

# Veri Madenciliği Vize ve Final Soruları 2020-2021

Soru )

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

Yukarıda verilen veri setini dikkate alarak  $P(RÜZGAR_{hafif} | HAVAY_{yağmurlu})$  olasılık değerini hesaplayınız.

- a. 0.33      b. 0.25      c. 0.66      d. 0.75      e. 0.50

Soru ) Veri Madenciliği araçlarından olan WEKA yazılımı, aşağıdaki programlama dillerinden hangisi ile geliştirilmiştir?

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

Soru )

		Öngörülen Sınıf	
		Sınıf=1	Sınıf=0
Doğru Sınıf	Sınıf=1	a	b
	Sınıf=0	c	d

a: TP (True Pozitif)      c: FP (False Pozitif)  
b: FN (False Negatif)      d: TN (True Negatif)

- a: TP (True Pozitif)      c: FP (False Pozitif)      b: FN (False Negatif)      d: TN (True Negatif)

Yukarıda verilen karışıklık matrisine göre sınıflandırıcının negatif sınıfını tahmin başarımı aşağıdakilerden hangisidir?

- a.  $a/(a+b)$       b.  $d/(b+d)$       c.  $(c+b)/(a+b+c+d)$       d.  $a/(a+c)$       e.  $d/(d+c)$

**Soru )** "Bebek bezi" alan müşterilerin "Mama" da aldığını gösteren birliktelik kuralı, aşağıdaki birliktelik kuralı türlerinden hangisine aittir?

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

---

**Soru )** Bu ders için hazırlanan ders videolarını yeterince izlediğinize, ödevleri yapmaya çalıştığınıza ve yeterli emeği verdiğinizize inanıyor musunuz? **1 ile 5** arasında bir puan veriniz.

Yanıt: **2**

---

**Soru )**  $x_1 = [3 \ 7 \ 1 \ 4 \ 5 \ 2]$  verisi ile  $x_2 = [2 \ 6 \ 11 \ 4 \ 8 \ 1]$  verisi arasındaki **Manhattan uzaklığı**nı hesaplayınız.

Yanıt: **16**

---

**Soru )**  $A \rightarrow B$  Birliktelik Kuralı için, aşağıdakilerden hangisi, diğerlerine göre **A ile B** arasındaki ilişkinin en güçlü olduğu **lift değeri**dir?

- a. -3.33
- b. 2.5**
- c. 0.33
- d. 0
- e. 0.87

---

**Soru )** Aşağıdakilerden hangisi, Birliktelik Kuralına ait bir metriktir?

- a. İlginçlik**
- b. özgüllük
- c. F-Ölçütü
- d. Minskowski
- e. Anma

---

**Soru )** Elimizde 2 adet zar bulunmaktadır Bu zarlardan birinde **TEK** sayılar (**1,3,5,7,9,11**) sayılar, diğerinde ise **ÇİFT** sayılar (**2,4,6,8,10,12**) bulunmaktadır. Rastgele seçilen bir zar **2 defa atıldığında** gelen sayıların toplamın 10'dan küçük olduğu bilindiğine göre, bu rastgele seçilen zarın **ÇİFT** sayıların olduğu zar olma ihtimali nedir?

- a. 0,25
  - b. 0,50
  - c. 0,66
  - d. 0,75
  - e. 0,33**
-

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

Yukarıda görülen veri seti üzerinde **Gini algoritması** kullanılarak karar ağacı oluşturulmak istenmektedir. "**Güneşli,Bulutlu <-->Yağmurlu**" aday bölünmesine ait **Gini değeri** aşağıdakilerden hangisidir.

- a. 0.50      b. 0.48      **c. 0.44**      d. 0.56      e. 0.37

Soru )

	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

Yukarıda verilen veri setine ait **dendrogram**, **minskowski** uzaklık ölçütü ve en yakın komşu algoritması kullanılarak oluşturulmak istenmektedir. Uzaklık matrisleri hesaplanırken, aşağıdaki değerlerden hangisi elde edilir. (**Not:** Normal yollarla çözmeye çalışırsanız soru çok zaman alacaktır. Mantık yürüterek çözmeye çalışınız.)

- a. 34.21      b. 54.82      c. 4.44      d. 17.83      **e. 12.47**

Soru )

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

Yukarıda görülen veri seti üzerinde **Twoing algoritması** kullanılarak **ISI niteliğini tahmin** edecek bir karar ağacı oluşturulmak istenmektedir. **ISI niteliğinin tahmin** edilebilmesi için **1. iterasyonda** kaç farklı aday bölünme elde edilmektedir.

Yanıt: ---3---

**Soru ) a: TP (True Pozitif) 12 b: FN (Falso Negatif) 3 c: FP (False Pozitif) 5 d: TN (True Negatif) 9**

		Gerçek	
		Pozitif	Negatif
Tahmin	Pozitif	12	3
	Negatif	5	9

**Precision:  $TP/(TP+FP)$ .**

Yukarıda verilen Karışıklık Matrisine göre, **Kesinlik (precision)** değeri aşağıdakilerden hangisidir?

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

**Soru )**

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

Yukarıda verilen veri setini dikkate alarak **P(RÜZGARhafif\HAVAyağmurlu)** olasılık değerini hesaplayınız.

a. 0,33      b. 0,25      c. 0,66      d. 0,75      e. 0,50

**Soru ) a: TP (True Pozitif) 12 b: FN (Falso Negatif) 3 c: FP (False Pozitif) 5 d: TN (True Negatif) 9**

		Gerçek	
		Pozitif	Negatif
Tahmin	Pozitif	12	3
	Negatif	5	9

Yukarıda verilen Karışıklık Martisine göre, test verilerinin kaç tanesi **Negatif sınıfa** aittir?

