\text{sayı}(B,H) / \text{sayı}(B) = 1 / 3 = 0.33$	X
$A \rightarrow BE$	$\text{sayı}(A,B,E) / N = 2 / 4 = 0.5$	$\text{sayı}(A,B,E) / \text{sayı}(A) = 2 / 4 = 0.5$	X

2.

Cinsiyet	Gelir	Marka	Satış
Erkek	26K	A	Hayır
Kadın	24K	B	Evet
Kadın	65K	B	Hayır
Erkek	26K	B	Evet
Erkek	34K	B	Evet
Erkek	27K	A	Hayır
Kadın	50K	B	Hayır
Erkek	60K	B	Hayır

a.

$$\text{Entropi} = \text{satış}_{\text{evet}} \log_2 \text{satış}_{\text{evet}} - \text{satış}_{\text{hayır}} \log_2 \text{satış}_{\text{hayır}} = -3/8 \log_2 3/8 - 5/8 \log_2 5/8 = \mathbf{0.954}$$

$$E(\text{Marka}_A) = -2/2 \log_2 2/2 = 0$$

$$E(\text{Marka}_B) = -3/6 \log_2 3/6 - 3/6 \log_2 3/6 = 1$$

$$\text{Gain}(S, \text{Marka}) = 0.954 - [2/8 * E(\text{Marka}_A) + 6/8 * E(\text{Marka}_B)] = \mathbf{0.204}$$

$$E(\text{Cinsiyet}_{\text{Erkek}}) = -2/5 \log_2 2/5 - 3/5 \log_2 3/5 = 0.971$$

$$E(\text{Cinsiyet}_{\text{Kadın}}) = -1/3 \log_2 1/3 - 2/3 \log_2 2/3 = 0.918$$

$$\text{Gain}(S, \text{Cinsiyet}) = 0.954 - [5/8 * E(\text{Cinsiyet}_{\text{Erkek}}) + 3/8 * E(\text{Cinsiyet}_{\text{Kadın}})] = \mathbf{0.003}$$

$$\text{Eşik değerleri} = \{25, 26, 30, 42, 55, 62\}$$

$$E(\text{Gelir}_{\leq 25}) = -1/1 \log_2 1/1 = 0$$

$$E(\text{Gelir}_{>25}) = -2/7 \log_2 2/7 - 5/7 \log_2 5/7 = 0.863$$

$$\text{Gain}(S, \text{Gelir}) = 0.954 - [1/8 * E(\text{Gelir}_{\leq 25}) + 7/8 * E(\text{Gelir}_{>25})] = \mathbf{0.2}$$

$$E(\text{Gelir}_{\leq 26}) = -1/3 \log_2 1/3 - 2/3 \log_2 2/3 = 0.918$$

$$E(\text{Gelir}_{>26}) = -1/5 \log_2 1/5 - 4/5 \log_2 4/5 = 0.722$$

$$\text{Gain}(S, \text{Gelir}) = 0.954 - [3/8 * E(\text{Gelir}_{\leq 26}) + 5/8 * E(\text{Gelir}_{>26})] = \mathbf{0.158}$$

$$E(\text{Gelir}_{\leq 30}) = -2/4 \log_2 2/4 - 2/4 \log_2 2/4 = 1$$

$$E(\text{Gelir}_{>30}) = -1/4 \log_2 1/4 - 3/4 \log_2 3/4 = 0.811$$

$$\text{Gain}(S, \text{Gelir}) = 0.954 - [4/8 * E(\text{Gelir}_{\leq 30}) + 4/8 * E(\text{Gelir}_{>30})] = \mathbf{0.049}$$

$$E(\text{Gelir}_{\leq 42}) = -3/5 \log_2 3/5 - 2/5 \log_2 2/5 = 0.971$$

$$E(\text{Gelir}_{>42}) = -3/3 \log_2 3/3 = 0$$

$$\text{Gain}(S, \text{Gelir}) = 0.954 - [5/8 * E(\text{Gelir}_{\leq 42}) + 3/8 * E(\text{Gelir}_{>42})] = \mathbf{0.347}$$

