



Makine Öğrenmesi

1. hafta

- Temel Terimler
- Danışmanlı – Danışmansız Öğrenme
- Veri Hazırlama
- Çapraz Geçerlik
- Aşırı Eğitim

1



Makine Öğrenmesi Nedir?

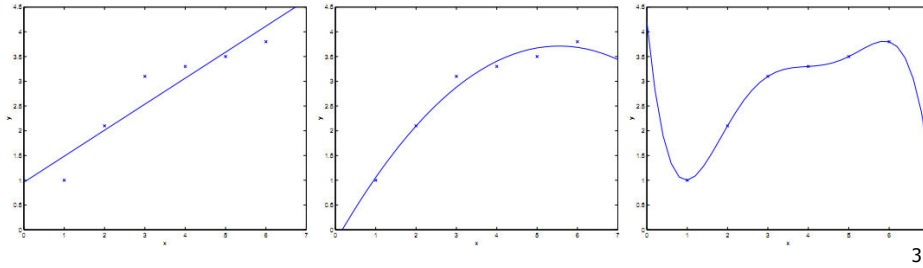
Makine Öğrenmesi, verilen bir problemi probleme ait ortamdan edinilen veriye göre modelleyen bilgisayar algoritmalarının genel adıdır. Yoğun çalışılan bir konu olduğu için önerilmiş birçok yaklaşım ve algoritma mevcuttur. Bu yaklaşımların bir kısmı tahmin (prediction) ve kestirim (estimation) bir kısmı da sınıflandırma (classification) yapabilme yeteneğine sahiptir.

2



Makine Öğrenmesi Yöntemleri

Önerilmiş birçok makine öğrenmesi yöntemi mevcuttur. Bunlar probleme yaklaşımlara göre farklılık gösterebilir ve bu yüzden farklı problemlerde farklı başarılarla sahip olabilirler.



3



Makine Öğrenmesi Terimleri

- Tahmin (prediction): Veriden öğrenen modellerde sistem çıkışının nicel olması durumunda kullanılan yöntemlerin ürettiği değerlerdir.
- Sınıflandırma (classification): Giriş verisine ait çıkışların nitel olduğu durumlarda kullanılan yöntemlerin her veri örneğinin hangi sınıfa ait olduğunu belirlemesidir.

4



Tahmin ve Kestirim

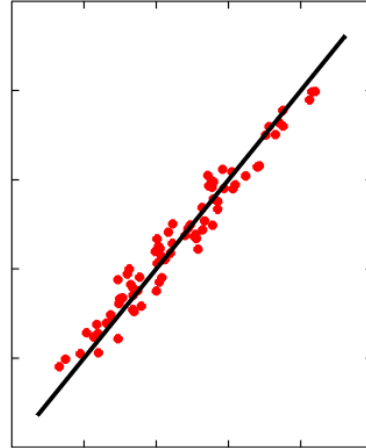
İstatistikte rasgele bir değişkenin bilinmeyen bir değerinin belirlenmesi için tahmin (prediction), bilinmeyen bir sabitin belirlenmesi içinse kestirim (estimation) kavramından bahsedilir. Çok yakın anlamları dolayısıyla bu iki terim literatürde çoğunlukla karıştırılarak aynı şeyi ifade etmede kullanılır.

5



Tahmin ve Kestirim

Örneğin sağda veri noktaları için bir doğrusal eğri fonksiyonu tespit edilmiştir. Herhangi bir x değerine karşılık gelen y değerini belirlemekte "tahmin", tahmini yapmamıza yardımcı olan eğrinin (eğim ve sapma değeri gibi) parametrelerinin belirlenmesinde ise "kestirim" terimleri tercih edilmelidir.



6

Bir Kestirim Örneği

Bir probleme ilişkin olarak sağdaki veri kümesi veriliyor. Amaç bir x girdisine karşı bir y değerini üretmektir. Buna göre veri üzerinde doğru ilişkiyi bulan öyle bir $f : X \rightarrow Y$ fonksiyonu tanımlanmalıdır ki en küçük hatayı üretsinsin.

Bu basit problemde $y=3*x$ çözümü kolayca görünmektedir.