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

**Soru )**

	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

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}** ve **c5={i,j}** olarak belirlenmiştir. Algoritmanın **birinci tekrarı** (iterasyon) sonunda **b** verisinin yeni kümesi aşağıdakilerden hangisi olur?

- a. c3                      b. c5                      c. c2                      d. c1                      e. c4

**Soru )**

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

Yukarıda verilen verilere göre **|1,|2->|3** birliktelik kuralına **ait Lift değeri** aşağıdakilerden hangisidir?

- a. 2                      b. 1,06                      **c. 1,77**                      d. 1,23                      e. 0,53

**Soru )** Zaman içerisinde olabildiğince birikmiş verilerin oluşturduğu bir veri yığınlarına ne denir?

- a. Veri Ambarı**                      b. Veri Madenciliği                      c. OLAP                      c. Veri Küprü                      c. OLAP

Soru )

Başvuru	EGİTİM	YAS	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

Şekilde verilen veri seti için **Twoing algoritması** ile karar ağacı bulunmak isteniyor. "**Orta,Yüksek<--> İlk**" aday bölünmesinin **P(Evet|tsol)** değeri aşağıdakilerden hangisidir?

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

Soru )

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
Bulutlu	Soğuk	Normal	Kuvvetli	Hayır

Yukarıda verilen veri setine, **ID3** algoritması uygulanarak **OYUN** sınıfını tahmin eden bir karar ağacı oluşturulmak istenmektedir. **NEM niteliğine ait Kazanç** değeri aşağıdakilerden hangisidir?

- a. 0,811      b. **0,857**      c. 0,811      d. 0,006      e. 0,0018

**Soru )** 100 kayıt, 5 giriş niteliği ve 1 sınıf niteliğinden oluşan veri setinin **%80'i** eğitim, **%20'si** test verisi olarak kullanılarak sınıflandırma işlemi yapılmaktadır. Elde edilen **başarım (doğruluk) değeri** aşağıdakilerden hangisi olamaz.

- a. 75      b. 50      c. 25      d. **68**      e. 60

Soru )

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

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

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

**Soru )** Aşağıdakilerden hangisi veri **ambarının** özelliklerinden biri **değildir**?

- a. Konuya yöneliktir.                      b. sadece okunabilir.                      c. Zaman boyutu vardır.  
**d. Dinamik bir yapıya sahiptir.**                      e. Bütünleşiktir.

**Soru )**

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

**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. E                      c. A                      d. D                      e. C

**Soru ) Çekiç** alan müşterilerin daha sonra **Çivi** de aldığını gösteren kural, aşağıdaki **birliktelik kuralı** türlerinden hangisine aittir?

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

**Soru )** Bir veri setine **GINI** algoritması uygulanarak karar ağacı oluşturulmak istenmektedir. **Gini(j)** aday bölünmesinin **Gini(sol)** değeri **0.56** ve **Gini(sağ)** değeri **0.64** olarak elde edilmiştir. Buna göre **Gini(j)** değeri aşağıdakilerden hangisi olabilir?

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

**Soru )** Aşağıdakilerden hangisi, Birliktelik Kuralında **A-->B** kuralının güven değeri olabilir

- a. 1,53 **b. 0,127** c. -0,556 d. 2,55 e. -0,71

**Soru )** Aşağıdaki algoritmalardan hangisi **entropiye dayalı bölümlmeyi** kullanan bir sınıflandırma algoritmadır?

- a. Apriori b. SVM c. Bayes **d. C4.5** e. YSA

**Soru )**

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

Yukarıda verilen verilere göre **|1->|2,|3** birliktelik kuralına **ait güven değeri** aşağıdakilerden hangisidir?

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



**Soru )**

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

Yukanda verilen tabloda, **8 müşteri** satın aldığı 5 farklı ürün listelenmektedir. Bu veri setine **MinSup=0,2** değeri için **Apriori** algoritması uygulandığında, **3 elemanlı** yoğun nesne kümesinin (**L3**) eleman sayısı kaç olur? (**Quiz-3 8. Soru ile benzer**)

- a. 1                      b. 2                      c. 5                      d. 4                      e. 3

---

**Soru ) 100 kayıt, 5 giriş niteliği ve 1 sınıf** niteliğinden oluşan veri setinin **%90'i eğitim, %10'u test verisi** olarak kullanılarak sınıflandırma işlemi yapılmaktadır. Elde edilen **başarım (doğruluk)** değeri aşağıdakilerden hangisi olabilir.

- a. 10                      b. 25                      c. 75                      d. 68                      e. 55

---

**Soru )**

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

Yukarıdaki tabloda, **8 müşteri** ve satın aldıkları ürünler görülmektedir. Bu verilere göre aşağıda istenen kurallara ait değerleri bulunuz. (**Not:** Bulduğunuz **Vim sonuçları ondalıklı** sayı olarak virgöl ile yuvarlama yapmadan giriniz ve virgülden sonra **sadece 2 basamak** kullanınız. **Örnek: 3,00**)



Soru )

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
Güneşli	Soğuk	Yüksek	Kuvvetli	Evet
Bulutlu	Ilık	Normal	Kuvvetli	Hayır

