

**ULUDAĞ ÜNİVERSİTESİ**  
**MÜHENDİSLİK FAKÜLTESİ**  
**BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ**  
**2023-2024 Güz Dönemi**  
**BMB3015 Yapay Zeka**  
**FİNAL SINAVI ÇÖZÜMÜ**

**Süre: 75 dakika**

- 1) Aşağıdaki cümlelerde geçen boşlukları uygun ifadelerle doldurunuz.
- a) Toplamda n tane değişken barındıran bir Bayes ağında  $X_1, X_2, \dots, X_k$  değişkenlerinin atası bulunmuyorsa bu değişkenlerin **bağımsız** olduğu anlaşılır
  - b) Çapraz geçerlemede veri kümesi genelde eğitim, test ve **validasyon** olarak üçe bölünür.
  - c) Bir modelin veri kümesine az uyumu halinde **yüksek sapma (bias)** aşırı uyumu halinde ise yüksek **değişim (varyans)** oluşur.
  - d) AdaBoost sınıflandırıcıda bir test örneği tüm sınıflandırıcıların verdiği etiketlerin **çoğunluğuna** göre sınıflandırılır.
  - e) ROC eğrisi değişen **eşik değerlerinde** yanlış pozitif oranına karşılık doğru pozitif oranını çizdirir.
  - f) İkili karar ağacı düğümlerinde bir **özellik** ve **değer** çifti yer alır.
  - g) **Entropi** eldeki bir veri kümesinin karmaşıklığını ölçmeye yardımcı olur.
  - h) **Budama** karar ağaçlarında aşırı uyumu önlemenin bir yoludur.
  - i) Tek bir yapay sinir hücresi XOR veya XNOR problemlerini **çözemaz**.
  - j) Çok katmanlı yapay sinir ağlarıyla iki boyutta çıkıntı (bump) benzeri karar sınırlarını öğrenebilmek için **üç veya daha fazla** katmana ihtiyaç vardır.
  - k) K-en yakın komşu algoritmasında bir veri örneğinin komşuları **mesafe** ölçülerek belirlenir.
  - l) **Sarmalayıcı** bir özellik seçim algoritmasıdır **PCA** bir özellik çıkartım algoritmasıdır.

- 2) Şekildeki kategorik veride 12 farklı müşterinin bir restoranda yemek yiyebilmek için şartlarını ifade eden altı farklı özelliğe bağlı olarak bekleyip beklememe kararını verdiği görülmektedir. Buna göre verinin yağış özelliğine göre bölünmesi sonucu elde edilecek bilgi kazancı ne olur? Verinin bölümden önceki ve sonraki entropi değerlerini yorumlayınız.

$$IG(\text{Rain}) = 1 - \left[ \frac{4}{12} I\left(\frac{1}{2}, \frac{1}{2}\right) + \frac{8}{12} I\left(\frac{1}{2}, \frac{1}{2}\right) \right] = 0$$

Veriyi Rain özelliğine göre bölmek hiçbir kazanç sağlamaz.  $E(\text{Restaurant}) = 1$ , dengeli sınıf dağılımından dolayı yüksek karmaşa  $E(\text{Rain}=T) = 1$ ,  $E(\text{Rain}=F) = 1$ , yağmur yağsa da yağmasa da sınıf dağılımı dengelidir. Dolayısıyla entropi yine en yüksek seviyededir.

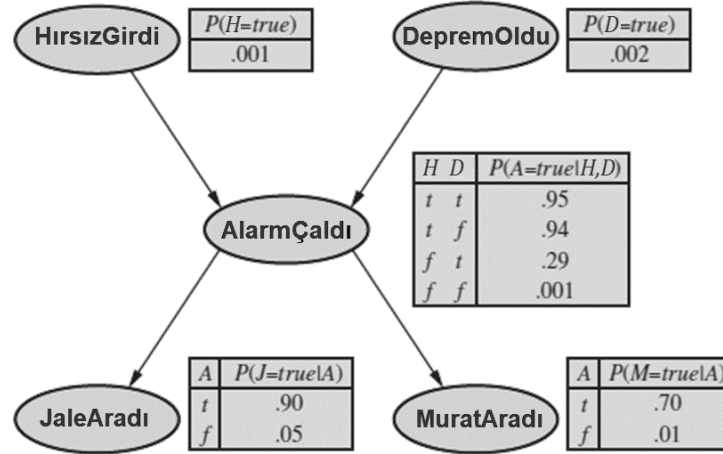
Örnekler	Özellikler						Hedef Bekle?
	Alt	Lobi	Cuma	Aç	Yağış	Rez	
$X_1$	T	F	F	T	F	T	T
$X_2$	T	F	F	T	F	F	F
$X_3$	F	T	F	F	F	F	T
$X_4$	T	F	T	T	F	F	T
$X_5$	T	F	T	F	F	T	F
$X_6$	F	T	F	T	T	T	T
$X_7$	F	T	F	F	T	F	F
$X_8$	F	F	F	T	T	T	T
$X_9$	F	T	T	F	T	F	F
$X_{10}$	T	T	T	T	F	T	F
$X_{11}$	F	F	F	F	F	F	F
$X_{12}$	T	T	T	T	F	F	T

