

| HAVA | ISI | NEM | RÜZGAR | OYUN |
|----------|-------|--------|----------|-------|
| Güneşli | Sıcak | Yüksek | Kuvvetli | Hayır |
| Güneşli | Sıcak | Yüksek | Hafif | Evet |
| Bulutlu | Ilık | Yüksek | Kuvvetli | Hayır |
| Yağmurlu | Ilık | Normal | Kuvvetli | Hayır |
| Yağmurlu | Soğuk | Normal | Hafif | Evet |
| Yağmurlu | Soğuk | Normal | Kuvvetli | Hayır |

1. Yukarıda görülen veri setine sınıflandırma işlemi uygulanmış ve elde edilen karar ağacında ISI niteliğinin hiç kullanılmadığı görülmüştür. Buna göre sınıflandırma işleminde aşağıdaki algoritmalarından hangisi kesinlikle kullanılmamıştır? (5,35 Puan)

- a. ID3
- b. Gini
- c. C4.5
- d. Twoing
- e. KNN

| | | Gerçek | |
|--------|---------|---------|---------|
| | | Pozitif | Negatif |
| Tahmin | Pozitif | 12 | 3 |
| | Negatif | 5 | 9 |

2. Yukarıda verilen Karışıklık Matrisine Göre, Hata Oranın değeri aşağıdakilerden hangisidir? (5,35 Puan)

- a. 21/29
- b. 9/12
- c. 9/14
- d. 8/21
- e. 8/29

3. 100 Kayıt, 5 Giriş niteliği ve 1 sınıf niteliğinden oluşan veri setinin %75'i Eğitim, %25'i Test verisi olarak kullanılarak sınıflandırma işlemi yapılmaktadır. Elde edilen Başarım (doğruluk) değeri aşağıdakilerden hangisi olabilir? (4,75 Puan)

- a. 10
- b. 50
- c. 25
- d. 75
- e. 60

4. Aşağıdakilerden hangisi, Birliktelik Kuralında A----->B kuralının güven değeri olabilir. (4,95 Puan)

- a. 1,11
- b. -0,35
- c. -0,51
- d. 3
- e. 0,693

5. Aşağıdakilerden hangisi, Birliktelik Kuralına ait bir Metriktir? (4,5 Puan)

- a. Öklid Uzaklığı
- b. Destek
- c. Kazanç
- d. Gini
- e. Uygunluk Değeri

| | a | b | c | d | e | f | g | h | i | j |
|----|---|---|----|----|----|----|---|----|----|---|
| x1 | 3 | 4 | 3 | -2 | -5 | -7 | 8 | 13 | 11 | 9 |
| x2 | 5 | 8 | -6 | 4 | 5 | 1 | 2 | 7 | -4 | 2 |

6. Yukarıdaki tabloda 10 adet veriye ait X1 ve X2 değerleri verilmektedir. Bu veriler, k-ortalamlar kümeleme algoritması ile 5 kümeye ayrılmak istenmektedir. Başlangıç kümeleri, C1={a, b}, C2={c, d}, C3={e, f}, C4={g, h}, C5={i, j} olarak belirlenmiştir. Algoritmanın birinci tekrarı (iterasyon) sonunda d verisinin yeni kümesi aşağıdakilerden hangisi olur. (5,75 Puan)

- a. C2
- b. C1
- c. C4
- d. C5
- e. C3

| Başvuru | EGİTİM | YAŞ | CİNSİYET | KABUL |
|---------|--------|-------|----------|-------|
| 1 | ORTA | YAŞLI | ERKEK | EVET |
| 2 | İLK | GENÇ | ERKEK | HAYIR |
| 3 | YÜKSEK | ORTA | KADIN | HAYIR |
| 4 | ORTA | ORTA | ERKEK | EVET |
| 5 | İLK | ORTA | ERKEK | EVET |
| 6 | YÜKSEK | YAŞLI | KADIN | EVET |
| 7 | İLK | GENÇ | KADIN | HAYIR |

7. Şekilde verilen veri seti Twoing algoritması ile karar ağacı bulmak isteniyor. “Orta, Yüksek <----->İlk” aday bölünmesinin P(Hayır|tsol) değeri aşağıdakilerden hangisidir? (5,25 Puan)

- a. 0,25
- b. 0,33
- c. 0,75
- d. 0,50
- e. 0,60

8. K-En Yakın Komşu algoritmasında ağırlıklı uzaklık kullanmak için uzaklığa (d) aşağıdaki işlemlerden hangisi uygulanır? (4,50 Puan)