X	Y
3	9
5	15
8	24
10	30
19	57
24	72
27	81
31	93
38	114
43	129

7

Bir Tahmin Örneği

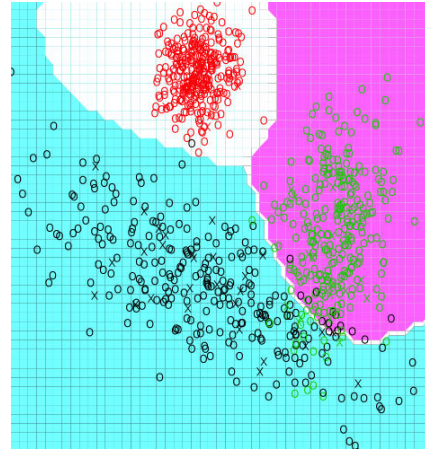
Bulunan $y=3*x$ çözümüne ilişkin olarak eğer $x=50$ değeri için y değeri istenirse kestirilen denklem yardımı ile $y=150$ tahmin değerini kolayca hesaplamak mümkün olur.

X	Y
3	9
5	15
8	24
10	30
19	57
24	72
27	81
31	93
38	114
43	129

8

Sınıflandırma

Amaç probleme ait tüm uzayın belirli sayıda sınıfa bölünmesidir. Sağdaki resimde her renk bir sınıfı göstermektedir. Sınıflandırma teknikleri sayesinde hiçbir veri örneğinin olmadığı bölgeler de renklendirilebilir.



9

Bir Sınıflandırma Örneği

Sağdaki veri kümesinde amaç X girdisine karşı Y etiket (label) değerini üretmektir. Öyle bir $f : X \rightarrow Y$ sınıflandırma fonksiyonu tanımlanmalıdır ki en az hata ile sınıflandırma yapılabilir.

$$Y = \begin{cases} 0, & X < 20 \\ 1, & X \geq 20 \end{cases}$$

X	Y
2	0
5	0
9	0
13	0
19	0
20	1
27	1
33	1
39	1
47	1

10



Makine Öğrenmesi Terimleri

Makine Öğrenmesi yöntemleri verinin yapısına göre ikiye ayrılır.

- Danışmanlı (supervised) Öğrenme: Veri, etkiye tepki prensibiyle çalışan sistemlerden alınır ve giriş-çıkış düzeninde organize edilir.
- Danışmansız (unsupervised) Öğrenme: Sınıf bilgisi olmayan veya verilmeyen veri içerisindeki grupları keşfetmeyi hedefler.

11



Takviyeli Öğrenme

Bazen öğretici, sisteme beklenen sonucu tam söyleyemez. Ama sistemin ürettiği sonuç için “doğru/yanlış” şeklinde fikir belirtir. Bu öğrenme şekli de takviyeli (reinforcement) öğrenme adıyla anılır.

Boltzmann makinesi, LVQ ve genetik algoritma örnek olarak sayılabilir.

12



Makine Öğrenmesi Terimleri

Sınıflandırma ve tahmin yöntemlerine danışmanlı öğrenme yöntemleri de denilebilir.

Kümeleme (clustering) yöntemleri ise danışmansız öğrenme yöntemleri olarak anılırlar.

Peki Kümeleme nedir?

13



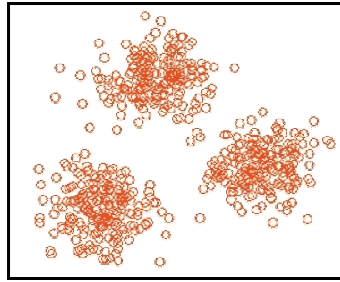
Kümeleme

Danışmansız öğrenmeyi baz alan Kümelemede amaç veri kümesi içerisindeki veri örneklerini sadece özellik (feature) vektörlerine göre gruplamaktır. Bunun için örneklerin birbirine benzerliği gözetilir. Kümeleme literatüründe benzerlik terimi uzaklık terimiyle zıt anlamda kullanılmaktadır.

