

Yapay Zeka Öğrenme Modelleri

Giriş



Yapay Zeka Öğrenme Modelleri

Giriş

Ders için kaynaklar:

Stephen Marsland, Machine Learning: An Algorithmic Perspective, Crc Press, 2009.

Sunila Gollapudi, Practical Machine Learning, Packt Publishing Ltd., Birmingham – Mumbai, January 2016.

Ashish Kumar, Learning Predictive Analytics with Python, Packt Publishing Ltd., Birmingham – Mumbai, February 2016.

Peter Harrington, Machine Learning in Action, Manning Publications Co., 2012.

Prof. Dr. Ercan Öztemel, **Yapay Sinir Ağları**, Papatya Yayıncılık

Dr. Yalçın Özkan, **Veri Madenciliği Yöntemleri**, Papatya Yayıncılık, 2016

Sınav

Arasınav : %40

Ödev+Proje: %20

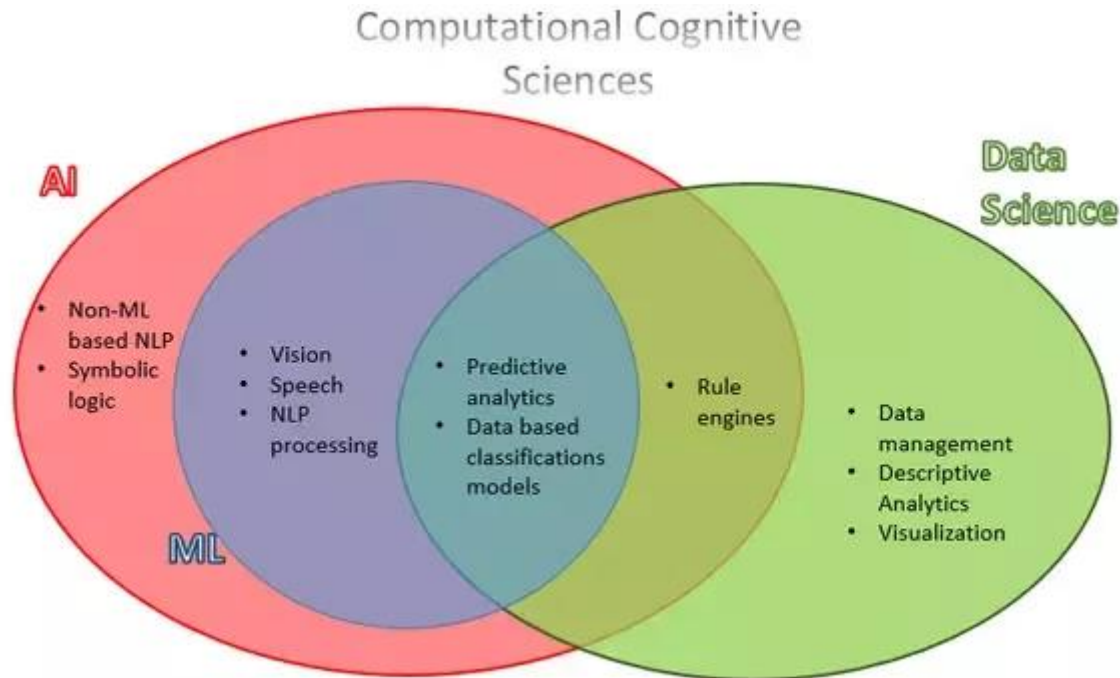
Final Sınavı %40

Yapay Zeka (Artificial Intelligence-AI), Makine Öğrenmesi (Machine Learning-ML) ve Veri Madenciliği (Data Mining)

- İnsan zekâsının bilgisayarlara aktarılabilmesi düşüncesi yapay zekâ çalışmalarının temelini oluşturmaktadır.
- Bu doğrultuda, öğrenebilen makinelerin geliştirilmesine etki eden unsurlardan birinin, insandaki öğrenmenin bilgisayarlara aktarılabilmesi yani bir yapay öğrenme elde edebilme düşüncesi olduğu söylenebilir.

Yapay Zeka (Artificial Intelligence-AI), Makine Öğrenmesi (Machine Learning-ML) ve Veri Madenciliği (Data Mining)

- Makine öğrenmesi, bir şemsiye terim olan Yapay zekanın bir alt kümesidir.
- Veri madenciliği toplanan veriyi işlemekle yükümlü olurken makine öğrenmesi bu veriyi öğrenerek yeni sonuçlar geliştirir.



Yapay Zeka (Artificial Intelligence-AI), Makine Öğrenmesi (Machine Learning-ML) ve Veri Madenciliği (Data Mining)

- Veri madenciliği, daha önce anlaşılamayan ve bilinmeyen büyük veri tabanlarından bilgi toplama süreci ve daha sonra ilgili kararları almak için bu bilgileri kullanma sürecidir.
- Veri madenciliği, verileri anlama konusunda daha fazla şey ifade eder.
- Makine Öğrenmesi ise öngörude bulunmak için verileri kullanmakla ilgilidir.

Yapay Zeka (Artificial Intelligence-AI), Makine Öğrenmesi (Machine Learning-ML) ve Veri Madenciliği (Data Mining)

- Tüm makine öğrenmeleri Yapay zeka olarak sayılır, fakat tüm yapay zekalar makine öğrenmesi olarak sayılmaz.
- Örneğin, sembolik mantık, uzman sistemler ve bilgi grafikleri hepsi yapay zeka olarak tanımlanabilir fakat hiçbirisi makine öğrenmesi değildir.

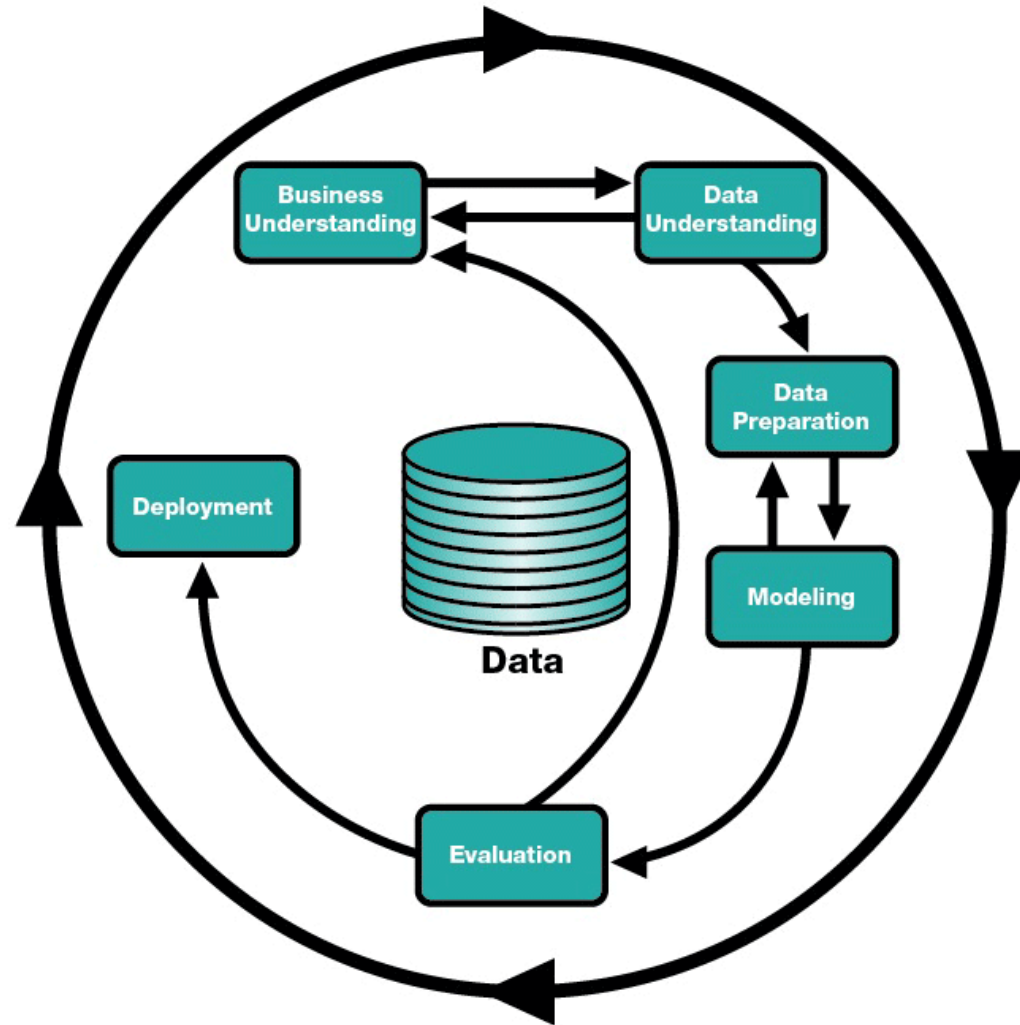
Makine Öğrenmesi

- Yapay Zekanın bir alt dalı olarak makine öğrenmesi, bir makinenin (genellikle bir bilgisayarın) bir işi yapmayı öğrenmesini hedeflemektedir.
- Makine öğrenmesinde öğrenme işi, kullanılan veri kümeleri ve algoritmalar ile sağlanmaktadır.
- Kısaca makine öğrenmesi için, **veriden “öğrenmeyi” sağlayan algoritmalar veya teknikler** denilebilir.