- a. $1/d^2$
- b. $1-d$
- c. d^2
- d. $d^{1,2}$
- e. $d^{2,2}$

| İşlemler | Satın Alınan Ürün Listesi |
|----------|---------------------------|
| T1 | I1, I2, I5 |
| T2 | I2, I4 |
| T3 | I2, I3, I5 |
| T4 | I1, I2, I4 |
| T5 | I1, I3 |
| T6 | I2, I3 |
| T7 | I1, I3 |
| T8 | I1, I2, I3, I5 |

9. Yukarıdaki verilere göre $I1 \rightarrow I2, I3$ birliktelik kuralına ait güven değeri aşağıdakilerden hangisidir? (5,50 Puan)

- a. 0,2
- b. 0,33
- c. 0,4
- d. 0,5
- e. 0,66

10. $X = [1\ 2\ 4\ 7\ 12\ 20\ 35\ 36]$ verisine ait dendrogram, en yakın komşu algoritması ile oluşturulmuştur. Bu dendrogram kullanılarak veriler 5 kümeye ayrılmak istendiğinde, eleman sayısı en fazla olan kümenin eleman sayısı kaç olur. (5.00 Puan)

- a. 2
- b. 6
- c. 3
- d. 5
- e. 4

11. Bir veri setine GINI algoritması uygulanarak karar ağacı oluşturmak istenmektedir. Gini (j) aday bölünmesinin Gini (sol) değeri 0,51 ve Gini (sağ) değeri 0,57 olarak elde edilmiştir. Buna göre Gini(j) değeri aşağıdakilerden hangisi olabilir? (4,75 Puan)

- a. 0,48
- b. 0,71
- c. 0,61
- d. 0,56
- e. 0,66

| HAVA | ISI | NEM | RÜZGAR | OYUN |
|----------|-------|--------|----------|-------|
| Güneşli | Sıcak | Yüksek | Kuvvetli | Hayır |
| Güneşli | Sıcak | Yüksek | Hafif | Evet |
| Bulutlu | Ilık | Yüksek | Kuvvetli | Hayır |
| Yağmurlu | Ilık | Normal | Kuvvetli | Hayır |
| Yağmurlu | Soğuk | Normal | Hafif | Evet |
| Yağmurlu | Soğuk | Normal | Kuvvetli | Hayır |

12. Yukarıda verilen veri setini dikkate alarak $P(\text{HAVA}=\text{Güneşli} \mid \text{ISI}=\text{Sıcak})$ olasılık değerini hesaplayınız? (4,90 Puan)

- a. 0,50
- b. 0,33
- c. 0,75
- d. 0,66
- e. 1,00

| İşlemler | Satın Alınan Ürün Listesi |
|----------|---------------------------|
| T1 | I1, I2, I5 |
| T2 | I2, I4 |
| T3 | I2, I3, I5 |
| T4 | I1, I2, I4 |
| T5 | I1, I3 |
| T6 | I2, I3 |
| T7 | I1, I3 |
| T8 | I1, I2, I3, I5 |

13. Yukarıda verilen tabloda, 8 müşterinin satın aldığı 5 farklı ürün listelenmektedir. Bu veri setine $\text{MinSup}=0,4$ değeri için Apriori algoritması uygulandığında, 2 elemanlı aday nesne kümesinin (C2) eleman sayısı kaç olur? (5,25 Puan)

- a. 5
- b. 6
- c. 4
- d. 2
- e. 3

14. Veri ambarının içeriğinin neler olduğunu belirten, kullanılan verinin yapısını ortaya koyan veri ambarının bileşeni aşağıdakilerden hangisidir? (4,50 Puan)

- a. Ayrıntı Veri
- b. Düşük Düzeyde Özetlenmiş Veri
- c. Metadata
- d. Yüksek Düzeyde Özetlenmiş Veri
- e. Eski Ayrıntı Veri

15. Aşağıdakilerden hangisi veri ambarının özelliklerinden biri değildir? (4,50 Puan)

- a. Konuya Yöneliktir.
- b. Bütünleşiktir.
- c. Sadece Okunabilir.
- d. Dinamik bir yapıya sahiptir.
- e. Zaman Boyutu Vardır.