14

Kümeleme

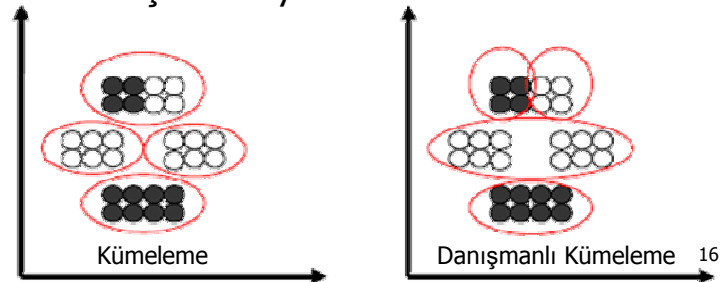
Birbirine benzer olan örnekler aynı kümeye ve birbirinden uzak olan örnekler farklı kümelere yerleştirilmeye çalışılır. Küme sayısı genellikle dışarıdan verilir.



15

Danışmanlı Kümeleme

Sınıf bilgisinden bağımsız olarak veri içerisinde bilgi keşfi sınıflandırma için destekleyici bir etkiye sahip olabilir. Sınıflandırma öncesi verinin kümelenmesi (supervised clustering), genellikle sınıflandırma başarısını yükseltmektedir.





Notasyonlar

Çalışmalarda

- İstenen sınıflar için D ,
 - Tahmin çıkışları için Y ,
 - Giriş verisi için X ,
 - Girişin her bir özelliği için X_j ,
 - Veri içerisindeki her bir örnek için x_i ,
 - Giriş veri kümesinin tamamı için **X**
- notasyonları tercih edilmektedir.

17



Öğrenme Zamanlaması

- Çevrimiçi (online) öğrenme: Gerçek zamanlı olarak çalışan sistemlerde öğrenme aşamasının çalışma anında da sürdürülmesi gerekiyorsa kullanılır.
- Çevrimdışı (offline) öğrenme: Uygulama sistemi henüz çalışmaz iken eğitim gerçekleştirilir, eğitilmiş yöntem sisteme yüklenir ve sistem çalıştırılır.

18



Öğrenme Kuralları

Veriden eğitim için önerilen her algoritma farklı bir öğrenme kuralı olsa da genel eğilim algoritmaları ortak özelliklerine göre gruplamak yönündedir. Öğrenme algoritmaları dört grupta toplanabilir:

- Hebb
- Delta
- Hopfield
- Kohonen

19



Hebb Öğrenme Kuralı

1949'da geliştirilen ilk öğrenme kuralıdır. "Bir hücre, bağlı olduğu diğer hücreleri etkiler" prensibine dayalıdır. Bu öğrenme kuralı geliştirilerek farklı öğrenme kuralları geliştirilmiştir.

20



Delta Öğrenme Kuralı

Beklenen sonuç ve hesaplanan sonuç arasındaki karesel fark, sistemin hatasıdır. Bu hatayı indirmek için hücreler arasındaki bağlantılar sürekli değiştirilir. Çok katmanlı algılayıcı ağlar bu kurala göre eğitilirler.

21



Hopfield Öğrenme Kuralı

Beklenen sonuç ile hesaplanan sonuç aynı ise hücreler arası bağlar belirli bir oranda güçlendirilir, aksi durumda bağlar zayıflatılır. Recurrent ve Elman ağları bu kural ile eğitilirler.

22

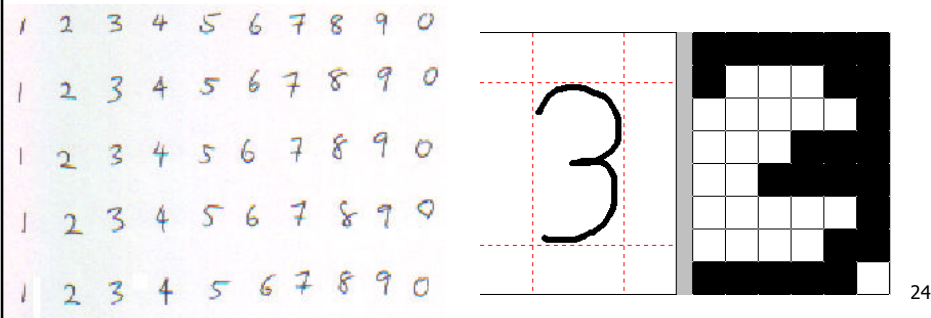
Kohonen Öğrenme Kuralı

Bu danışmansız modelde hücreler yarış halindedir. En büyük sonucu üreten hücre yarış kazanır. Kazanan hücrenin ve komşularının bağları güçlendirilir. ART (Adaptive Resonance Theory) ve Kohonen tarafından geliştirilen SOM (Self Organizing Map) ağları örnek olarak verilebilir.