Makine Öğrenmesi Süreci

- Veri madenciliğiyle yakın ilişkisi dikkate alınırsa makine öğrenmesinde, veri madenciliğinde izlenen sürece benzer bir yolun izlenmesi olanaklıdır.
- CRISP (CRoss-Industry Standard Process for Data Mining) modeli,
 - *Problemin Tanımlanması*
 - *Veriyi Anlama*
 - *Veri Hazırlama*
 - *Modelin Kurulması*
 - *Modeli Değerlendirme*
 - *Modelin Uygulamaya Geçirilmesi*
- olmak üzere altı aşamadan oluşmaktadır.

Makine Öğrenmesi Süreci



CRISP (CRoss-Industry Standard Process for Data Mining) modeli

1. Problemin Tanımlanması

- Üzerinde çalışılacak olan öğrenme probleminin iyi tanımlanması gerekir.
- Bu aşamada, ihtiyaç belirlenerek makine öğrenmesi ile kurulacak modelin amacının ne olduğu ortaya konmalıdır.
- Problem için hangi verileri toplayacağımız (E), makine öğrenme modelinin hangi kararları vermesi gerektiği (T) ve sonuçlarını nasıl değerlendireceğimiz (P) hakkında net bir şekilde düşünmemiz gerekir.

1. Problemin Tanımlanması

- Örneğin,
- Bir kanser teşhisi öğrenme problemi için:
- **Deneyim (Experience) E:** Kanserli ve kanserli olmayan hastalara ait örnek olarak alınan kayıtlar veya gözlemler.
- **Görev (Task) T:** Hastaların kanser ya da kanser değil biçiminde sınıflandırılması.
- **Performans (Performance) P:** Sınıflandırma doğruluğu. Doğru sınıflandırılan hastaların yüzdesi.

2. Veriyi Anlama

- Makinenin deneyim olarak yararlanacağı veri, ele alınan probleme uygun bir biçimde temin edilir.
- Toplanan veri yapılandırılmış (veritabanı düzenindeki gibi), ya da yapılandırılmamış olabilir.
- Veri toplama aşamasında farklı kaynaklardan yararlanılabilmektedir.
- Örneğin, internette yer alan hazır veri kümelerinden yararlanılabilir.

2. Veriyi Anlama

- İnternette erişilebilen veri kümeleri, erişim ve kullanım kolaylığı bakımından makine öğrenmesi çalışmalarına avantaj sağlamaktadır.
- University of California Irvine (UCI) Machine Learning Repository, Kaggle Datasets, Microsoft Research Open Data gibi paylaşım açılan veri kümelerine erişmek mümkündür.



2. Veriyi Anlama

- Verinin analizler için hazırlaması belki de makine öğrenmesi sürecinin en zaman alıcı aşaması olarak kabul edilebilir.
- Modelin kurulmasından önce mevcut verinin iyi anlaşılması ve iyi analiz edilmesi gerekmektedir.
- Veri kümesi temin edildikten sonra veri hakkında ön fikir edinilmesi için bazı basit istatistiksel hesaplamalar yapılabilir ve grafikler çizilebilir.

3. Veri Hazırlama

Veri Hazırlama,

- ***Veri temizleme***
 - ***Veri bütünleştirme***
 - ***Veri indirgeme***
 - ***Veri dönüştürme, veri normalleştirme***
 - ***Veriyi eğitim ve test kümelerine ayırma***
- olmak üzere beş aşamada incelenebilir.

3. Veri Hazırlama

- **Veri Temizleme**

- Bazı uygulamalarda, üzerinde çözümleme yapılacak verilerin istenen özelliklere sahip olmadığı görülebilir.
- Örneğin **eksik verilerle ve uygun olmayan verilerin** oluşturduğu tutarsız verilerle karşılaşılabilir.
- Bu gibi durumlarda verinin söz konusu sorunlardan temizlenmesi gerekecektir. Eksik verilerin yerine yenileri belirlenerek konulmalıdır.

3. Veri Hazırlama

Veri Temizleme

Veri Temizleme için aşağıda belirtilen yöntemlerden biri kullanılabilir.

a) Eksik değer içeren kayıtlar veri kümesinden atılabilir.

b) Kayıp değerlerin yerine bir genel sabit kullanılabilir. Bütün kayıp değerler için aynı sabit kullanılabilir.

Örneğin "bilinmiyor" değeri bu eksik veri yerine kullanılabilir.

Ancak bütün özniteliklerde kayıp değerler yerine aynı sabit değer kullanımı bazen sorun oluşturabilir.

3. Veri Hazırlama

Veri Temizleme

- c) Özniteliğin tüm verileri kullanılarak ortalaması hesaplanır ve eksik değer yerine bu değer kullanılabilir.
- d) Özniteliğin tüm verileri yerine, sadece bir sınıfa ait örneklerin değişken ortalaması hesaplanarak eksik değer yerine kullanılabilir.
- e) Verilere uygun bir tahmin yapılarak, örneğin regresyon ya da karar ağacı modeli kurularak eksik değer tahmin edilebilir ve eksik değer yerine kullanılabilir.

3. Veri Hazırlama

- **Veri Bütünleştirme**

- Farklı veri tabanlarından ya da veri kaynaklarından elde edilen verilerin birlikte değerlendirmeye alınabilmesi için farklı türdeki veriler tek türe dönüştürülür.
- Yani veri bütünleştirilmesi söz konusu olacaktır.

3. Veri Hazırlama

- **Veri İndirgeme**

- Makine öğrenmesi uygulamalarında bazen çözümleme işlemi uzun süre alabilir.
- Eğer çözümlemeden elde edilecek sonucun değişmeyeceğine inanılıyorsa veri sayısı ya da özniteliklerin sayısı azaltılabilir.
- Veri indirgeme çeşitli biçimlerde yapılabilir;
 - a. Veriyi birleştirme veya veri küpü
 - b. Boyut indirgeme
 - c. Veri sıkıştırma
 - d. Örnekleme
 - e. Genelleme

3. Veri Hazırlama

- **Veri İndirgeme**

- Veriyi indirgeme aşamasında verilerin çok boyutlu veri küpleri biçimine dönüştürmek söz konusu olabilir. Böylece çözümlmeler sadece belirlenen boyutlara göre yapılır.
- Veriler arasında bir seçme işlemi yapılarak, gereksiz veriler veri tabanından çıkarılır ve boyut azaltılması sağlanabilir.

3. Veri Hazırlama

- **Veri İndirgeme**

- Veri sıkıştırma aşamasında, büyük veri kümelerinin sıkıştırılarak daha az yer işgal etmeleri sağlanabilir.
- Örnekleme aşamasında ise, büyük veri topluluğu yerine onu temsil eden daha küçük veri kümelerinin oluşturulması amaçlanır.
- Genelleme verilerin tek tek değil genel kavramlarla ifade edilmesini sağlar.

3. Veri Hazırlama

- **Veri Dönüştürme**

- Veriyi bazı durumlarda makine öğrenmesi çözümlmelerine aynen katmak uygun olmayabilir.
- Özniteliklerin ortalama ve varyansları birbirlerinden önemli ölçüde farklı olduğu takdirde büyük ortalama ve varyansa sahip özniteliklerin diğerleri üzerindeki baskısı daha fazla olur ve onların rollerini önemli ölçüde azaltır.

3. Veri Hazırlama

- **Veri Dönüştürme**

- Ayıca özniteliklerin sahip olduğu çok büyük ve çok küçük değerler de çözümlemelerin sağlıklı biçimde yapılmasını engeller.
- Bu nedenle bir dönüşüm yöntemi uygulayarak söz konusu **örneklerin normalleştirilmesi veya standartlaştırılması** uygun bir yol olacaktır.

3. Veri Hazırlama

- **Veri Normalleştirme :**

- Veriler normalize edilerek aşırı uç değerlerden kurtarılabilmekte ve böylece eldeki veriler daha iyi biçimde modellenebilmektedir.
- Veri normalleştirme, işlem süreci sırasında kullanılan verilerin kümülatif toplamların oluşturacağı olumsuzlukların engellenmesini sağlamaktadır.

3. Veri Hazırlama

Veri Normalleştirme

1. [0, +1] aralığında doğrusal dönüşüm:

$$x_n = (x_0 - x_{\min}) / (x_{\max} - x_{\min})$$

2. [a,b] aralığında doğrusal dönüştürme:

$$x_n = (b-a)(x_0 - x_{\min}) / (x_{\max} - x_{\min}) + a$$

3. İstatiksel normalizasyon:

$$x_n = (x_0 - \mu) / s$$

4. Basit normalizasyon:

$$x_n = x_0 / x_{\max}$$

Formüllerdeki **xn normalleştirilmiş** ve **x0 orijinal veriyi**, x_{\min} , x_{\max} , μ , s satır veya sütün boyunca sırasıyla **minimum**, **maksimum**, **ortalama** ve **standart sapma** değerlerini ifade etmektedir.