16. Ekmek alan müşterilerin Simit almadığını gösteren kural, aşağıdaki birliktelik kuralı türlerinden hangisine aittir? (4,55 Puan)

- a. Ağırlıklandırılmış Birliktelik Kuralı
- b. Negatif Birliktelik Kuralı
- c. Periyodik Birliktelik Kuralı
- d. Boolean Birliktelik Kuralı
- e. Hiyerarşik Birliktelik Kuralı

| | | Gerçek | |
|--------|---------|---------|---------|
| | | Pozitif | Negatif |
| Tahmin | Pozitif | 12 | 3 |
| | Negatif | 5 | 9 |

17. Yukarıda verilen Karışıklık Matrisine göre, test verilerinin kaç tanesi Pozitif sınıfına aittir? (5,30 Puan)

- a. 15
- b. 21
- c. 12
- d. 9
- e. 17

| HAVA | ISI | NEM | RÜZGAR | OYUN |
|----------|-------|--------|----------|-------|
| Güneşli | Sıcak | Yüksek | Kuvvetli | Hayır |
| Güneşli | Sıcak | Yüksek | Kuvvetli | Evet |
| Bulutlu | Ilık | Yüksek | Hafif | Hayır |
| Yağmurlu | Ilık | Normal | Kuvvetli | Hayır |
| Yağmurlu | Soğuk | Normal | Hafif | Evet |
| Yağmurlu | Soğuk | Normal | Kuvvetli | Hayır |
| Bulutlu | Soğuk | Normal | Hafif | Hayır |

18. Yukarıda verilen veri setine, ID3 algoritması uygulanarak OYUN sınıfını tahmin eden bir karar ağacı oluşturmak istenmektedir. NEM niteliğine ait Kazanç değeri aşağıdakilerden hangisidir? (5,10 Puan)

- a. 0,184
- b. 0,006
- c. 0,857
- d. 0,918
- e. 0,811

| ID | X1 | X2 | X3 | Sınıf |
|----|----|----|----|-------|
| 1 | 3 | 7 | 6 | C |
| 2 | 2 | 2 | 2 | B |
| 3 | 5 | 6 | 7 | C |
| 4 | 3 | 5 | 8 | C |
| 5 | 1 | 9 | 2 | D |
| 6 | 4 | 3 | 2 | A |
| 7 | 5 | 2 | 6 | B |
| 8 | 6 | 1 | 8 | E |

19. $X = \{4, 4, 4\}$ değerlerine sahip bir veri K-En Yakın Komşu (KNN) algoritması ile sınıflandırıldığında $k=6$ değeri için X verisinin sınıfı aşağıdakilerden hangisi olur? (5,55 Puan)

- a. A
- b. E
- c. C
- d. D
- e. B

20. ????????????? (4,70 Puan)

| | | | |
|--------|---------|---------|---------|
| | | Gerçek | |
| | | Pozitif | Negatif |
| Tahmin | Pozitif | 12 | 3 |
| | Negatif | 5 | 9 |

21. Yukarıda verilen Karışıklık Matrisine Göre, sınıflandırıcının doğru tahmin ettiği kayıt sayısı aşağıdakilerden hangisidir?

- a. 12
- b. 21
- c. 15
- d. 17
- e. 9

| ID | X1 | X2 | X3 | Sınıf |
|----|----|----|----|-------|
| 1 | 3 | 7 | 6 | C |
| 2 | 2 | 2 | 2 | B |
| 3 | 5 | 6 | 7 | C |
| 4 | 3 | 5 | 8 | C |
| 5 | 1 | 9 | 2 | D |
| 6 | 4 | 3 | 2 | A |
| 7 | 5 | 2 | 6 | B |
| 8 | 6 | 1 | 8 | E |

22. $X = \{8, 1, 6\}$ değerlerine sahip bir veri K-En Yakın Komşu (KNN) algoritması ile sınıflandırıldığında $k=1$ değeri için X verisinin sınıfı aşağıdakilerden hangisi olur?

- a. B
- b. D
- c. E
- d. A
- e. C

23. Bir veri setine GINI algoritması uygulanarak karar ağacı oluşturmak istenmektedir. Gini (j) aday bölünmesinin Gini (sol) değeri 0,65 ve Gini (sağ) değeri 0,70 olarak elde edilmiştir. Buna göre Gini(j) değeri aşağıdakilerden hangisi olabilir?

- a. 0,48
- b. 0,66
- c. 0,71
- d. 0,61
- e. 0,56

| İşlemler | Satın Alınan Ürün Listesi |
|----------|------------------------------|
| T1 | I1, I2, I5 |
| T2 | I2, I4 |
| T3 | I2, I3, I5 |
| T4 | I1, I2, I4 |
| T5 | I1, I3 |
| T6 | I2, I3 |
| T7 | I1, I3 |
| T8 | I1, I2, I3, I5 |