Yukarıda verilen veri seti kullanılarak, **HAVA=Güneşli, ISI=Soğuk, NEM=Yüksek, RÜZGAR=Hafif** verisine ait **OYUN** sınıfı, **Bayes sınıflandırıcı** ile tahmin edilmek istenmektedir. **Ci=EVET** sınıfı için  $P(X|Ci)P(Ci)$  olasılığını hesaplayınız. (**Not:** Bulduğunuz sonucu ondalıklı sayı olarak virgöl ile yuvarlama yapmadan giriniz ve virgülden sonra **sadece 2 basamak** kullanınız. **Örnek: 3,00**)

Yanıt: ----

Yukarıda verilen veri seti kullanılarak, **HAVA=Bulutlu, ISI=Soğuk, NEM=Yüksek, RÜZGAR=Kuvvetli** verisine ait **OYUN** sınıfı, **Bayes sınıflandırıcı** ile tahmin edilmek istenmektedir. **Ci=HAYIR** sınıfı için  $P(X|Ci)P(Ci)$  olasılığını hesaplayınız. (**Not:** Bulduğunuz sonucu ondalıklı sayı olarak virgöl ile yuvarlama yapmadan giriniz ve virgülden sonra **sadece 2 basamak** kullanınız. **Örnek: 3,00**)

Yanıt: ----

Yukarıda verilen veri seti kullanılarak, **HAVA=Bulutlu, ISI=Ilık, NEM=Normal, RÜZGAR=Kuvvetli** verisine ait **OYUN** sınıfı, **Bayes sınıflandırıcı** ile tahmin edilmek istenmektedir. **Ci= HAYIR** sınıfı için  $P(X|Ci)P(Ci)$  olasılığını hesaplayınız. (**Not:** Bulduğunuz sonucu ondalıklı sayı olarak virgöl ile yuvarlama yapmadan giriniz ve virgülden sonra **sadece 2 basamak** kullanınız. **Örnek: 3,00**)

Yanıt: -----

Yukarıda verilen veri seti kullanılarak, **HAVA=Yağmurlu, ISI=Sıcak, NEM=Yüksek, RÜZGAR=Kuvvetli** verisine ait **OYUN** sınıfı, **Bayes sınıflandırıcı** ile tahmin edilmek istenmektedir. **Ci=EVET** sınıfı için  $P(X|Ci)P(Ci)$  olasılığını hesaplayınız. (**Not:** Bulduğunuz sonucu ondalıklı sayı olarak virgöl ile yuvarlama yapmadan giriniz ve virgülden sonra **sadece 2 basamak** kullanınız. **Örnek: 3,00**)

Yanıt: ----

Soru )

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

Yukarıdaki tabloda, **8 müşteri** ve satın aldıkları ürünler görülmektedir. Bu veriler üzerinde **Apriori** algoritması **minDestek** değeri=**0,35** için çalıştırıldığında **3** Elemanlı **Yoğun Nesne** kümesinin kaç tane elemanı olacağını bulunuz. (**L3**'ün eleman sayısı kaçtır?)

Yanıt: -----

Yukarıdaki tabloda, **8 müşteri** ve satın aldıkları ürünler görülmektedir. Bu veriler üzerinde **Apriori** algoritması **minDestek** değeri=**0,28** için çalıştırıldığında **3** Elemanlı **Aday Nesne** kümesinin kaç tane elemanı olacağını bulunuz. (**C3**'ün eleman sayısı kaçtır?)

Yanıt: -----

---

Soru )

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

**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. B                      b. E                      c. A                      d. D                      e. C

Soru )

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
Güneşli	Soğuk	Yüksek	Kuvvetli	Evet
Bulutlu	Ilık	Normal	Kuvvetli	Hayır

Yukarıdaki tabloda verilen veri seti için **Gini algoritması** ile karar ağacı oluşturulmak istenmektedir. "Ilık,Soğuk<-->Sıcak" aday bölünmesinin **Gini değerini** hesaplayınız? (Not: Bulduğunuz sonucu **ondalıklı sayı olarak virgül ile yuvarlama yapmadan** giriniz ve virgülden **sonra sadece 2 basamak** kullanınız. Örnek: 3,00)