3. Veri Hazırlama

- Veriyi eğitim ve test kümelerine ayırma
- Veri kümesi eğitim ve test kümeleri olarak ayrılır.
- Bu ayırma işlemi çeşitli şekillerde yapılabilir.
- Örneğin veri kümesinin %75'lik bir kısmını eğitim %25'lük bir kısmını ise test için ayırmak gibi.
- Eğitim veri kümesi ile model eğitildikten sonra test veri kümesi ile model başarısının sınanması yapılır.

3. Veri Hazırlama

- **Veriyi eğitim ve test kümelerine ayırma**
- Model doğrulama ve ayarlama için isteğe bağlı bir üçüncü doğrulama veri kümesi de ayrılabilir.
- Genel olarak, daha fazla eğitim verisi, daha iyi bir model ve potansiyel performans ile sonuçlanır.
- Daha fazla test verileri, model performansının ve genel genelleme kabiliyetinin daha fazla değerlendirilmesiyle sonuçlanır.

4. Modelin kurulması

- Veri hazır hale getirildikten sonra konuyla ilgili makine öğrenmesi algoritmaları uygulanır.
- Seçilen algoritma ve bu algoritma yardımıyla elde edilen model, uygulamadaki gelişime göre değişebilir.
- Bir model seçerken bazı hususlar dikkate alınmalıdır.
 - *Doğruluk (accuracy)*
 - *Hız (eğitim, test etme ve gerçek zamanlı işleme)*
 - *Ölçeklenebilirlik (scalability)*
 - *Sadelik (simplicity)*
 - *Yorumlanabilirlik (Interpretability)*

5. Modelin değerlendirilmesi

- Makine öğrenmesi algoritması eğitim veri kümesi üzerinde uygulanıp model oluşturulduktan sonra, model değerlendirmesi test veri kümesi kullanılarak gerçekleştirilir ve sonuçlar değerlendirmeye alınır.
- Değerlendirme, hangi makine öğrenmesi modelinin daha iyi olduğunu bulmak ve gelecekte modelin ne kadar iyi çalışacağını anlamak için yardımcı olur.

6. Modelin uygulanması

- Kurulan makine öğrenme modeli, model performans değerlendirme yöntemleri ve ölçülerine göre değerlendirilir.
- Değerlendirme işlemi sonucunda model performansı yeterli görülüyor ve herhangi bir sorun veya eksiklik görülmiyorsa modelin uygulamaya geçilmesi sağlanır.

Öğrenme Yöntemleri

- Makine öğrenmesinde kullanılan öğrenme yöntemleri daha çok **örneklerden öğrenme (learning from examples)** şeklindedir ve dört başlıkta incelenebilir,

1. Denetimli öğrenme (Supervised Learning)

- ☐ Sınıflandırma (Classification)
- ☐ Öngörü (Prediction)

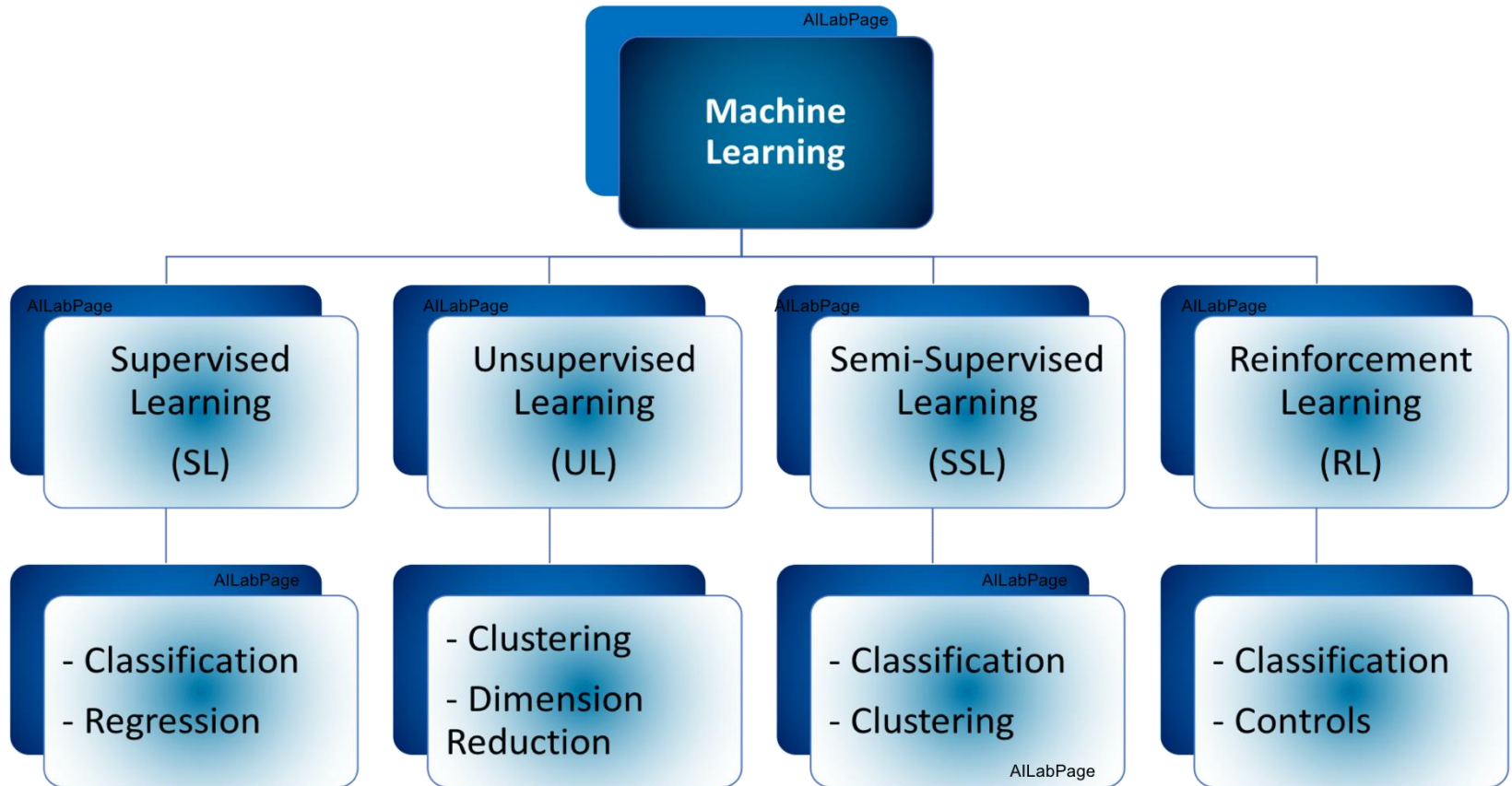
2. Denetimsiz öğrenme (Unsupervised Learning)

- ☐ Kümeleme (Clustering)

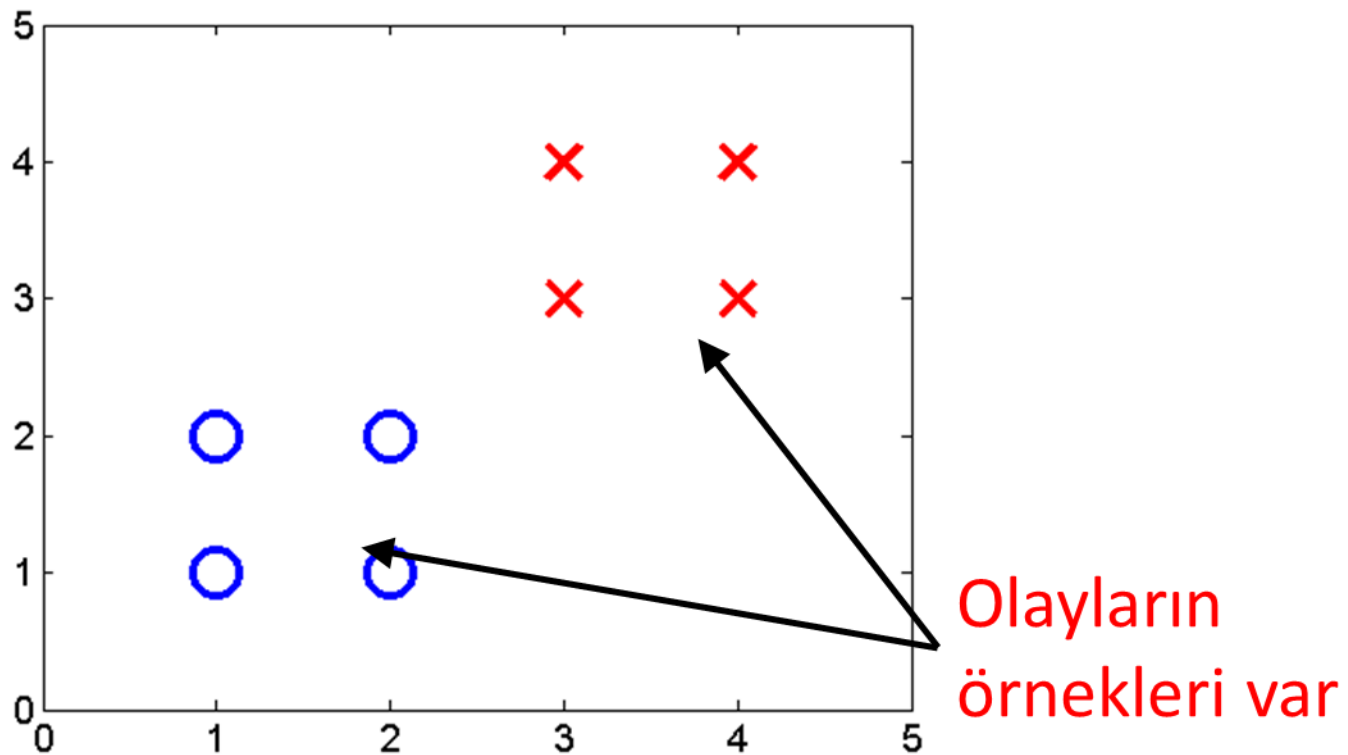
3. Yarı-denetimli öğrenme (Semi-supervised learning)