24. Yukarıdaki verilere göre $I1 \rightarrow I2, I3$ birliktelik kuralına ait Lift değeri aşağıdakilerden hangisidir?

- a. 0,53
- b. 1,77
- c. 2
- d. 1,33
- e. 1,06

| ID | X1 | X2 | X3 | Sınıf |
|----|----|----|----|-------|
| 1 | 3 | 7 | 6 | C |
| 2 | 2 | 2 | 2 | B |
| 3 | 5 | 6 | 7 | C |
| 4 | 3 | 5 | 8 | C |
| 5 | 1 | 9 | 2 | D |
| 6 | 4 | 3 | 2 | A |
| 7 | 5 | 2 | 6 | B |
| 8 | 6 | 1 | 8 | E |

25. $X = \{5, 5, 5\}$ değerlerine sahip bir veri K-En Yakın Komşu (KNN) algoritması ile sınıflandırıldığında $k=3$ değeri için X verisinin sınıfı aşağıdakilerden hangisi olur?

- a. E
- b. D
- c. C
- d. B
- e. A

26. Aşağıdakilerden hangisi, Birliktelik Kuralında $A \rightarrow B$ kuralının güven değeri olabilir.

- a. -0,556
- b. 2,55
- c. 0,333
- d. -0,71
- e. 1,53

| ID | X1 | X2 | X3 | Sınıf |
|----|----|----|----|-------|
| 1 | 3 | 7 | 6 | C |
| 2 | 2 | 2 | 2 | B |
| 3 | 5 | 6 | 7 | C |
| 4 | 3 | 5 | 8 | C |
| 5 | 1 | 9 | 2 | D |
| 6 | 4 | 3 | 2 | A |
| 7 | 5 | 2 | 6 | B |
| 8 | 6 | 1 | 8 | E |

27. $X = \{-12, 14, 41\}$ değerlerine sahip bir veri K-En Yakın Komşu (KNN) algoritması ile sınıflandırıldığında $k=8$ değeri için X verisinin sınıfı aşağıdakilerden hangisi olur?

- a. B
- b. C
- c. A
- d. E
- e. D

28. Meyve alan müşterilerin Meyve Suyu aldığını gösteren kural, aşağıdaki birliktelik kuralı türlerinden hangisine aittir?

- a. Hiyerarşik Birliktelik Kuralı
- b. Hiçbiri
- c. Boolean Birliktelik Kuralı
- d. Periyodik Birliktelik Kuralı
- e. Ağırlıklandırılmış Birliktelik Kuralı

29. $X = [3 \ 7 \ 13 \ 4 \ 5 \ 10]$ verilerinin standart sapma değeri aşağıdakilerden hangisidir?

- a. 3,84
- b. 12,41
- c. -2,14
- d. 0,00
- e. 7,00

30. Veri Madenciliği araçlarından olan WEKA, aşağıdaki programlama dillerinden hangisi ile geliştirilmiştir.

- a. MatLab
- b. Python
- c. C#
- d. C++
- e. Java

| İşlemler | Satın Alınan Ürün Listesi |
|----------|------------------------------|
| T1 | I1, I2, I5 |
| T2 | I2, I4 |
| T3 | I2, I3, I5 |
| T4 | I1, I2, I4 |
| T5 | I1, I3 |
| T6 | I2, I3 |
| T7 | I1, I3 |
| T8 | I1, I2, I3, I5 |

31. Yukarıdaki verilere göre $I2 \rightarrow I5$ birliktelik kuralına ait Güven değeri aşağıdakilerden hangisidir?

- a. 0,66
- b. 0,2
- c. 0,33
- d. 0,5
- e. 0,4

| İşlemler | Satın Alınan Ürün Listesi |
|----------|------------------------------|
| T1 | I1, I2, I5 |
| T2 | I2, I4 |
| T3 | I2, I3, I5 |
| T4 | I1, I2, I4 |
| T5 | I1, I3 |
| T6 | I2, I3 |
| T7 | I1, I3 |
| T8 | I1, I2, I3, I5 |

32. Yukarıdaki verilere göre $I2 \rightarrow I5$ birliktelik kuralına ait Lift değeri aşağıdakilerden hangisidir?

- a. 1,77
- b. 2
- c. 0,53
- d. 1,33
- e. 1,06

33. Birliklik Kuralında, $X = \text{destek}(A, B) - \text{destek}(A) * \text{destek}(B)$ formülü ile hesaplanan değer aşağıdakilerden hangisidir?

