

Örüntü Tanıma

Emir Öztürk

Oğuz Kırat

Örüntü Tanımada Lineer Cebir

- Lineer cebir makine öğrenmesi için gerekli
 - Verinin işlenmesi ve değiştirilmesi için
- Genellikle veri vektörler veya matrisler ile ifade edilir.
- Vektör veya matris hesaplarını hızlı yapabilecek araçlar ve donanım
 - Performans kazancı amaçlanır.

Örüntü Tanımada Lineer Cebir

- Veri dönüşümü, boyut küçültme (PCA) gibi karmaşık işlemlerin basitleştirilmesi için kullanılır.
- Matris çarpımı, özdeğer, özvektör ve lineer işlemler ağırlıklıdır.
- Örüntü tanıma ve makine öğrenmesinde skalarlar, vektörler ve matrisler kullanılır.
- Bunların tamamı Tensor kavramı altında incelenebilmektedir.



Tensor

- Çok boyutlu verinin simgelenebildiği geometrik nesnelerdir.
- Çok boyutlu veri kümeleri için kullanılır.
- Boyut sayısı tensörün derecesidir.
- Skalerin derecesi 0
- Vektörün derecesi 1
- Matrisin derecesi 2
- Tensörlerin derecesi 3 ve üstü de olabilir
 - RGB resimler
 - Video çerçeveleri (frame)

Skalar

- Bir sayı değerinden oluşur.
- Hangi uzayda olduklarının belirtilmesi gerekir.
- \mathbb{R} Reel sayılar
- \mathbb{N} Pozitif tamsayılar
- \mathbb{Z} Tamsayılar
- \mathbb{Q} Rasyonel sayılar

Vektörler

- Geometrik bir objedir.
- Vektör uzaylarının üyeleridir.
- Vektör uzayı
 - Bir boyutta bulunabilecek tüm vektör olasılıklarını içeren koleksiyon
- İki özelliğe sahiptir.
 - Sayısal büyüklük
 - Yön

Vektörler

- Farklı gösterimlere sahiptir
- $V = (3, -2, 5)$
- $V = [3 \ -2 \ 5]$
- $V = \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix}$

Vektörler

- Çok boyutlu olması sayesinde birden fazla özellik saklanabilir.
- Örüntü tanımada özellik vektörleri kullanılır.
- Vektör içerisindeki her değer bir özelliği ifade eder.

Matrisler

- Matrisler satır ve sütunlara sahip sayılar dizisidir.
- Lineer eşitlik sistemleri ifade edilebilir.
- Örüntü verisi için veri dönüştürme aşamasında kullanılır.
- $M = \begin{bmatrix} 11 & 12 \\ 21 & 22 \end{bmatrix}$
- İki indis ile ifade edilirler
- m satır ve n sütun olmak üzere mxn şeklinde verilirler.

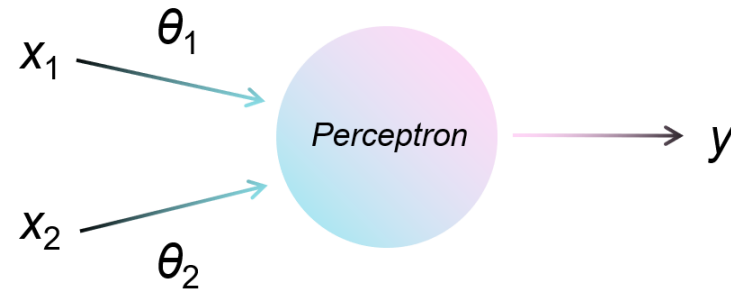
Matrisler

- Bir vektör aynı zamanda $1 \times n$ ya da $m \times 1$ bir matristir.
- Matrisler üzerinde toplama çıkarma ve çarpım işlemleri gerçekleştirilebilir.
- Bu işlemler örüntü tanıma işlemlerinde sıkça kullanılan işlemlerdir.
 - Özellik vektörlerinin dönüşümü
 - Model parametrelerinin hesaplanması
 - Feedforward ve backpropagation