4. Pekiştirmeli öğrenme (Reinforcement learning)

Öğrenme Yöntemleri



Denetimli öğrenme (Supervised learning)



Denetimli öğrenme için örnek veri kümesi

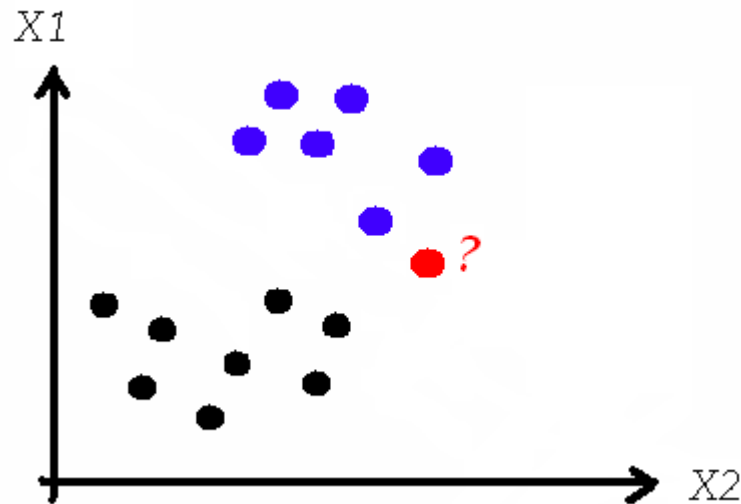
(Öznitelikler-features)

(Sınıf Etiketi-Class Label)

	Yaş	Boy	Kilo	Cinsiyet
→	20	175	70	Erkek
→	21	179	80	Erkek
→	19	162	50	Kız
→	22	169	55	Kız
→	20	183	90	Erkek
→	19	181	75	Erkek
→	21	171	57	Kız

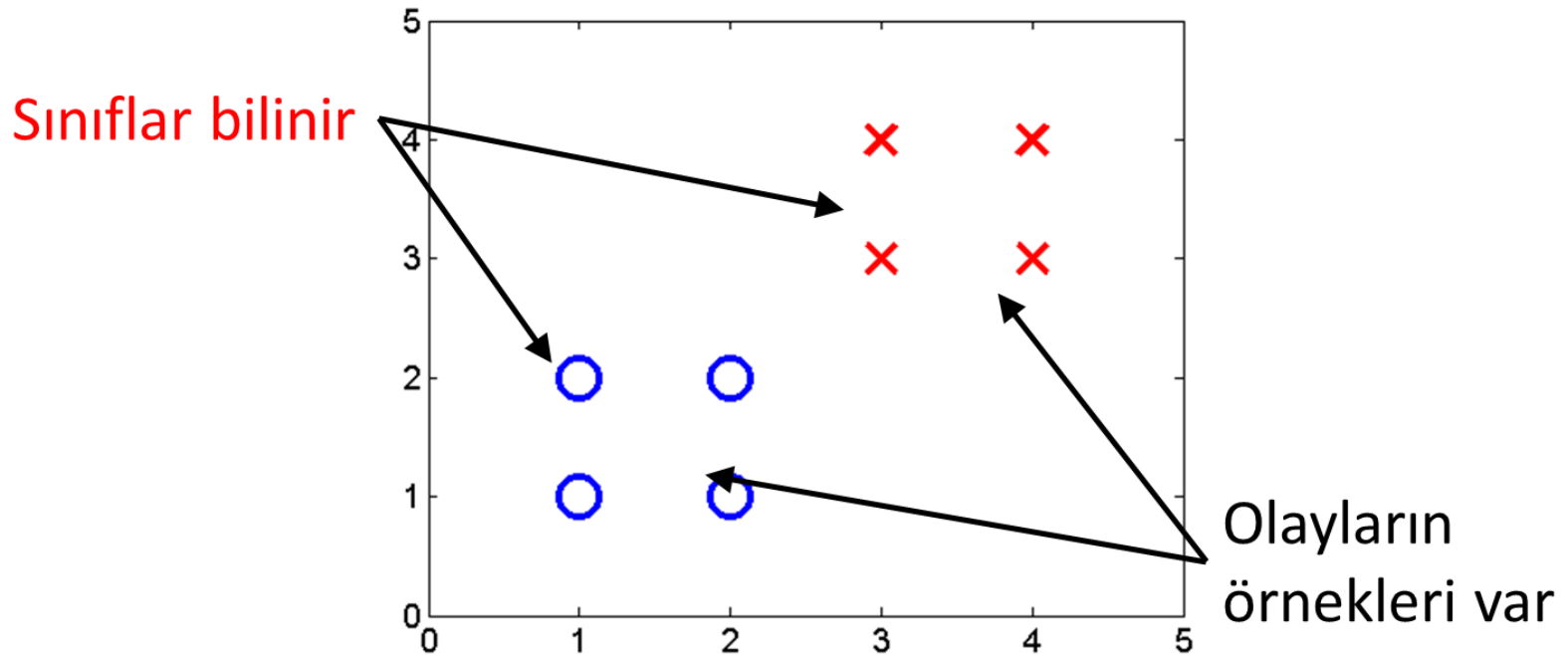
→(Örnekler-samples)

Denetimli öğrenme (Supervised learning)



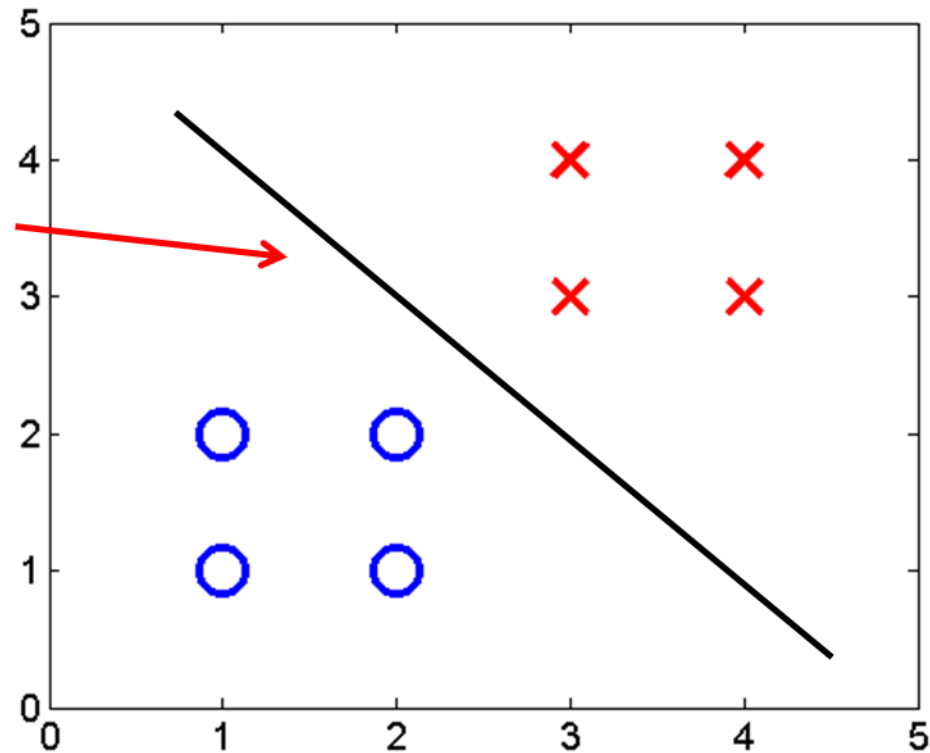
Kırmızı hangi sınıftan?