Yanıt: -----

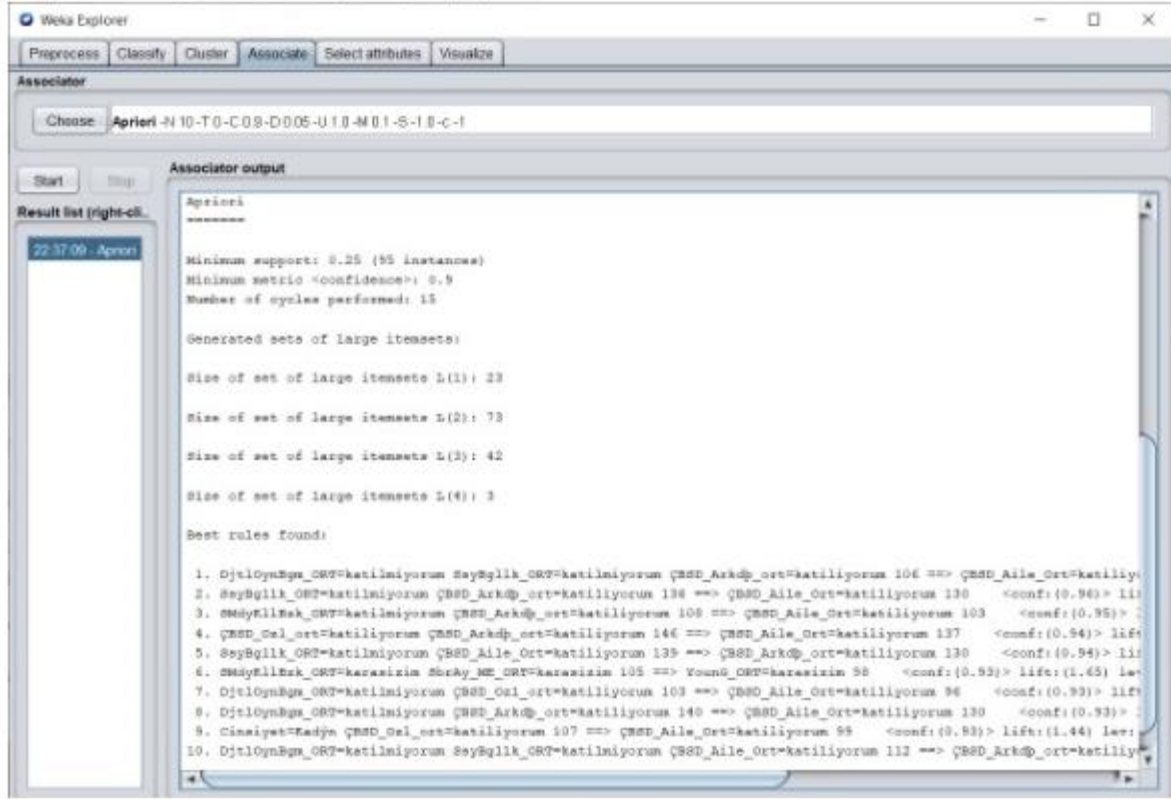
Yukarıdaki tabloda verilen veri seti için **Gini algoritması** ile karar ağacı oluşturulmak istenmektedir. "Güneşli,Soğuk<--> Yağmurlu" aday bölünmesinin **Gini değerini** hesaplayınız? (Not: Bulduğunuz sonucu **ondalıklı sayı olarak virgül ile yuvarlama yapmadan** giriniz ve virgülden **sonra sadece 2 basamak** kullanınız. Örnek: 3,00)

Yanıt: -----

---

**Soru )** Bir veri setinde verileri normalize etme aşağıdakilerden hangisi olarak adlandırılır.

- a. Sıkıştırma
  - b. Bütünleştirme
  - c. İndirgeme**
  - d. Dönüştürme
  - e. Temizleme
-



Yukarıdaki resimde, Weka yazılımı ile Birliktelik Kuralı analizi yapılan bir çalışma görülmektedir. Buna göre aşağıda verilen boşlukları doldurunuz. (**Not:** Boşlukları doldururken şekilde görülen bilgilerle aynı formatta giriş yapınız. **Örnek:** Algoritmanın çalıştırıldığı saat=**22:37:09** gibi)

Çalışmada,  adet veri kullanılmıştır.

Çalışmada,  algoritması kullanılmıştır.

Algoritmada, minimum destek değeri  olarak alınmıştır.

Algoritma, en fazla  elemanlı yoğun nesne kümesi bulmuştur.

6 nolu kuralın güven değeri  olarak elde edilmiştir.

Yukarıdaki resimde, Weka yazılımı ile **Birliktelik Kuralı analizi** yapılan bir çalışma görülmektedir. Buna göre aşağıda verilen boşlukları doldurunuz. (**Not:** Boşlukları doldururken şekilde görülen bilgilerle aynı formatta giriş yapınız. **Örnek:** Algoritmanın çalıştırıldığı saat=**22:37:09** gibi)

Çalışmada, ----- adet veri kullanılmıştır.

Çalışmada, ----- algoritması kullanılmıştır.

Algoritmada, minimum destek değeri ----- olarak alınmıştır.

Algoritma, en fazla ----- elemanlı yoğun nesne kümesi bulmuştur.

6 nolu kuralın güven değeri ----- olarak elde edilmiştir.

Soru )

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

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

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

Soru )

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

**Bayes sınıflandırıcı** kullanarak yukarıda görülen veri seti ile sınıflandırma yapılacaktır Birini **Dışarıda Bırak Çapraz Doğrulama** Test yöntemi kullanılarak elde edilecek başarımlar değeri aşağıdakilerden hangisi olabilir

- a. 0.5                      b. 83                      c. 25                      d. 1.00                      e. 0.30                      f. 0                      g. 75

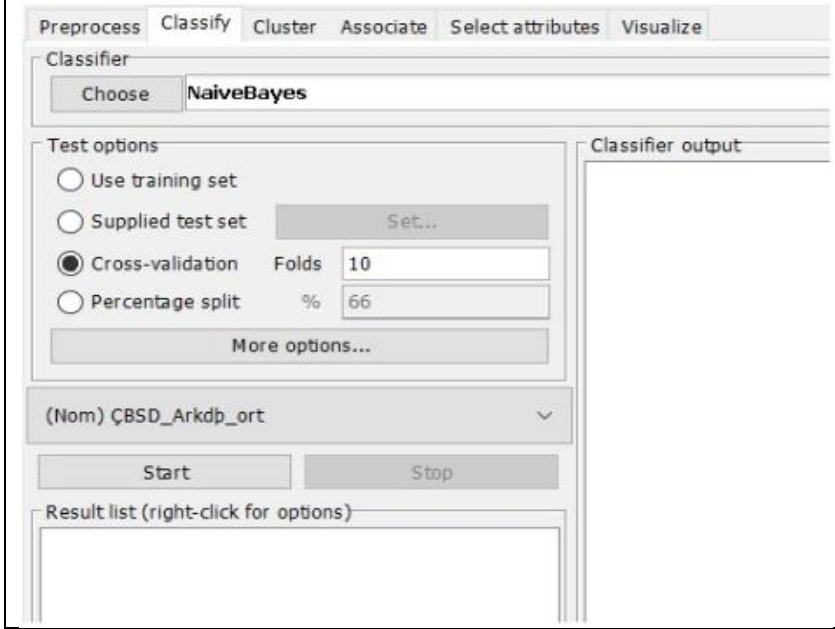
Soru )  $x=[1\ 2\ 4\ 6\ 12\ 20]$  verisine ait **dendrogram** en **uzak komşu algoritması** ile oluşturulmak istenmektedir. Bu dendrogram kullanılarak veriler **2** kümeye ayrılmak istenirse, oluşan her **iki kümenin eleman sayıları** arasındaki fark kaç olur?