Matrisler

- Derin öğrenme uygulamalarında ağırlıklar matris formunda saklanır.
- Girdiler vektör olarak saklanır.
- Özellikle ağırlık hesaplarının ifade edilebilmesi için nokta çarpım (dot product) kullanılır.

$$\mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^n a_i b_i = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$



Lineer dönüşümler

- Matrisler rotasyon ve transformasyon işlemlerini kodlayabilirler.
 - Bir modeli 3 boyutta dönüştürmek için önceden tanımlanmış bir rotasyon matrisi ile çarpılabilir
- Linear Translation (Çevirme)
 - Bir verinin normalize edilmesi ya da ortalanması (centering data) için kullanılır.
- Scaling (Ölçeklendirme)
 - Vektörleri genişletme veya sıkıştırma işlemleri
- Rotation (Döndürme)
 - Verileri bir eksen üzerinde çevirme işlemi

Özdeğer ve Özvektör

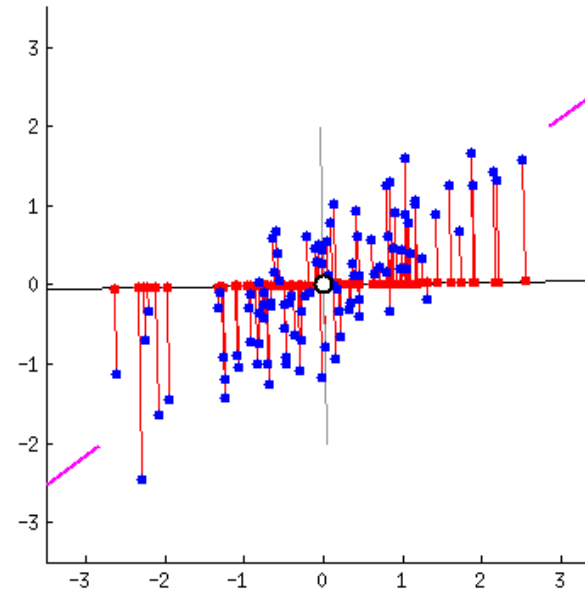
- Bir A kare matrisi için, sıfır olmayan v vektörü öyle ki:
- $A \cdot v = \lambda \cdot v$
- Burada $v \neq 0$ olmalıdır.
- Bu vektöre özvektör adı verilir.
- λ skalerine özdeğer adı verilir.
- Matrisin özvektör üzerindeki etkisini gösterir.
 - Vektörü yalnızca ölçeklendirir.
 - Yönünü değiştirmez.

Özdeğer ve Özvektör

- Verilen
- $A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$
- $v = [1, 0]$ için:
 - $A \cdot [1, 0] = [2, 0] = 2 \cdot [1, 0]$
 - $\lambda = 2, [1, 0]$ ise bu özdeğere ait özvektördür.
- $v = [0, 1]$ için:
 - $A \cdot [0, 1] = [0, 3] = 3 \cdot [0, 1]$
 - $\lambda = 3, [0, 1]$ ise bu özdeğere ait özvektördür.

Özdeğer ve Özvektör

- Karmaşık verilerde boyut düşürme sağlar ve böylece performans artışı
- PCA
 - Özdeğer ve özvektörler elde edildiğinde sadece n adet özvektör seçilerek boyut düşürülebilir.