Denetimli öğrenme (Supervised learning)



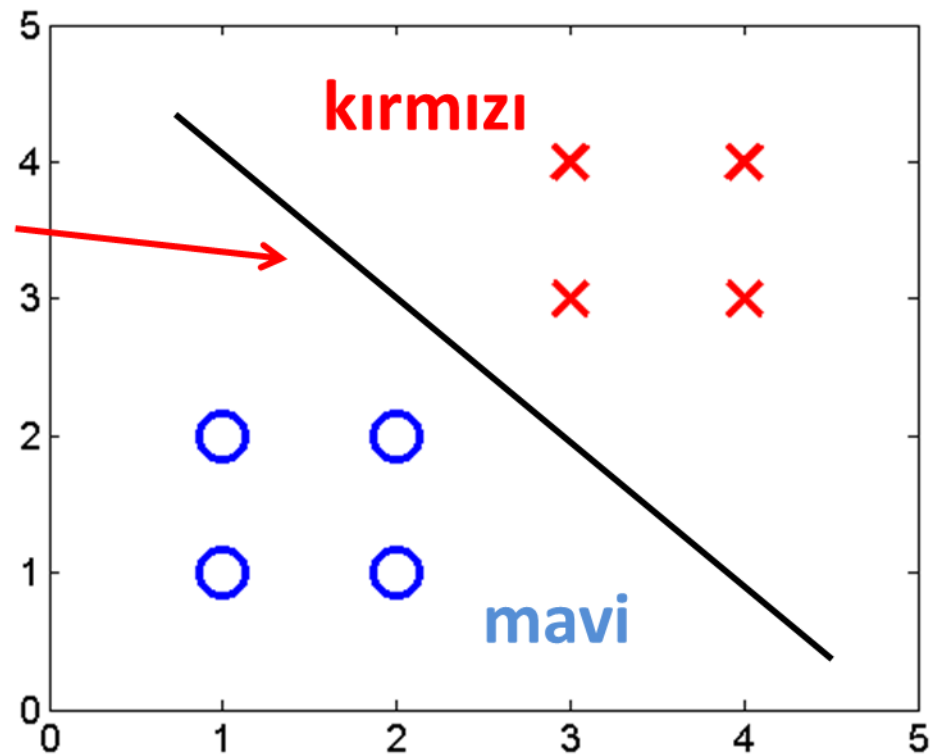
Denetimli öğrenme (Supervised learning)

Bir mümkün
karar modeli
– bu çizgi



Denetimli öğrenme (Supervised learning)

Bir mümkün
karar modeli
– bu çizgi



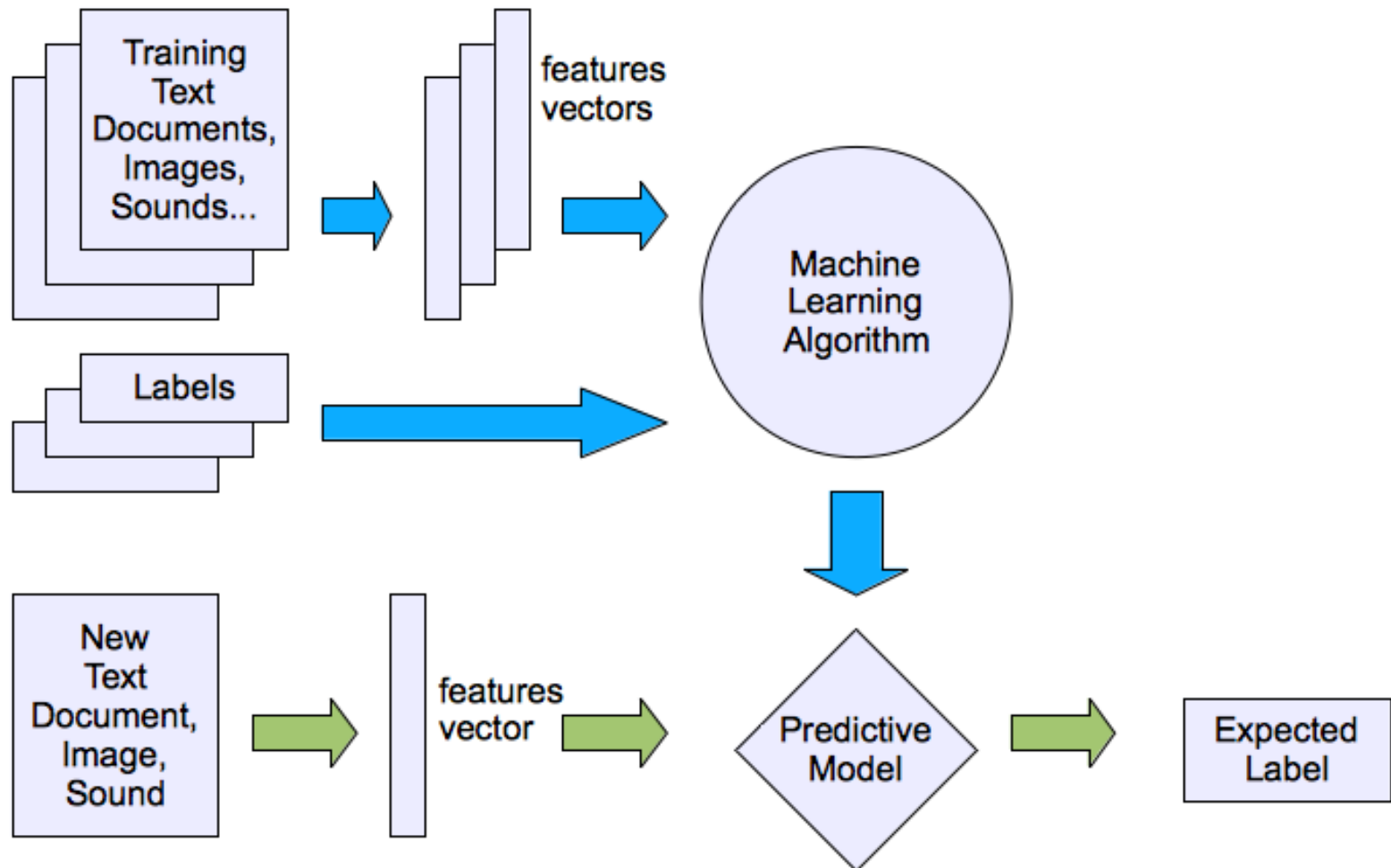
Denetimli öğrenme (Supervised learning)

- Bu öğrenmede oluşturulan modelin ürettiği çıktılar ile hedef (beklenen) çıktılar arasındaki fark hata olarak ele alınır ve bu hata minimize edilmeye çalışılır.
- Denetimli öğrenme algoritmasının bir “öğretmene” veya “danışmana” ihtiyacı vardır.
- Widrow-Hoff tarafından geliştirilen Delta kuralı ve Rumelhart ve McClelland tarafından geliştirilen genelleştirilmiş Delta kuralı veya geriyayılım (backpropagation) algoritması **denetimli öğrenme algoritmalarına** örnek olarak verilebilir.

Denetimli öğrenme (Supervised learning)

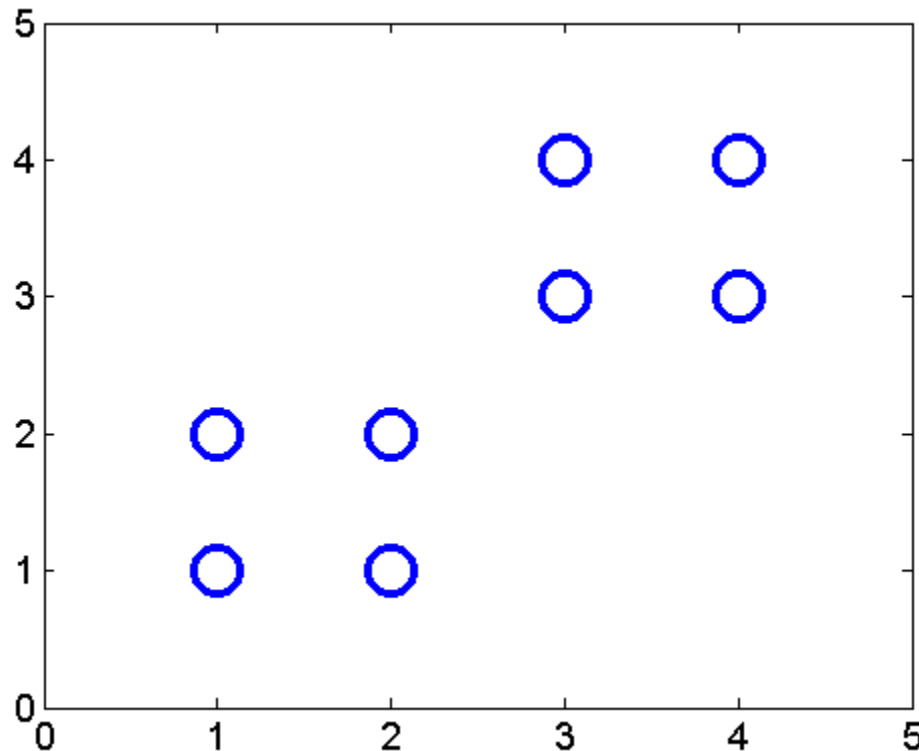
- Denetimli öğrenme en yaygın kullanılan öğrenme tipidir.
 - Olayların verileri ve bu verilere karşılık gelen çıktılar, sonuçlar, yada kararlar **(sınıf etiketleri)** bilinmektedir.
 - Var olan olayların örnekleri kullanarak makine öğrenme modeli genellemeye çalışmaktadır.
 - Önceden görülen olayın örnekleri kullanılarak gelecek durumlar için sonuç öngörülme çalışılır.