- 3) İkili bir sınıflandırma probleminin çözümü için geliştirilen model test aşamasında sağda verilen karmaşa matrisini vermektedir. Bu matrisin elemanlarını kullanarak aşağıdaki değerleri hesaplayınız.

Tahmin \ Gerçek	Pozitif	Negatif
Pozitif	80	30
Negatif	40	50

- a) **TP = 80**
- b) **FP = 40**
- c) **TN = 50**
- d) **FN = 30**
- e) **Doğruluk (Accuracy) =  $(TP+TN)/(TP+TN+FP+FN) = 130/200 = \%65$**
- f) **Kesinlik (Precision) =  $TP/(TP +FP) = 80/120 = \%67$**
- g) **Hassasiyet (Sensitivity) =  $TP/(TP +FN) = 80/110 = \%73$**

- 4) Aşağıda verilen Bayes ağında beş farklı ikili (boolean) rasgele değişken bağımlılık ilişkileri ve koşullu olasılık tablolarıyla birlikte tanımlanmıştır. Bu Bayes ağına göre mesela ev sahibi isteyken evdeki alarmin çalması direkt olarak eve hırsız girmesi veya deprem olması değişkenlerine; Jale'nin veya Murat'ın ev sahibini araması değişkenleri ise direkt alarmin çalması değişkenine bağlıdır.



- a) Deprem olmaması, hırsız girmesi, alarmin çalması, Jale'nin araması fakat Murat'ın aramaması birleşik olasılığı yani  $P(j \wedge \neg m \wedge a \wedge h \wedge \neg d)$  nedir?

$$P(j \wedge \neg m \wedge a \wedge h \wedge \neg d) = P(j|a) \cdot P(\neg m|a) \cdot P(a|h, \neg d) \cdot P(h)P(\neg d) \\ = 0.9 \times 0.3 \times 0.94 \times 0.001 \times 0.998 = 0.00026$$

- b) Hırsız girmediği ve deprem olduğu bilindiğinde alarmin çalması olasılığı yani  $P(\neg a | \neg h \wedge d)$  nedir?

$$P(\neg a | \neg h, d) = 1 - 0.29 = 0.71$$

- c) Herhangi bir kanıt olmadığında Murat'ın ev sahibini araması olasılığı yani  $P(m)$  nedir?

$$P(m) = P(m|a)P(a) + P(m|\neg a) \cdot P(\neg a) = 0.7 \times 0.0025 + 0.01 \times 0.9975 = 0.0117$$

$$P(a) = P(a|h, d) \cdot P(h)P(d) + P(a|h, \neg d) \cdot P(h)P(\neg d) + P(a|\neg h, d) \cdot P(\neg h)P(d) \\ + P(a|\neg h, \neg d) \cdot P(\neg h)P(\neg d) \\ = 0.95 \times 0.001 \times 0.002 + 0.94 \times 0.001 \times 0.998 + 0.29 \times 0.999 \times 0.002 \\ + 0.001 \times 0.999 \times 0.998 = 0.0025$$

$$P(\neg a) = 1 - P(a) = 0.9975$$

- d) Jale aramadığında alarmin çalması olasılığı yani  $P(a | \neg j)$  nedir? İpucu: Bayes'in kuralını uygulayınız.