- a. 4                      **b. 3**                      c. 2                      d. 1                      e. 0

**Soru )** Aşağıdakilerden hangisi veri indirgeme yöntemlerinden biri değildir?

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

**Soru )**



Şekilde, **Weka programına ait** ekran görüntüsü verilmektedir. **Birini dışarıda bırak (leave one out) çapraz doğrulama test** yöntemi uygulanmak istenirse, **weka programında** şekildeki seçeneklerden hangisi kullanılmalıdır?

a. Supplied test set  
b. Use training set  
c. More options  
d. Percentage split  
**e. Cross-validation**

**Soru )** 100 kayıt, **5 giriş niteliği ve 1 sınıf** niteliğinden oluşan veri setinin **%70'i** eğitim, **%30'u** test verisi olarak kullanılarak sınıflandırma işlemi yapılmaktadır. Elde edilen **başarım (doğruluk)** değeri aşağıdakilerden hangisi olabilir?

- a. 68      b. 55      c. 75      d. 25      e. 10

**Soru )** a: **TP** (True Pozitif) **12**      b: **FN** (False Negatif) **3**      c: **FP** (False Pozitif) **5**      d: **TN** (True Negatif) **9**

		Gerçek	
		Pozitif	Negatif
Tahmin	Pozitif	12	3
	Negatif	5	9

Yukarıda verilen Karışıklık Matrisine göre, **sınıflandırıcı kaç tane kaydı Negatif olarak tahmin** etmiştir?

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



**Soru )** Bir veri setine **GINI** algoritması uygulanarak karar ağacı oluşturulmak istenmektedir. **Gini(j)** aday bölünmesinin **Gini(sol)** değeri **0.56** ve **Gini(sağ)** değeri **0.64** olarak elde edilmiştir. Buna göre **Gini(j)** değeri aşağıdakilerden hangisi olabilir?

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

---

**Soru )**  $x=[2\ 3\ 4\ 61\ 1\ 12\ 15\ 16\ 17\ 24\ 25]$  verisi, **k-ortalamalar** algoritması ile k tane kümeye ayrılacak istenmektedir. **Eleman sayısı en az olan küme ile en fazla olan kümenin eleman sayıları arasındaki farkın 3 olduğu net** olarak bilindiğine göre **k** değeri aşağıdakilerden hangisidir.

Yanıt: ---**3**---

---

**Soru )** Risk={kötü,kötü,kötü,iyi,iyi,iyi} veri grubunun **H(T) Entropi** değerini hesaplayınız.

Yanıt: ---**1**---

---

**Soru )**  $x=[3\ 7\ 13\ 4\ 5\ 10]$  verilerinin **standart sapma** değeri aşağıdakilerden hangisidir.

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

---

**Soru )**  $x=[1\ 2\ 4\ 7\ 12\ 20\ 35\ 36]$  verisine ait **dendrogram**, en yakın komşu algoritması ile oluşturulmuştur. Bu dendrogram kullanılarak veriler **6 kümeye** ayrılacak istendiğinde, eleman sayısı en fazla olan kümenin eleman sayısı kaç olur?

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

---

**Soru )**  $x=[3\ 7\ 13\ 4\ 510]$  verilerine **max-min** normalleştirme uygulanacaktır. Bu işlem yapılırken, **A şıkında** görülen yerinin **yeni değeri** hesaplanırken, **hangi şıkta** verilen değer ile farkının alınması gerekir?