Denetimli öğrenme (Supervised learning)



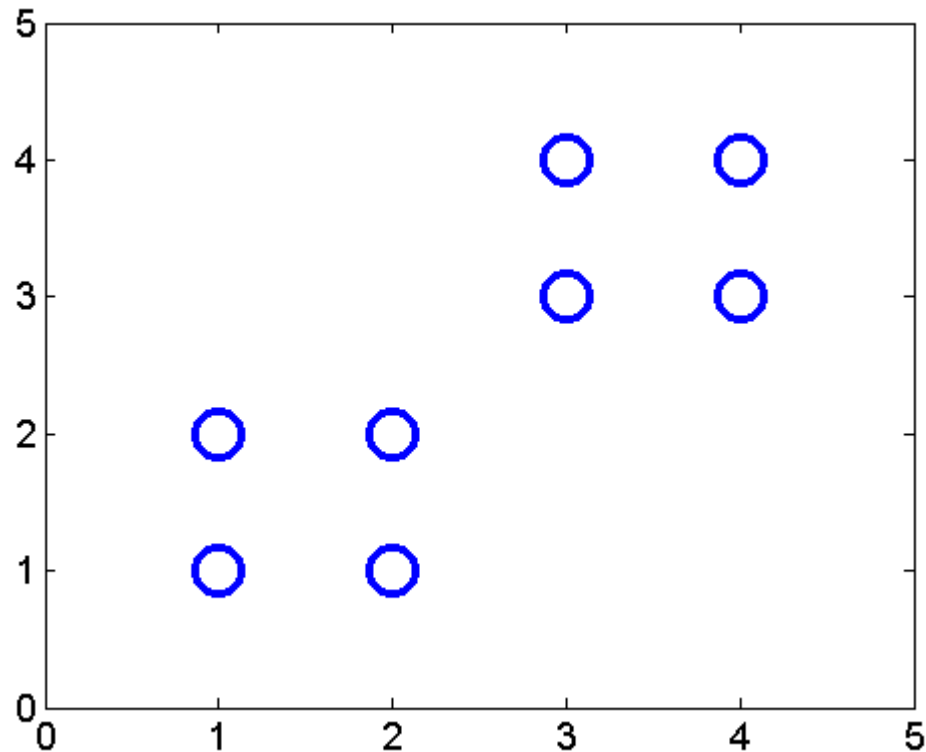
Denetimsiz öğrenme (Unsupervised learning)

Olayların örnekleri bilinir, ama onlara karşı gelen sonuçları bilinmez.



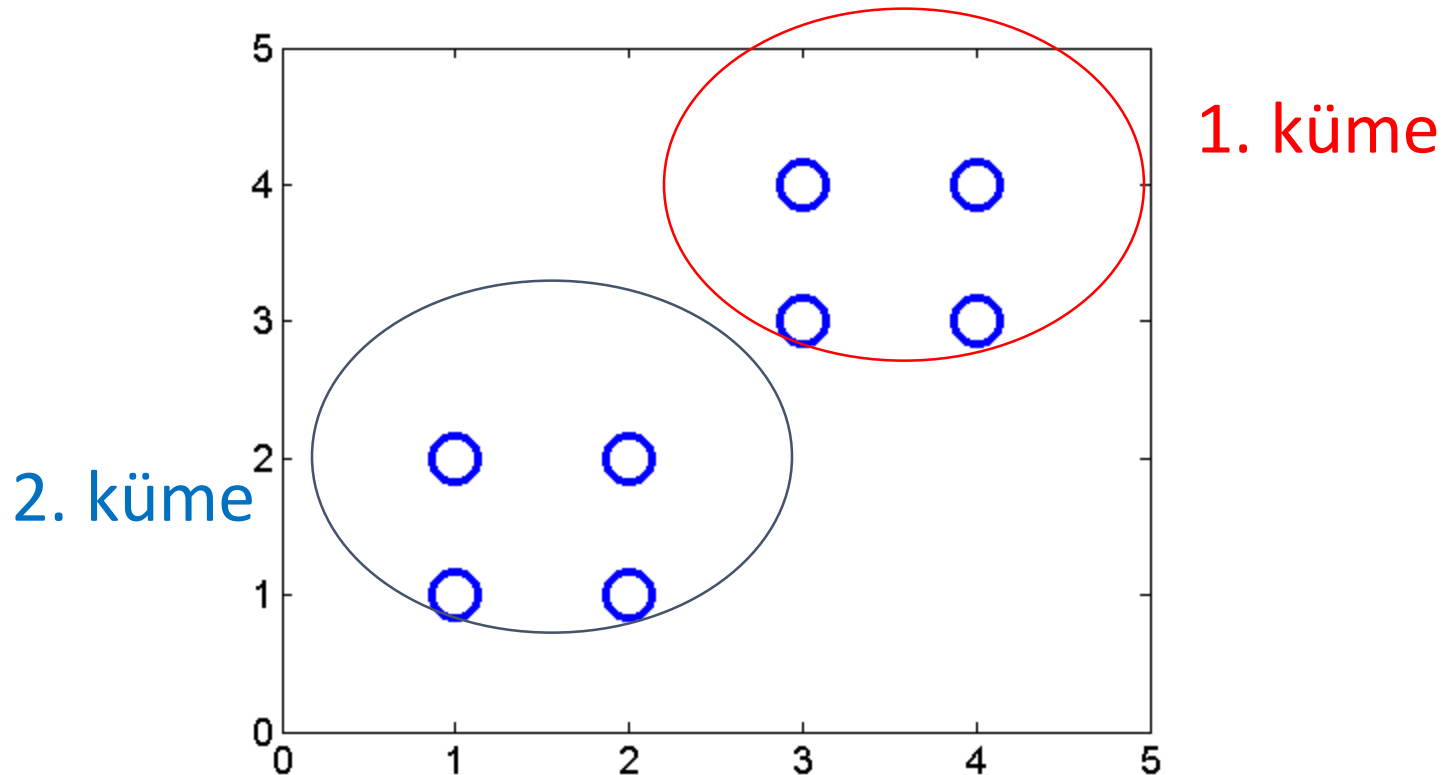
Denetimsiz öğrenme (Unsupervised learning)

Kurulan modelin kendi kendine çalışarak verilerin var olan yapısını bulması gerekiyor.



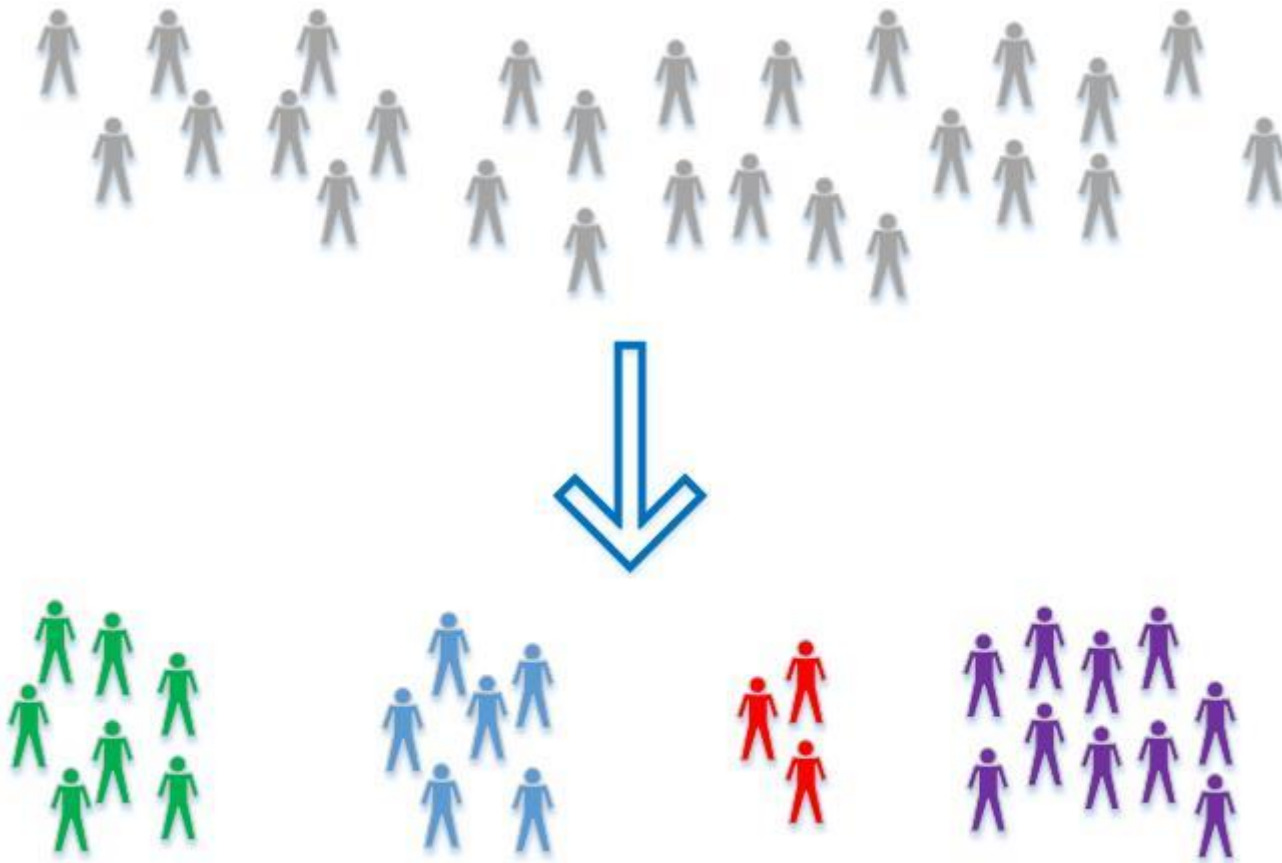
Denetimsiz öğrenme (Unsupervised learning)

Bu işlem “**kümeleme**” yada “**clustering**” olarak adlandırılır. Yani var olan olayların birkaç uygun küme/sınıfa konulması gerekiyor.