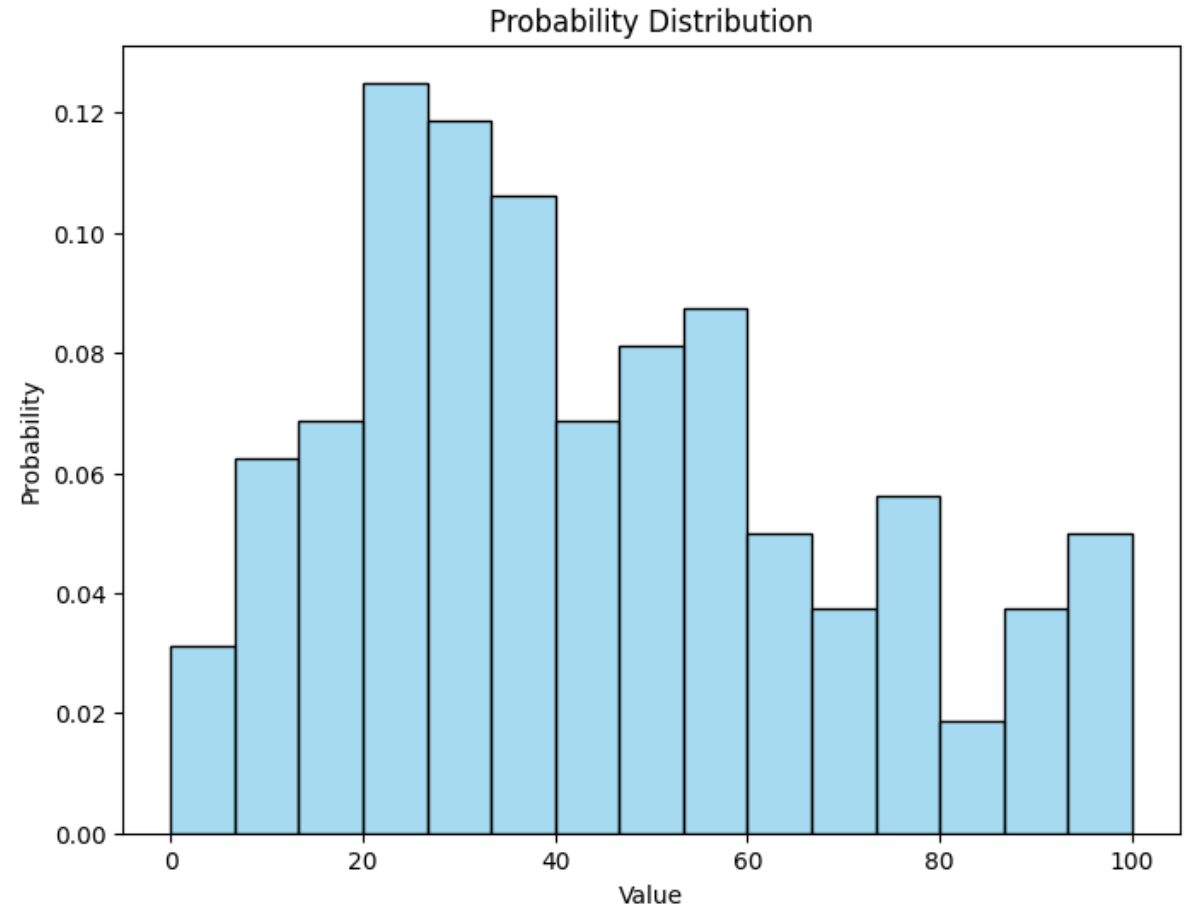
Olasılık Dağılımları

- Olasılık dağılımı, bir değişkenin alabileceği farklı değerlerin olasılıklarının dağılımıdır.
- Bir yazı tura uygulaması için
 - Yazı 0.5
 - Tura 0.5
- Frekans dağılımı
 - Bir verisetinde her veriden ne kadar geçtiği
 - Olasılık dağılımını belirler
- Olasılık dağılımını gerçek dünya örneğinde imkansız
 - Sonsuz büyüklükte örnek alma ihtimali yok

Ayrık Olasılık Dağılımları

- Tüm değerlerin olası olduğu durumdur.
- Dağılım içerisinde olasılığı 0 olan eleman bulunmaz.
- Örneğin zar atmak
- Ayrık olasılık dağılımının toplamı 1 eder.
- Binomial
- Ayrık uniform
- Poisson