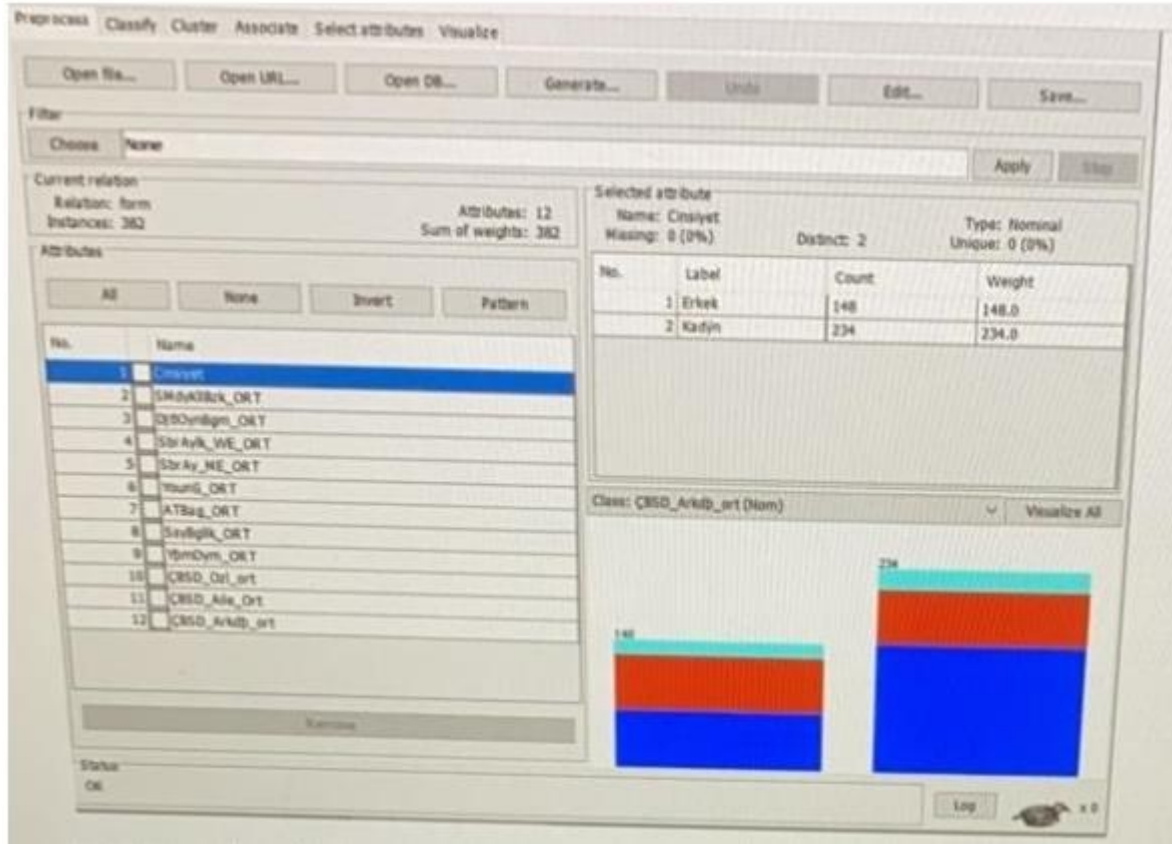
- a. 13      **b.3**      c.4      c.4      e.10

---

**Soru )** Bir veri setine **GINI** algoritması uygulanarak karar ağacı oluşturulmak istenmektedir. **Gini(j)** aday bölünmesinin **Gini(sol)** değeri **0.75** ve **Gini(sağ)** değeri **0.96** olarak elde edilmiştir. Buna göre **Gini(j)** değeri aşağıdakilerden hangisi olabilir?

- a. 0,81      b. 0,56      c. 0,61      d. 0,66      e. 0,71
-

Soru )



Şekilde verilen **Weka ekran** görüntüsüne göre, **Weka'nın varsayılan** olarak kabul ettiği sı niteliği kaç **sınıftan** oluşmaktadır?

- a. 4                      b. 3                      **c. 2**                      c. Hiçbiri                      c. 5

**Soru )** a: **TP** (True Pozitif) **12**      b: **FN** (False Negatif) **3**      c: **FP** (False Pozitif) **5**      d: **TN** (True Negatif) **9**

		Gerçek	
		Pozitif	Negatif
Tahmin	Pozitif	12	3
	Negatif	5	9

Yukarıda verilen Karışıklık Matrisine göre, **sınıflandırıcı doğru tahmin** ettiği kayıt sayısı aşağıdaki hangisi?

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

**Soru )**  $x=[1\ 2\ 4\ 7\ 12\ 20\ 35]$  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. 3                      b. 5                      c. 2                      d. 6                      e. 4

**Soru )**  $x=[1\ 2\ 4\ 7\ 12\ 20\ 35]$  verisine ait **dendrogram**, en yakın komşu algoritması ile oluşturulmuştur. Bu dendrogram kullanılarak veriler **3 kümeye** ayrılmak istendiğinde, eleman sayısı en fazla olan kümenin eleman sayısı kaç olur?

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

**Soru )** Bir veri setine **GINI** algoritması uygulanarak karar ağacı oluşturulmak 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. 0,71                      b. 0,48                      c. 0,66                      d. 0,56                      e. 0,61

**Soru )** a: **TP** (True Pozitif) **12**    b: **FN** (False Negatif) **3**    c: **FP** (False Pozitif) **5**    d: **TN** (True Negatif) **9**

		Gerçek	
		Pozitif	Negatif
Tahmin	Pozitif	12	3
	Negatif	5	9

Yukarıda verilen Karışıklık Matrisine göre, **sınıflandırıcı hatalı tahmin** ettiği kayıt sayısı aşağıdaki hangisi?

- a. 17                      b. 12                      c. 15                      d. 8                      e. 9

**Soru )** Veri ambarı içeriğinin neler olduğunu belirten, kullanılan **yerinin yapısını ortaya koyan** veri ambarının bileşeni aşağıdakilerden hangisidir?

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

**Soru )** Aşağıdaki algoritmalardan hangisi **entropiye dayalı** bölümlmeyi kullanan bir sınıflandırma algoritması ?

- a. YSA                      b. Bayes                      c. Apriori                      d. SVM                      e. **C4.5**

**Soru )** Aşağıdakilerden hangisi, **Birliktelik Kuralına** ait bir metriktir?

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

**Soru ) a**

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

Yukarıda verilen verilere göre **|5->|2,|3** birliktelik kuralına **ait Lift değeri** aşağıdakilerden hangisidir?

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

**Soru )** Bir veri setine **GINI** algoritması uygulanarak karar ağacı oluşturulmak 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?

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

**Soru )** Bir veri setine **GINI** algoritması uygulanarak karar ağacı oluşturulmak istenmektedir. **Gini(j)** aday bölünmesinin **Gini(sol)** değeri **0.68** ve **Gini(sağ)** değeri **0.75** olarak elde edilmiştir. Buna göre **Gini(j)** değeri aşağıdakilerden hangisi olabilir?