23

Veri Hazırlama

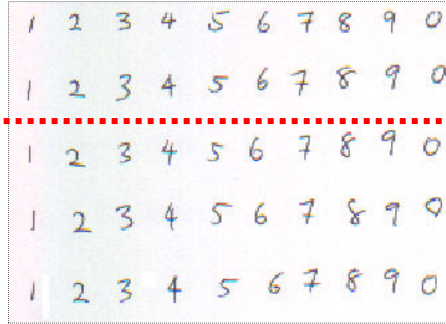
Sınıflandırma veya tahmin için problem uzayını tam olarak temsil edebilen veri kümesi hazırlamak gerekmektedir.



24

Veri Hazırlama

Çözüm için hazırlanan veri, hem eğitim hem de başarı ölçümünde kullanıldığı için genellikle ikiye ayrılır.



1	2	3	4	5	6	7	8	9	0
1	2	3	4	5	6	7	8	9	0
1	2	3	4	5	6	7	8	9	0
1	2	3	4	5	6	7	8	9	0
1	2	3	4	5	6	7	8	9	0

25

Sınıflandırma Geçerliği

Veriye dayalı olarak eğitim yapılmasının temel amacı eğitilen sistemin benzer konuda hiç bilinmeyen bir örneğe mantıklı bir cevap üretebilmesidir. Eldeki sınırlı veri kullanılarak sistemin hem eğitilmesi hem de başarısının tarafsız bir şekilde ölçülebilmesi gerekmektedir. Bunun için çapraz geçerlik (cross validation) adıyla anılan yöntemler kullanılmalıdır.

26

Çapraz Geçerlik (Cross Validation)

Bu temel prensibe dayanarak önerilmiş birkaç geçerlik yöntemi vardır. Ama hepsinde temel mantık aynıdır. Sistemin başarısını ölçebilmek için mevcut veri kümesi ikiye bölünür. Birisi eğitim için (train set) diğeri de sistemin hiç görmediği olası örnekleri temsilen (test set) kullanılır. Sistem, seçilen eğitim algoritması ile eğitim kümesini öğrenir. Eğitilen sistemin başarısı daha sonra test kümesi üzerinde hesaplanır.

27

Çapraz Geçerlik Yöntemleri

Üç tip çapraz geçerlik yöntemi önerilmiştir:

- Rasgele örnekleme
- K parçalı
- Birini hariç tut

28

Rasgele Örneklemeli Geçerlik

Rasgele örneklemede kümeler rasgele seçilir. Optimum başarı için birkaç kez tekrarlanır. Çapraz geçerlik içerisinde en yüksek başarıyı sağlayan yöntemdir.

DATASET	TRAIN SET	TEST SET

1

29

K Parçalı ve Birini Hariç Tut

K parçalıda veri, K adet kümeye ayrılır. Birisi test kümesi için ve diğer K-1 küme birleştirilip eğitim kümesi için seçilir. Bu işlem kümeler değiştirilerek K kez tekrarlanır. Genel başarı için K adet başarı değerinin ortalaması alınır.

DATASET	TRAIN SET	TEST SET

1

Birini hariç tut yöntemi ise N adet örnek içeren veri için N parçalı çapraz geçerliktir.

30



Aşırı Eğitim (Overfitting)

Veri ile iterasyonel eğitim yapan modellerin hepsinde öğrenme sürecinin zamanı gelince durdurulması gerekir. Eğitim durdurulmazsa öğrenilmesi gereken veri içerisindeki tüm örnekler sistem tarafından ezberlenir ve bilinmeyen örnekleri tahmin yeteneği azalır. Aşırı eğitim (overfitting) denilen bu istenmeyen durum makine öğrenmesinin temel amacı olan genelleştirme kavramına zarar verir.

31



Eğitim Süresi

Eğitimin ne zaman durdurulması gerektiği kullanılan algoritmaya bağlıdır. Erken durdurma yapılırsa sistem veriyi tam öğrenemediği için başarı düşecektir. Aşırı eğitimde de modelin genelleştirme yeteneği zarar göreceği için benzer şekilde başarı kötü yönde etkilenecektir.

32