Olasılık Dağılımları



Sürekli Olasılık Dağılımları

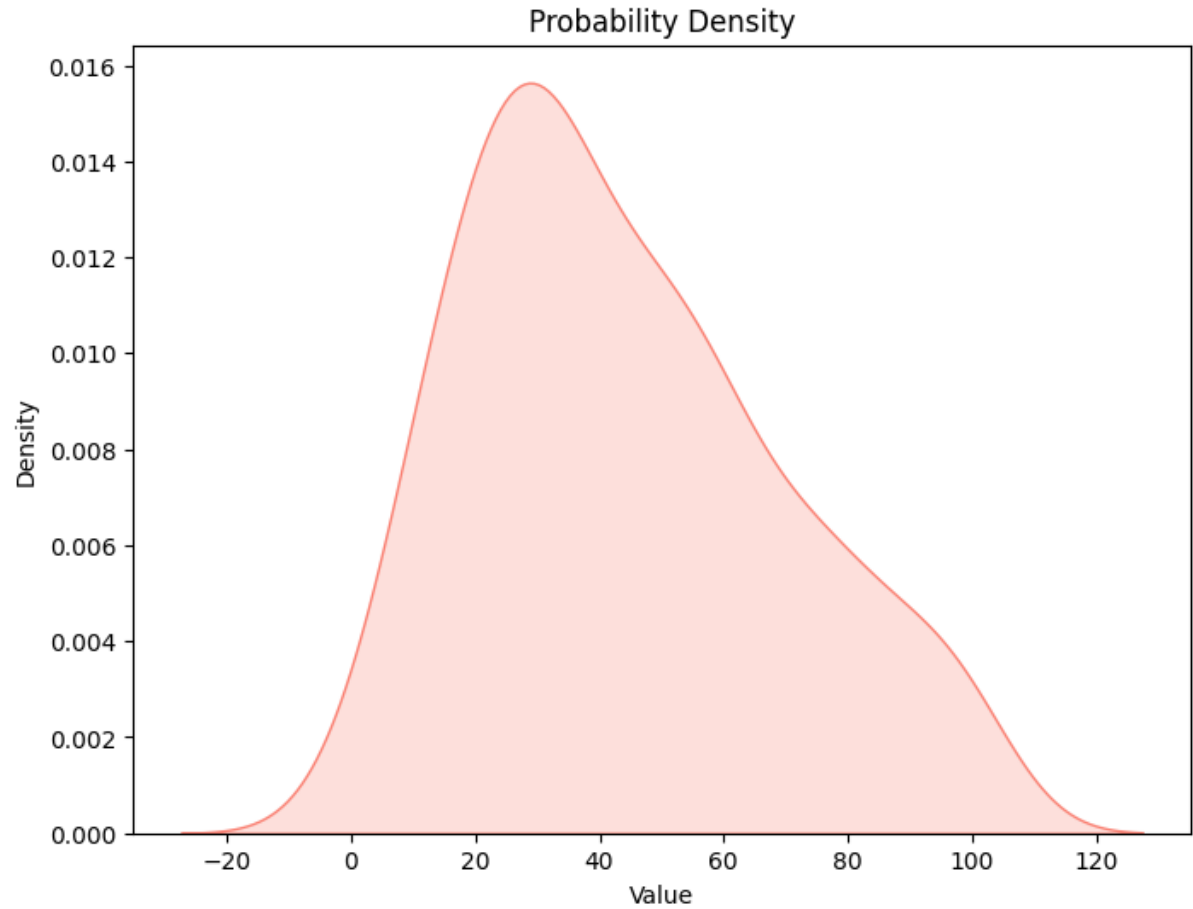
- Sürekli değer en alt ve en üst değerler arasındaki her değeri içerir.
- Değişkenin aralığındaki tüm değerleri almak mümkündür.
- Sürekli bir olasılık dağılımında herhangi bir sayının olasılık değerinin olma ihtimali çok küçüktür.
- Bu değer 0 sayılabilir.
- Bunun yerine bir değer belirlenen bir aralıkta olma ihtimali alınabilir.
- Olasılık yoğunluk fonksiyonu