- a. 0,81      b. 0,71      c. 0,66      d. 0,61      e. 0,56

**Soru )**

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

Yukarıda verilen veri setini dikkate alarak **P(ISOsoğuk\HAVAyağmurlu)** olasılık değerini hesaplayınız.

- a. 0,50      b. 0,33      c. 0,66      d. 0,75      e. 0,25

**Soru )** Meyve alan müşterilerin **Meyve Suyu** da aldığı gösteren birliktelik kuralı, aşağıdaki birliktelik kuralı türlerinden hangisine aittir?

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

**Soru )**

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

**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. E**      c. A      d. D      e. C

**Soru )**

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
----------	-------	--------	----------	-------

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. Twoing      **b. ID3**      c. Gini      d. YSA      e. C4.5

---

**Soru )**

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

Yukarıda verilen verilere göre **|3->|2,|5** birliktelik kuralına **ait güven değeri** aşağıdakilerden hangisidir?

- a. 0,2      b. 0,33      b. 0,4      b. 0,66      b. 0,5
- 

Yukarıda verilen verilere göre **|2->|5** birliktelik kuralına **ait güven değeri** aşağıdakilerden hangisidir?

- a. 0,2      b. 0,33      b. 0,4      b. 0,66      b. 0,5
- 

Yukarıda verilen verilere göre **|5->|2,|3** birliktelik kuralına **ait güven değeri** aşağıdakilerden hangisidir?

- a. 0,33      b. 0,2      b. 0,4      b. 0,5      b. 0,66

---

**Soru )** Aşağıdakilerden hangisi, Birliktelik Kuralında **A-->B** kuralının **güven değeri** olabilir

- a. 3,51      b. -1,156      c. 4,50      **d. 0,011**      e. -0,171
-

**Soru ) K-En Yakın Komşu algoritmasında ağırlıklı uzaklık kullanmak için uzaklığa (d) aşağıdaki işlemlerden hangisi uygulanır?**

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

**Soru ) 100 kayıt, 5 giriş niteliği ve 1 sınıf niteliğinden oluşan veri setinin %75'i eğitim verisi olarak kullanılarak sınıflandırma işlemi yapılmaktadır. Elde edilen başarımlar değeri aşağıdakilerden hangisi olabilir.**

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

**Soru ) Aşağıdakilerden hangisi, Birliktelik Kuralında  $A \rightarrow B$  kuralının güven değeri olabilir**

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

**Soru ) a: TP (True Pozitif) 12      b: FN (False Negatif) 3      c: FP (False Pozitif) 5      d: TN (True Negatif) 9**

		Gerçek	
		Pozitif	Negatif
Tahmin	Pozitif	12	3
	Negatif	5	9

Yukarıda verilen Karışıklık Matrisine göre, **Duyarlık (sensitivity)** değeri aşağıdaki hangisi?

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

**Soru ) a: TP (True Pozitif) 12      b: FN (False Negatif) 3      c: FP (False Pozitif) 5      d: TN (True Negatif) 9**

		Gerçek	
		Pozitif	Negatif
Tahmin	Pozitif	12	3
	Negatif	5	9

Yukarıda verilen Karışıklık Matrisine göre, **sınıflandırıcı kaç tane kaydı Pozitif olarak tahmin etmiştir?**

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

**a. 68**      b. 55      c. 75      d. 25      e. 10

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

a. 3                      b. 1                      c. 4                      d. 5                      e. 2

a. 3                      b. 1                      c. 4                      d. 5                      e. 2

a. 10                      b. 68                      c. 80                      d. 25                      e. 75



**Soru )** a: **TP** (True Pozitif) **12**    b: **FN** (False Negatif) **3**    c: **FP** (False Pozitif) **5**    d: **TN** (True Negatif) **9**

		Gerçek	
		Pozitif	Negatif
Tahmin	Pozitif	12	3
	Negatif	5	9

Yukarıda verilen Karışıklık Martisine göre, **Hata Oranın** değeri aşağıdaki hangisi?

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

Yukarıda verilen Karışıklık Martisine göre, test verilerinin kaç tanesi **Pozitif sınıfa** aittir?

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

**Soru )** Aşağıdakilerden hangisi, Birlikte Kuralında **A-->B** kuralının **güven değeri** olabilir.

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

**Soru )**  $x=[1\ 2\ 4\ 7\ 12\ 20\ 35\ 36]$  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?

- a. 2      b. 6      a. 3      b. 5      b. 4

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

Yukarıda verilen veri setini dikkate alarak **P(HAVAGüneşli\ISIsıcak)** olasılık değerini hesaplayınız.

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

Soru )

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

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

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

---

**Soru )** Birliktelik 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

---

Soru )

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

Yukanda verilen tabloda, **8 müşterinin** satın aldığı **5 farklı ürün** listelenmektedir. Bu veri setine **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. 5                      e. 7

**Soru )** Aşağıdakilerden hangisi, Birliktelik Kuralında **A-->B** kuralının **güven değeri** olabilir.

**a. 0,01**      b. -2,35      c. 2      d. 1,01      e. -0,5

Güven değeri sadece **0 ile 1** arasında olabilir      yani olasılıklar sadece **0 – 1** olabilir

**Soru )**

		Öngörülen Sınıf	
		Sınıf=1	Sınıf=0
Doğru Sınıf	Sınıf=1	<b>a</b>	<b>b</b>
	Sınıf=0	<b>c</b>	<b>d</b>