Denetimsiz öğrenme (Unsupervised learning)

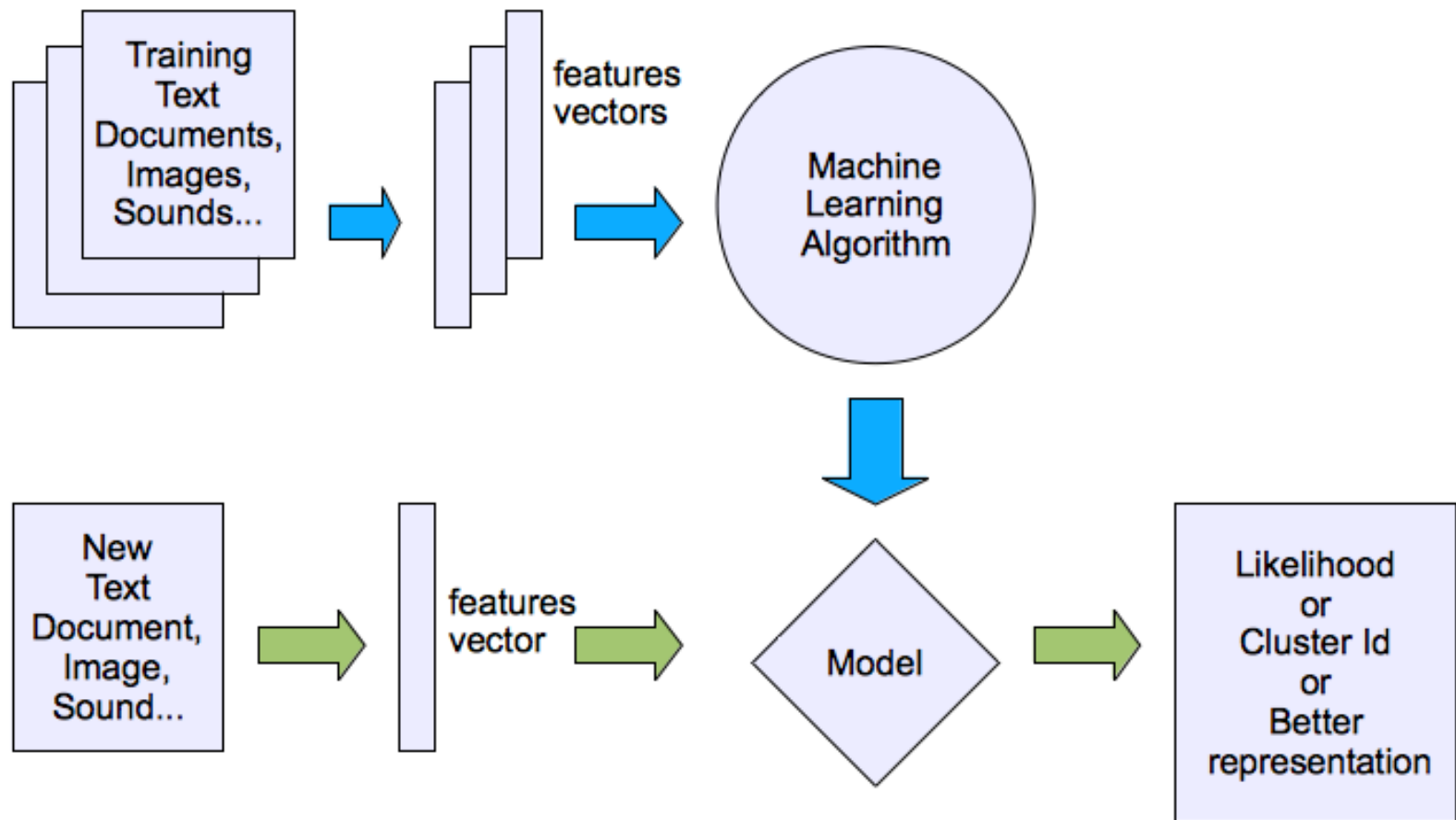
- Kümeleme-Clustering



Denetimsiz öğrenme (Unsupervised learning)

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

Denetimsiz öğrenme (Unsupervised learning)



Etiketlenmiş ve Etiketlenmemiş veri (Labeled and Unlabeled data)



Etiketlenmemiş veri
Ucuz ve bol!



Human expert/
Special equipment/
Experiment



“Crystal” “Needle” “Empty”

“0” “1” “2” ...

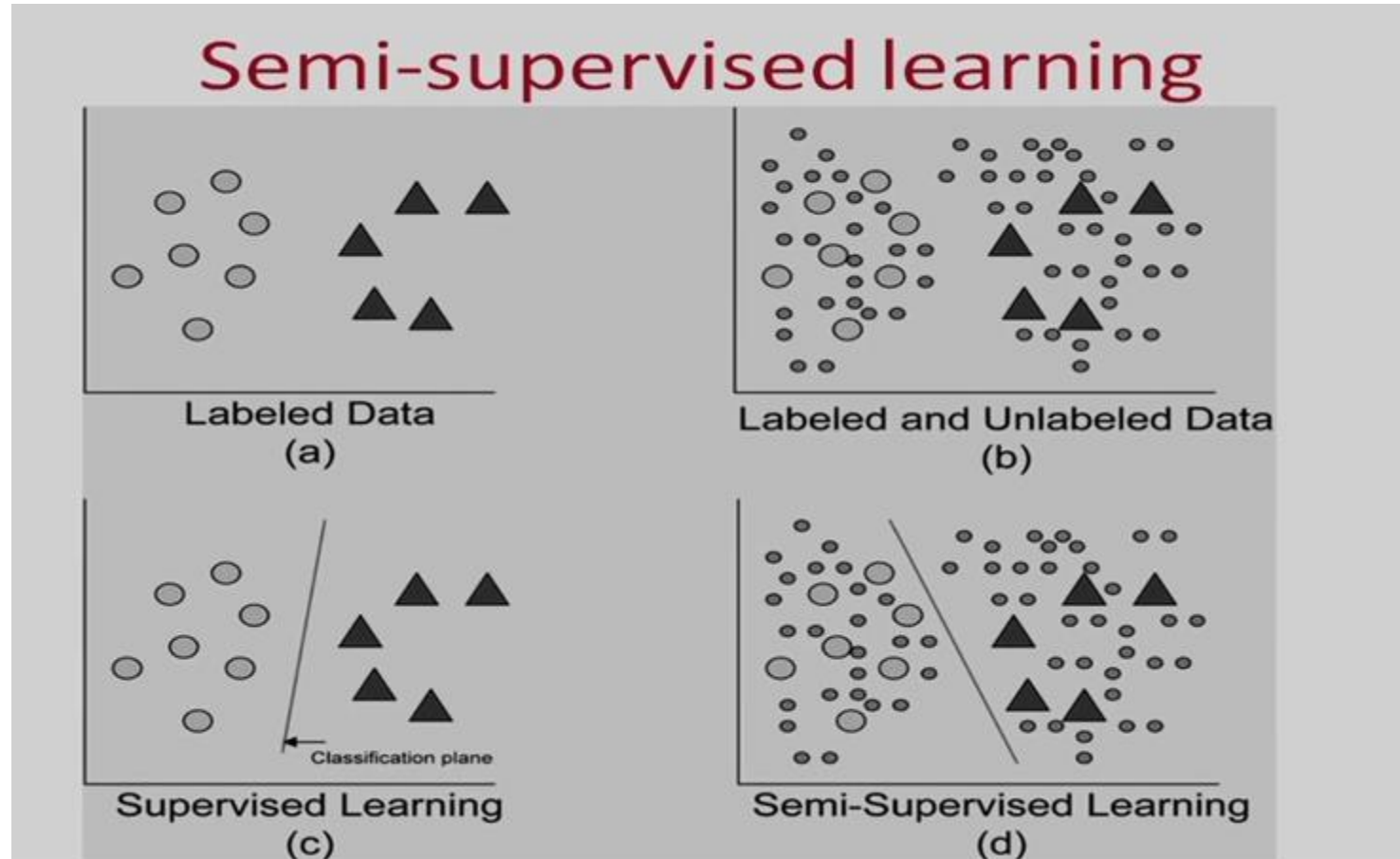
“Sports”
“News”
“Science”
...

Etiketlenmiş veri
Pahalı ve kıt!