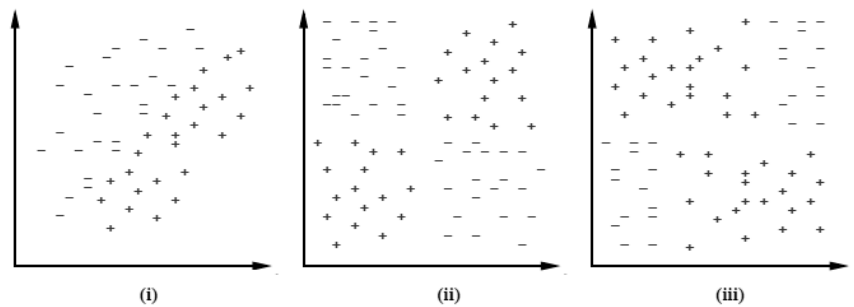
$$P(a | \neg j) = \frac{P(\neg j|a) \times P(a)}{P(\neg j)} = \frac{0.1 \times 0.0025}{0.948} = 0.000264$$

$$P(\neg j) = P(\neg j|a) \times P(a) + P(\neg j|\neg a) \times P(\neg a) = 0.1 \times 0.0025 + 0.95 \times 0.9975 = 0.948$$

- e) Eve hırsız girdiği bilindiğinde alarmin çalması olasılığı yani  $P(a | h)$  nedir?

$$P(a|h) = P(a|h, d) \times P(d) + P(a|h, \neg d) \times P(\neg d) = 0.95 \times 0.002 + 0.94 \times 0.998 = 0.94$$

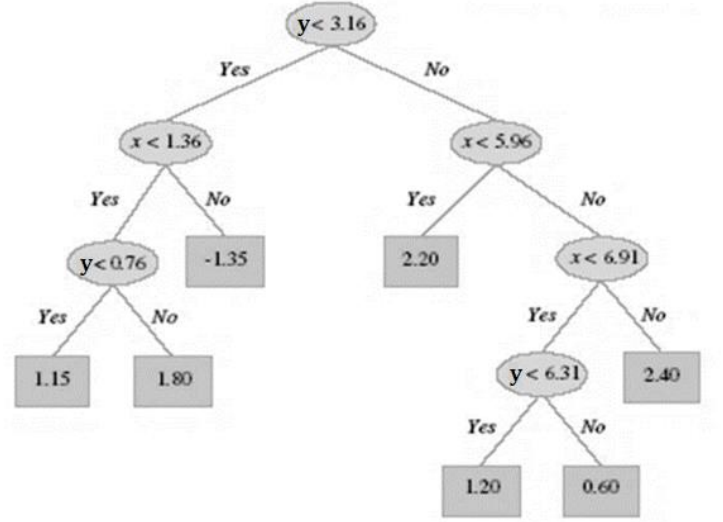
- 5) Şekilde verilen üç farklı veri seti + ve - ile ifade edilmiş iki sınıfa ait iki boyutlu veri örneklerinden oluşmaktadır. Bu üç farklı veri setinin her biri için lojistik regresyon, karar ağacı veya yapay sinir ağı ile sınıflandırma tekniklerinden hangisi uygun olur? Sebeplerini açıklayınız.



- a) Veri kümesi (i) için **lojistik regresyon (Lineer bir sınırla ayrılabilir iki sınıf vardır.)**
- b) Veri kümesi (ii) için **karar ağaçları (Veri if-else benzeri görece basit dallanmalarla ayrılabilir.)**
- c) Veri kümesi (iii) için **yapay sinir ağı (Daha karmaşık bir veri dağılımı olup sınıf örnekleri lineer olmayan kompleks sınırlarla ayrılabilir.)**

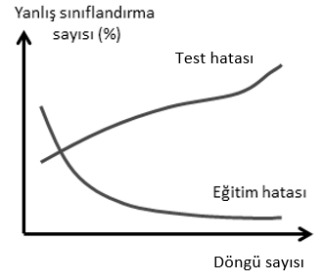
- 6) Şekilde verilen regresyon ağacına göre aşağıdaki veriler test edildiğinde tahmin edilen çıktı değeri ne olacağını ifade ediniz.

