



Puan: 5,00

Soru 1

Aşağıdakilerden hangisi sınıflama modeline ait algoritmalardan biri değildir?

- A** Gini
- B** K-en yakın komşu
- C** ID3
- D** Twoing

E Apriori

Seçimi Boş Bırakmak İstiyorum

Soru 2

Puan: 5,00

Olasılık kuramı ve istatistikten iki rassal değişken arasındaki **doğrusal ilişkinin yönünü ve gücünü belirten** kavram aşağıdakilerden hangisidir?

- A** Scale
- B** Korelasyon
- C** Benzerlik Ölçümü
- D** Regresyon
- E** Nominal

Seçimi Boş Bırakmak İstiyorum

Soru 3

Puan: 5,00

Bayes Sınıflandırıcı, sadece sabit veri setleri için uygun bir sınıflandırma yöntemidir.

A Yanlış

B Doğru

Seçimi Boş Bırakmak İstiyorum

74:56



Veri Madenciliği standart süreci aşağıdakilerden hangisinde tanımlıdır?

A OLAP

B OLTP

C WEKA

D CRISP-DM

E SPSS

Seçimi Boş Bırakmak İstiyorum

Soru 5

Puan: 5,00

Sınıflandırma çalışmalarında performans ölçümü için ROC ANALİZİ kullanılabilir.











A Doğru

B Yanlış

Seçimi Boş Bırakmak İstiyorum

Soru 6

Puan: 5,00

Kişiler	Saç Uzunlukları (inç)	Ağırlık	yaş	Sınıf
 Hasan	0	250	36	E
 Meral	10	150	34	K
 Bahadır	2	90	10	E
 Lale	6	78	8	K
 Melike	4	20	1	K
 Ali	1	170	70	E
 Selma	8	160	41	K
 Osman	10	180	38	E
 Kemal	6	200	45	E
 Cemal	8	290	38	?

Yukarıdaki öğrenme seti için genel entropi değeri kaçtır?

74:56

**B** 0,8111**C** 0,9232**D** 0,9911**E** 0

Seçimi Boş Bırakmak İstiyorum

Soru 7

Puan: 5,00

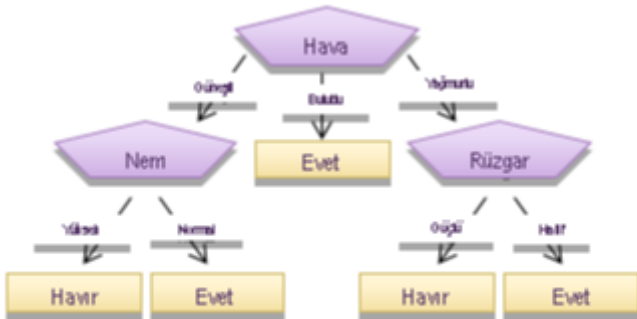
CLUCDUH kategorik verilerin kümelenmesi için kullanılabilecek bir algoritmadır.

A Yanlış**B** Doğru

Seçimi Boş Bırakmak İstiyorum

Soru 8

Puan: 5,00



Yukarıdaki karar ağacına göre hangisi yazılamaz?

A Eğer Hava=Yağmurlu ise veya Eğer Rüzgar=Hafif ise Migren =Hayır**B** hiçbiri**C** Eğer Hava=Bulutlu ise Migren =Evet**D** Eğer Hava=Güneşli ise ve Eğer Nem=Yüksek ise Migren=Hayır**E** Eğer Hava=Güneşli ise ve Eğer Nem=Normal ise Migren =Evet

Seçimi Boş Bırakmak İstiyorum



.... algoritması nitelik değerlerinin sol ve sağda olmak üzere ikili bölünmeler şeklinde ayrılması temeline dayanır.

A GINI

B C4.5

C ID3

D K_ORTALAMA

E APRIORI

Seçimi Boş Bırakmak İstiyorum

Soru 10

Puan: 5,00

Veri madenciliği için kullanılacak olan veri setinde, eksik veri niteliklerinin doldurulması için bir yöntem yoktur.

A Doğru

B Yanlış

Seçimi Boş Bırakmak İstiyorum

Soru 11

Puan: 5,00

Sınıflandırıcı Sonucu	Gerçek durum			
		POZİTİF	NEGATİF	TOPLAM
	POZİTİF	85	250	335
	NEGATİF	6	21	27
	TOPLAM	91	271	362

Yukarıdaki verilere göre ROC doğruluk değeri %92 dir

A Doğru

B Yanlış

Seçimi Boş Bırakmak İstiyorum

Puan: 5,00



gerekmektedir. Kurulan bu modele regresyon modeli denilmektedir.