**a: TP**(True Pozitif)  
**b: FN**(False Negatif)  
**c: FP**(False Pozitif)  
**d: FN**(True Negatif)

Yukarıda verilen Karışıklık Matrisine göre, **sınıflandırıcı hatalı tahmin** ettiği kayıt sayısı aşağıdaki hangisi?

a. c      b. (c+d)      c. (b+d)      **d. (b+c)**      e. d

**Soru )**

Başvuru	EGİTİM	YAS	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

Şekilde verilen veri seti için **Gini algoritması** ile karar ağacı bulunmak isteniyor. "**İlk,Orta <--> Yüksek**" aday bölünmesinin **Gini(sol)** değeri aşağıdakilerden hangisidir?

a. 0,44      b. 0,38      c. 0,50      d. 0,32      **e. 0,48**

**Soru )** Veri tabanında yer **alan tutarsız ve hatalı** verilere ..... denir.

Yanıt: ---**gürültü**---

Başvuru	EGİTİM	YAS	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

Şekilde verilen veri seti için **Twoing algoritması** ile karar ağacı bulunmak isteniyor. **Psol=4/7** değeri aşağıdakilerden hangisidir?

- a. **Erkek <--> Kadın**                      b. Orta <--> Genç, Yaşlı  
c. Yüksek,Genç <--> İlk                      d. Orta <--> İlk,Yüksek                      d. Orta, Yaşlı <--> Genç

Şekilde verilen veri seti için **Twoing algoritması** ile karar ağacı bulunmak isteniyor. "**Orta,Yaşlı<--> Genç**" aday bölünmesinin **P(Evet | tsağ)** değeri aşağıdakilerden hangisidir?

- a. 0,33                      b. 0,25                      c. 0,67                      d. 0,75                      **e. 0**

		Öngörülen Sınıf	
		Sınıf=1	Sınıf=0
Doğru Sınıf	Sınıf=1	<b>a</b>	<b>b</b>
	Sınıf=0	<b>c</b>	<b>d</b>

**a: TP**(True Pozitif)  
**b: FN**(False Negatif)  
**c: FP**(False Pozitif)  
**d: FN**(True Negatif)

Test veri setinde 1 sınıfına ait toplam kaç kayıt bulunmaktadır?

- a. a+b+c                      b. a                      c. a+c                      **d. a+b**                      e. a+b+c+d

**Soru )**  $x=[1\ 2\ 4\ 7\ 12\ 20\ 33]$  verisine ait **dendrogram**, en yakın komşu algoritması ile oluşturulmuştur. Bu dendrogram kullanılarak veriler **2 kümeye** ayrılmak istendiğinde, eleman sayısı en fazla olan kümenin eleman sayısı kaç olur?

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

**Soru )**

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

Yukarıda verilen verilere göre **|1->|3** birliktelik kuralına **ait güven değeri** aşağıdakilerden hangisidir?

- a. 0,6                      b. 0,2                      c. 0,33                      d. 0,66

**Soru )**

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

Yukanda verilen **Apriori** algoritması çalıştırıldığında, **2 elemanlı yoğun** nesne kümesinin eleman sayısı **3** olarak bulunmuştur. Bun göre Algoritmada kullanılan **Minimum Destek** değeri aşağıdakilerden hangisidir

- a. 0,2                      b. 0,35                      c. 0,45                      d. 0,40                      e. 0,25

**Soru )** Kola alan müşterilerin **Cips de aldığı**nı gösteren kural, aşağıdaki birliktelik kuralı türlerinden hangisine aittir?

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

		Öngörülen Sınıf	
		Sınıf=1	Sınıf=0
Doğru Sınıf	Sınıf=1	<b>a</b>	<b>b</b>
	Sınıf=0	<b>c</b>	<b>d</b>

**a: TP**(True Pozitif)  
**b: FN**(False Negatif)  
**c: FP**(False Pozitif)  
**d: FN**(True Negatif)

Yukarıda verilen karışıklık matrisine göre **doğruluğun(başarım)** değeri aşağıdakilerden hangisidir?

- a.  $a/(a+b)$       b.  $d/(b+d)$       c.  $(c+b)/(a+b+c+d)$       d.  $a/(a+c)$       e.  $d/(d+c)$

**Soru )**

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

Yukarıda verilen veri setine, **ID3** algoritması uygulanarak **OYUN** sınıfını tahmin eden bir karar ağacı oluşturulmak istenmektedir. Oluşacak karar ağacı **kökünü** , aşağıdaki niteliklerden hangisinin oluşturduğunu bulunuz.

- a. OYUN      b. ISI      c. RÜZGAR      d. NEM      e. HAVA

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

Yukarıda verilen veri setine corona virüsü bulaştığı için bazı kayıtlar görülememektedir. Ancak **Gini(j)** aday bölünmesinin **Gini(sol)** değeri **0.48** ve **Gini(sağ)** değeri **0.55** olduğu bilinmektedir. Buna göre **Gini(j)** değeri aşağıdakilerden hangisi olabilir?

- a. 0,57      b. 0,41      c. 0,60      d. 0,36      e. 0,50

Soru )

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

6 kayıt bulunan yukanda veri setine corona virüsü bulaştığı için, **2 niteliğin 2 şer tane veri görülememektedir**. Ancak [ **Yüksek,Düşük<-->Normal** ] aday bölünmesinin Gini değeri ile **Gini(sol)** değerinin eşit olduğu bilinmektedir buna göre **Gini(sağ)** değeri aşağıdakilerden hangisidir?

- a. 0,32      b. 0,51      c. 0      d. 0      e. 0

Soru ) a

Soru ) a