- a. Conviction
- b. Confidence
- c. Coverage
- d. Leverage
- e. Lift

34. Zaman içerisinde olabildiğince birikmiş verilerin oluşturduğu bir veri yığınlarına ne denir?

- a. Veri Ambarı
- b. Veri Küpü
- c. OLTP
- d. OLAP
- e. Veri Madenciliği

35. Aşağıdakilerden hangisi veri indirgeme yöntemlerinden biri değildir?

- a. Genelleme
- b. Örneklem
- c. Sıkıştırma
- d. Bütünleştirme
- e. Veri Küpü

| | | Gerçek | |
|--------|---------|---------|---------|
| | | Pozitif | Negatif |
| Tahmin | Pozitif | 12 | 3 |
| | Negatif | 5 | 9 |

36. Yukarıda verilen Karışıklık Matrisine Göre, Duyarlık (sensitivity) değeri aşağıdakilerden hangisidir?

- a. 12/31
- b. 9/14
- c. 12/17
- d. 12/15
- e. 9/12

37. Çekiş alan müşterilerin Çivi de aldığını gösteren kural, aşağıdaki birliktelik kuralı türlerinden hangisine aittir?

- a. Periyodik Birliktelik Kuralı
- b. Boolean Birliktelik Kuralı
- c. Ağırlıklandırılmış Birliktelik Kuralı
- d. Sıralı Örüntüler
- e. Hiyerarşik Birliktelik Kuralı

| HAVA | ISI | NEM | RÜZGAR | OYUN |
|----------|-------|--------|----------|-------|
| Güneşli | Sıcak | Yüksek | Kuvvetli | Hayır |
| Güneşli | Sıcak | Yüksek | Hafif | Evet |
| Bulutlu | Ilık | Yüksek | Kuvvetli | Hayır |
| Yağmurlu | Ilık | Normal | Kuvvetli | Hayır |
| Yağmurlu | Soğuk | Normal | Hafif | Evet |
| Yağmurlu | Soğuk | Normal | Kuvvetli | Hayır |

38. Yukarıda görülen veri setine sınıflandırma işlemi uygulanmış ve elde edilen karar ağacında ISI niteliğinin hiç kullanılmadığı görülmüştür. Buna göre sınıflandırma işleminde aşağıdaki algoritmalarından hangisi kesinlikle kullanılmamıştır?

- a. YSA
- b. ID3
- c. Twoing
- d. C4.5
- e. Gini

| Başvuru | EĞİTİM | YAŞ | CİNSİYET | KABUL |
|---------|--------|-------|----------|-------|
| 1 | ORTA | YAŞLI | ERKEK | EVET |
| 2 | İLK | GENÇ | ERKEK | HAYIR |
| 3 | YÜKSEK | ORTA | KADIN | HAYIR |
| 4 | ORTA | ORTA | ERKEK | EVET |
| 5 | İLK | ORTA | ERKEK | EVET |
| 6 | YÜKSEK | YAŞLI | KADIN | EVET |
| 7 | İLK | GENÇ | KADIN | HAYIR |

39. Şekilde verilen veri seti Twoing algoritması ile karar ağacı bulmak isteniyor. “Orta, Yüksek <----->İlk” aday bölünmesinin P(Evet | tsağ) değeri aşağıdakilerden hangisidir?

- a. 0,50
- b. 0,60
- c. 0,33
- d. 0,25
- e. 0,75

40. $X = [1\ 2\ 4\ 7\ 12\ 20\ 35]$ verisine ait dendrogram, en yakın komşu algoritması ile oluşturulmuştur. Bu dendrogram kullanılarak veriler 4 kümeye ayrılmak istendiğinde, eleman sayısı en fazla olan kümenin eleman sayısı kaç olur.

- a. 2
- b. 4
- c. 6
- d. 5
- e. 3

| | | Gerçek | |
|--------|---------|---------|---------|
| | | Pozitif | Negatif |
| Tahmin | Pozitif | 12 | 3 |
| | Negatif | 5 | 9 |

41. Yukarıda verilen Karışıklık Matrisine Göre, sınıflandırıcı kaç tane kaydı Negatif olarak tahmin etmiştir?

- a. 12
- b. 17
- c. 15
- d. 9
- e. 14

| İşlemler | Satın Alınan Ürün Listesi |
|----------|---------------------------|
| T1 | I1, I2, I5 |
| T2 | I2, I4 |
| T3 | I2, I3, I5 |
| T4 | I1, I2, I4 |
| T5 | I1, I3 |
| T6 | I2, I3 |
| T7 | I1, I3 |
| T8 | I1, I2, I3, I5 |

42. Yukarıda verilen tabloda, 8 müşterinin satın aldığı 5 farklı ürün listelenmektedir. Bu veri setine $\text{MinSup}=0,2$ değeri için Apriori algoritması uygulandığında, 2 elemanlı yoğun nesne kümesinin (L2) eleman sayısı kaç olur?

- a. 3
- b. 4
- c. 6
- d. 7
- e. 5