$x$	$y$	değer
4,2	5,7	<b>2,20</b>
0,8	2,3	<b>1,80</b>
6,4	7,1	<b>0,60</b>

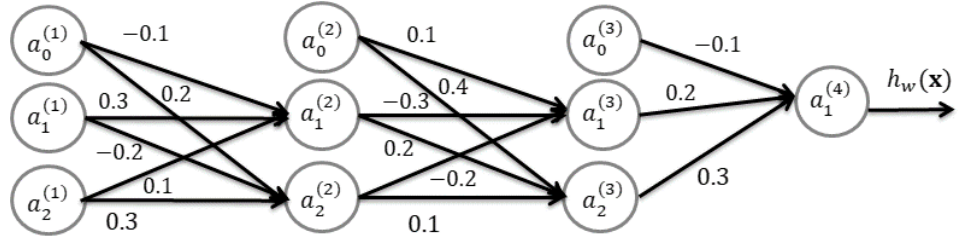


- 7) Bir sınıflandırma modelinin öğrenilmesi süresince döngü sayısına bağlı olarak test ve eğitim veri kümelerindeki örnekler için yanlış sınıflandırma yüzdesi sağdaki grafikte çizdirilmiştir. Grafikte belirtilen en büyük döngü sayısı tamamlandığında öğrenilmiş olan modelin temel problemi nedir? Problemin çözümü için dört farklı öneride bulununuz.

**Aşırı uyum problemi mevcuttur, mevcut özellikler içerisinde özellik seçimi yapılabilir, daha fazla örnek ile eğitim gerçekleştirilebilir, PCA veya LDA benzeri algoritmalarla özellikler farklı bir uzaya indirgenebilir, uygun bir düzenleme parametresi kullanılabilir vb.**



- 8) Sapma terimi olarak 1 değerini ve aktivasyon adımı için sigmoid fonksiyonunu kullanan şekildeki çok katmanlı yapay sinir ağı için,



- a)  $[a_1^{(3)} \ a_2^{(3)}] = [0.1, 0.2]$  ise

$h_w(x)$  nedir?

$$h_w(x) = g(1 \cdot -0.1 + 0.1 \cdot 0.2 + 0.2 \cdot 0.3) = g(-0.02) = 0.495$$

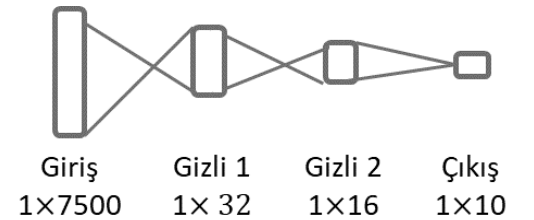
- b)  $a_1^{(2)} = 0.2$  ve  $[\delta_1^{(3)} \ \delta_2^{(3)}] = [0.1, 0.4]$  ise  $\delta_1^{(2)}$  nedir?

$$\delta_1^{(2)} = (w_{11}^{(2)} \cdot \delta_1^{(3)} + w_{21}^{(2)} \cdot \delta_2^{(3)}) \cdot a_1^{(2)} \cdot (1 - a_1^{(2)}) = (0.1 \cdot -0.3 + 0.4 \cdot 0.2) \cdot 0.2 \cdot 0.8 = 0.05 \cdot 0.2 \cdot 0.8 = 0.008$$

- c)  $\alpha$  öğrenme oranı 0.1 ve  $a_1^{(1)} = 0.8$  ise hesaplanan  $\delta_1^{(2)}$  değerine göre  $w_{11}^{(1)}$  nasıl güncellenir?

$$w_{11}^{(1)} = w_{11}^{(1)} + \alpha(a_1^{(1)} \cdot \delta_1^{(2)}) = 0.3 + 0.1(0.8 \cdot 0.008) = 0.30064$$

- 9) Girdisi  $50 \times 50 \times 3$  boyutlu renkli görüntülerden ve çıktısı 10 farklı görüntü türüne ait bir etiketten oluşan bir sınıflandırma problemi yapay sinir ağları ile çözülecektir. Ağın gizli katmanlarında sırasıyla 32 ve 16 nöron bulunmaktadır. Nöronlarda sapma terimi kullanıldığına göre,



- a) Yapay sinir ağının mimarisini basitçe çiziniz.

- b) Yapay sinir ağında öğrenilmesi gereken toplam ağırlık (parametre) sayısını hesaplayınız.

$$7500 \cdot 32 + 32 + 32 \cdot 16 + 16 + 16 \cdot 10 + 10 = 240032 + 528 + 170 = 240730$$

- c) İki gizli katmandan önce nöron bağlantıları %50 oranında rasgele kopartıldığında (dropout) öğrenilmesi gereken toplam ağırlık (parametre) sayısını hesaplayınız.

$$7500 \cdot 16 + 32 + 32 \cdot 8 + 16 + 16 \cdot 10 + 10 = 120032 + 272 + 170 = 120474$$