$$E(\text{Gelir}_{\leq 55}) = -3/6 \log_2 3/6 - 3/6 \log_2 3/6 = 1$$

$$E(\text{Gelir}_{>55}) = -2/2 \log_2 2/2 = 0$$

$$\text{Gain}(S, \text{Gelir}) = 0.954 - [6/8 * E(\text{Gelir}_{\leq 55}) + 2/8 * E(\text{Gelir}_{>55})] = \mathbf{0.204}$$

$$E(\text{Gelir}_{\leq 62}) = -3/7 \log_2 3/7 - 4/7 \log_2 4/7 = 0.985$$

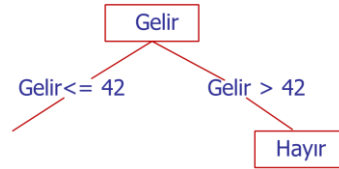
$$E(\text{Gelir}_{>62}) = -1/1 \log_2 1/1 = 0$$

$$\text{Gain}(S, \text{Gelir}) = 0.954 - [7/8 * E(\text{Gelir}_{\leq 62}) + 1/8 * E(\text{Gelir}_{>62})] = \mathbf{0.092}$$

$$\text{Gain}(S, \text{Marka}) = 0.204$$

$$\text{Gain}(S, \text{Cinsiyet}) = 0.003$$

$$\text{Gain}(S, \text{Gelir}) = \mathbf{0.347}$$



$$\text{Entropi} = -2/5 \log_2 2/5 - 3/5 \log_2 3/5 = \mathbf{0.971}$$

$$E(\text{Marka}_A) = -2/2 \log_2 2/2 = 0$$

$$E(\text{Marka}_B) = -3/3 \log_2 3/3 = 0$$

$$\text{Gain}(S_{\text{Gelir}}, \text{Marka}) = 0.971 - [2/5 * E(\text{Marka}_A) + 3/5 * E(\text{Marka}_B)] = \mathbf{0.971}$$

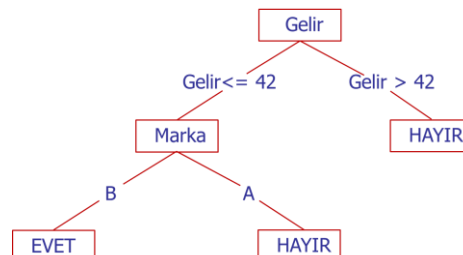
$$E(\text{Cinsiyet}_{\text{Erkek}}) = -2/4 \log_2 2/4 - 2/4 \log_2 2/4 = 1$$

$$E(\text{Cinsiyet}_{\text{Kadın}}) = -1/1 \log_2 1/1 = 0$$

$$\text{Gain}(S_{\text{Gelir}}, \text{Cinsiyet}) = 0.971 - [4/5 * E(\text{Cinsiyet}_{\text{Erkek}}) + 1/5 * E(\text{Cinsiyet}_{\text{Kadın}})] = \mathbf{0.171}$$

$$\text{Gain}(S_{\text{Gelir}}, \text{Marka}) = \mathbf{0.971}$$

$$\text{Gain}(S_{\text{Gelir}}, \text{Cinsiyet}) = 0.171$$



**b.**

- If (Gelir $\leq$ 42)  $\wedge$  (Marka=B) Then Satış= EVET
- If (Gelir $\leq$ 42)  $\wedge$  (Marka=A) Then Satış= HAYIR
- If (Gelir $>$ 42) Then Satış= HAYIR

**3.** Veri madenciliği uygulamalarının güvenilirliğini artırmak için veri ön işleme yapılmalıdır. Toplanan veriler tam ve tutarlı olmayabilir, gürültülü veri içerebilir. Verinin kalitesinin ve sonuçların doğruluğunun artırılması için veri ön işleme gereklidir.

**4.** Bilgi keşfi sürecinin adımları:

- Problemin tanımlanması,
- Verilerin hazırlanması,
- Modelin kurulması ve değerlendirilmesi,
- Modelin kullanılması,
- Modelin izlenmesi