Regresyon analizi yapılırken, gözlem değerlerinin ve etkilenilen olayların bir matematiksel gösterimle yani bir fonksiyon yardımıyla ifadesi gerekmektedir. Kurulan bu modele regresyon modeli denilmektedir.

A Yanlış

B Doğru

Seçimi Boş Bırakmak İstiyorum

Soru 13

Puan: 5,00

Sıra	Kıtap Türü
1	Migren, Anemi, Bronşit, Astım
2	Migren, Anemi, Astım
3	Astım, Diyabet
4	Diyabet, Troyid
5	Migren, Anemi
6	Migren, Anemi, Diyabet

Yukarıdaki öğrenme setine , %30 destek değerine (Bir üst değere yuvarlama yapınız) Apriori algoritması uygulandığında hangi hastalık ilk süreçte elenir?

A Migren

B Astım

C Anemi

D Troyid

E Diyabet

Seçimi Boş Bırakmak İstiyorum

Soru 14

Puan: 5,00

Aşağıdakilerden hangisi k-en yakın komşu algoritması için doğru bir tanımlama değildir?

A Kayıtlar, bir veri uzayındaki noktalar olarak düşünülürse, birbirine yakın olan kayıtlar, birbirinin civarında (yakın komşusu) olur

B Olayların ait olduğu sınıflar sayılır ve yeni durum sayısı fazla olan sınıfa dahil edilir.

C Olaylar arasındaki uzaklıklar hesaplandıktan sonra, yeni olayların sınıflandırılması için hâlihazırda sınıflandırılmış olan durumlar temel olarak alınır.

D K en yakın komşuluğunda temel düşünce “komşunun yaptığıını sakın yapma” dır.

E Komşuluk hesaplamaları yapılırken, daha yakın komşulara daha büyük ağırlık değerleri atanabilir

Seçimi Boş Bırakmak İstiyorum



“Doğal bağışıklık sisteminden esinlenerek, doğal sistemin prensiplerinin mühendislik açısından incelenerek karmaşık problemlerin çözümünde ve sistemlerin gelişiminde kullanılan sezgisel yöntemlerden biri” şeklinde tanımlanan yöntem hangisidir?

- A** Yapay sinir ağları
- B** Yapay Bağışıklık sistemi
- C** Uzman sistemler
- D** Bulanık sistemler
- E** Genetik algoritmalar

Seçimi Boş Bırakmak İstiyorum

Soru 16

Puan: 5,00

Veri Madenciliği Sürecinde son süreç aşağıdakilerden hangisidir?

- A** Bilginin üretilmesi, eylem planına dönüştürülmesi ve sonuçların ölçülüp değerlendirilmesi
- B** Verilerin bulunması ve analizi
- C** Modelin oluşturulması
- D** Kullanılacak verilerin seçilmesi ve hazırlanması
- E** Problemin Tanımı

Seçimi Boş Bırakmak İstiyorum

Soru 17

Puan: 5,00

Bilgi "information" için aşağıdakilerden hangisi doğru bir açıklama değildir?

- A** Verilerin istatistiksel ve matematiksel yöntemlerle analiz edilerek anlamlandırılmasıdır
- B** Grafikselleştirilerek verilerin sunulmasıdır
- C** İnsanlar, nesneler, işlemler, uygulamalar, olaylarla ilgili gerçekleri yansıtan niceliksel veya niteliksel değerlerdir
- D** Verilerin toplandıktan sonra sınıflandırılmasıdır
- E** Ortalama, mod, standart sapma ve benzeri istatistiksel ölçümlerle verilerin özetlenmesidir

74:56

**Soru 18**

Puan: 5,00

Klonal Seçim Prensibi ile verileri çoğaltabiliriz.

A Yanlış**B** Doğru

Seçimi Boş Bırakmak İstiyorum

Soru 19

Puan: 5,00

Veri Ambarlarının Kullanım amacı için hangisi söylenbilir?

A Müşterilerin gizli kalmış satın alma eğilimlerini tespit etmek**B** Satış analizi ve trendler üzerine odaklanmak**C** Finansal analiz yapabilmek,**D** HEPSİ**E** Karar destek sistemleri için Stratejik Analiz yapabilmek.

Seçimi Boş Bırakmak İstiyorum

Soru 20

Puan: 5,00

Seçicilik: Testin, gerçek pozitif durumlar içinden pozitif olan durumları ayırma yeteneğidir.

A Yanlış**B** Doğru

Seçimi Boş Bırakmak İstiyorum

Cevapları Gözden Geçir

74:56