Olasılık Yoğunluk Fonksiyonu

- Olasılık yoğunluk fonksiyonu
- μ beklenen değer
- σ standart sapma

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Olasılık Yoğunluğu



Örüntü Tanımada Olasılık Dağılımları

- Veri üretimi için Gaussian Mixture Model
- Variational Autoencoder'lar için
 - Latent değişkenler
- Sınıflandırma için
 - Bayes sınıflandırıcılar
 - Lojistik regresyon ve softmax
- Kayıp fonksiyonları
 - Log likelihood

Bayes Teoremi

- Koşullu olasılık
 - Başka bir olayın olma durumunda olayın olma olasılığı
- $P(A|B) = \frac{P(A \cap B)}{P(B)}$
- $P(A \cap B)$ ortak olasılık
- Havanın soğuk olma olasılığı?
- Kış olması durumunda havanın soğuk olma olasılığı?

Bayes Teoremi

- Thomas Bayes
- Koşullu olasılıktan türetilmiş
- $P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$
- $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$

Bayes Teoremi

- D Hasta olma durumu
- \bar{D} Hasta olmama durumu
- T Testin pozitif olma durumu
- Örneğin;
- $P(D) = 0.1$
- $P(\bar{D}) = 0.9$
- $P(T|D) = 0.8$
- $P(T|\bar{D}) = 0.1$
- Testin pozitif olması durumunda kişinin hasta olma olasılığı?

Bayes Teoremi

- $P(D|T) = 0.1$
- $P(\bar{D}) = 0.9$
- $P(T|D) = 0.8$
- $P(T|\bar{D}) = 0.1$
- $P(T) = P(T|D)P(D) + P(T|\bar{D})P(\bar{D})$
- $P(T) = 0.8 * 0.1 + 0.1 * 0.9 = 0.17$
- $P(D|T) = \frac{P(T|D)P(D)}{P(T)} = \frac{0.8*0.1}{0.17} = 0.47$

Örüntü Tanımada Bayes Teoremi

- Bir girdi ile bir çıktının arasındaki ilişkinin tespiti için kullanılır.
- $$P(Sınıf|Veri) = \frac{P(Veri|Sınıf)P(Sınıf)}{P(Veri)}$$
- Sınıflandırmanın temeli olarak kabul edilebilir.
- Naïve Bayes
 - Bayes
 - Hesaplamanın kolaylaştırılması gerekir.
 - Olasılıkların birbirini takibi yerine olma olasılıklarını saklar

Örüntü Tanımada Bayes Teoremi

- Örneğin;
- A B C kelimesinde bayes için A'dan sonra gelen B kelimesinden sonra gelen C kelimesinin olasılığının hesaplanması gerekir.
- Naive bayes için ise A, B ve V kelimelerinin gelme olasılıklarının kullanılması yeterlidir.
- Sıra önemli değildir.

Bayes Teoremi

Koşullu Olasılık	Bayes Teoremi	Naive Bayes Algoritması
Bir olayın, başka bir olayın gerçekleştiği durumda meydana gelme olasılığıdır.	Koşullu olasılık kavramını tersine çevirmek için kullanılan formüldür	Bayes teoremine dayalıdır, fakat tüm özelliklerin birbirinden bağımsız olduğunu varsayar.
Kullanım Alanı: Genel olasılık hesaplamaları Risk analizi	Kullanım Alanı: Ön bilgi (prior) güncellemesi Tıbbi teşhis ve karar destek sistemleri Yeni veriye göre olasılıkların revizesi	Kullanım Alanı: Spam filtreleme, metin sınıflandırması, duygu analizi gibi makine öğrenmesi ve veri madenciliği uygulamaları

Türev (Optimizasyon)

- Çıktı ve istenen değerin arasındaki hatanın minimize edilmesi istenir.
- Diferansiyel hesaplar bu işlem için yapılır.
- Parametrelere bağlı olarak modelin performansının türevlerini incelemek, parametrelerin düzeltilmesi için kullanılır.
- Amaç daha iyi sonuçlar elde etmektir.