Yarı-denetimli öğrenme (Semi-supervised learning)

- Yarı denetimli öğrenme, denetimsiz öğrenme (etiketli eğitim verisi olmadan) ve denetimli öğrenme (etiketli eğitim verileriyle) arasında bir öğrenme yöntemidir.
- Yarı denetimli öğrenme, eğitim için hem etiketlenmiş hem de etiketlenmemiş veri kümelerini kullanır.
- Genellikle etiketlenmemiş veri sayısı etiketli verilerden daha çok miktardadır.

Yarı-denetimli öğrenme (Semi-supervised learning)



Yarı-denetimli öğrenme (Semi-supervised learning)

- Veri kümesi örneklerinin tümünü etiketlemek uzmanlar için yoğun zaman ve maliyet gerektiren bir iştir.
- Veri kümesi örneklerinin az bir kısmı etiketlendiği için, yarı-denetimli öğrenme, etiketleme işlemiyle ilişkili maliyeti ve zamanı oldukça azaltır.
- Doğru bir model üretmek için yeterli etiketli veri olmadığında ve daha fazlasını elde etme yeteneğiniz veya kaynaklarınız olmadığında, yarı denetimli öğrenme kullanılabilir.

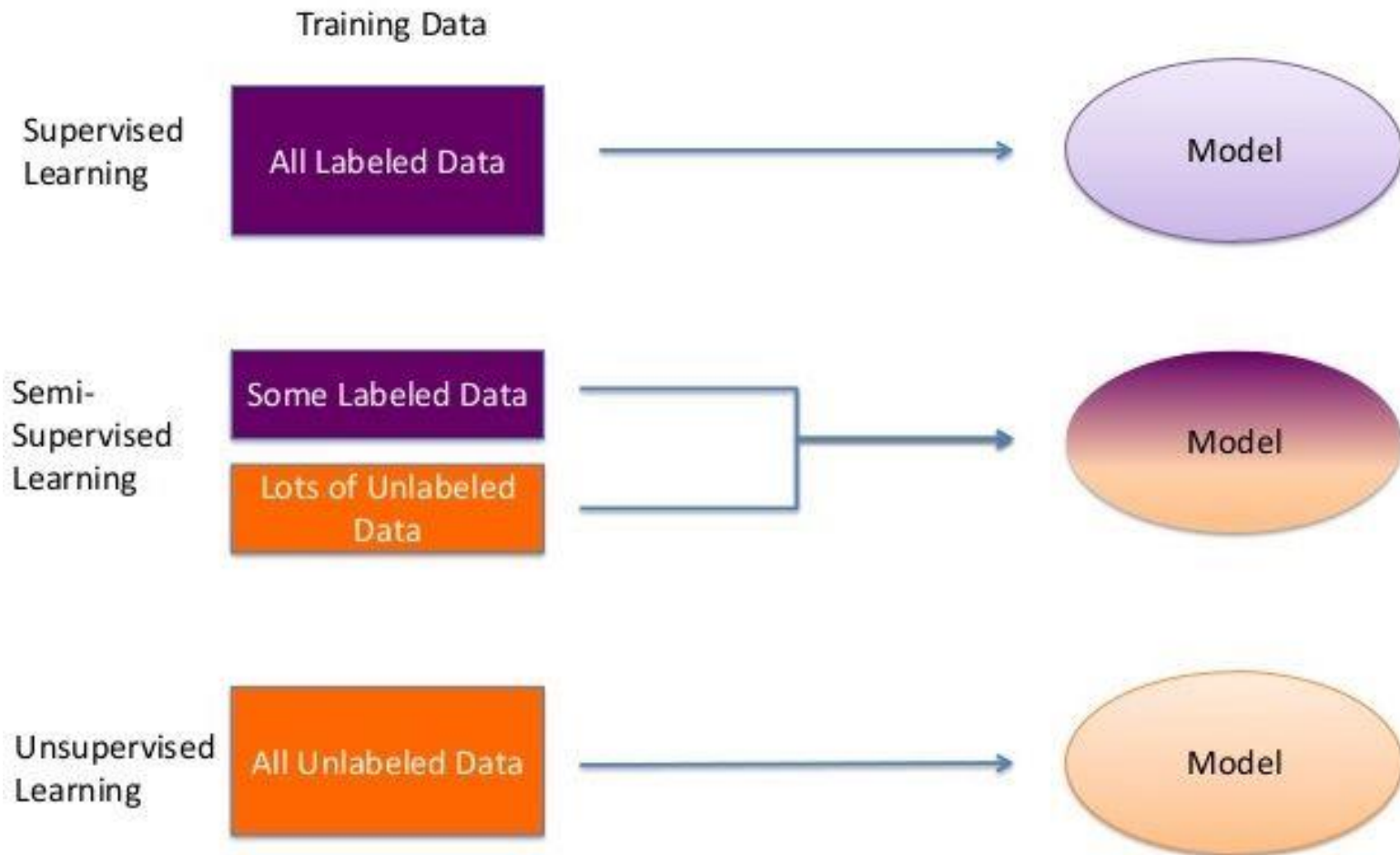
Yarı-Denetimli Öğrenme Algoritmaları

- Self-Training
- Generative methods, mixture models
- Graph-based methods
- Co-Training
- Semi-supervised SVM
-

Denetimli, Denetimsiz ve Yarı-Denetimli Öğrenme

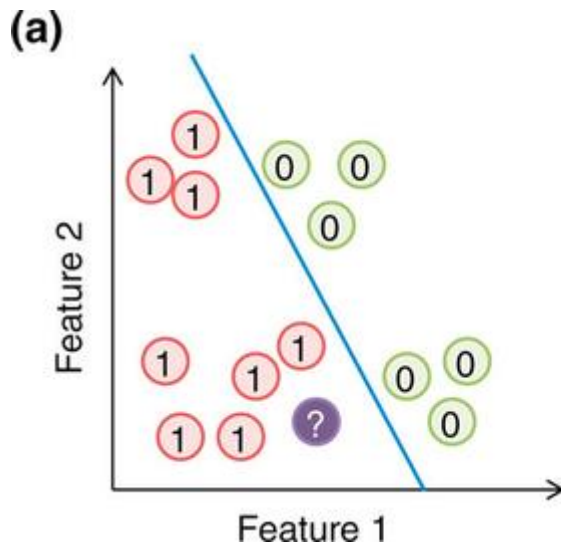
- Denetimli öğrenme (Supervised learning) :
Sınıflandırma, Öngörü
- Veri kümesindeki bütün örnekler etiketlenmiştir
- Denetimsiz öğrenme (Unsupervised learning) :
Kümeleme (Demetleme, öbekleme,...)
- Veri kümesinde hiçbir örnek etiketlenmemiştir.
- Yarı-denetimli Öğrenme (Semi-supervised learning)
- Veri kümesinde az miktarda örnek etiketlenmiş, çok miktarda örnek ise etiketlenmemiştir.

Denetimli, Denetimsiz ve Yarı-Denetimli Öğrenme

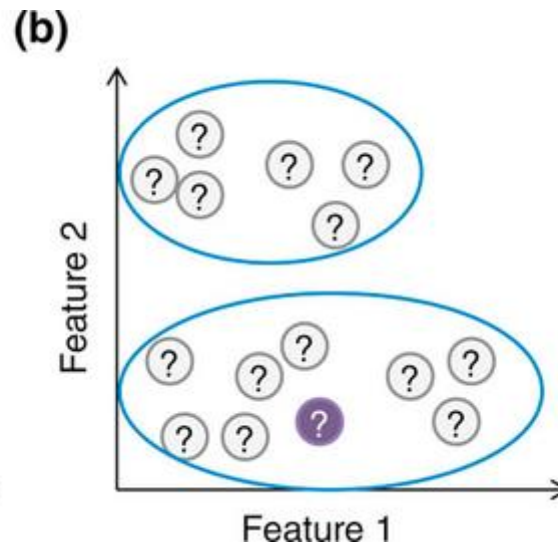


Denetimli, Denetimsiz ve Yarı-Denetimli Öğrenme

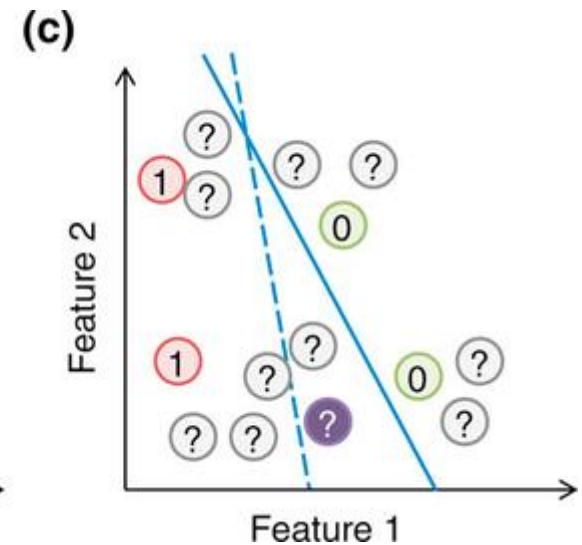
- a) Denetimli
- Sınıflar bilinir



- b) Denetimsiz
- Sınıflar bilinmez



- c) Yarı-Denetimli
- Sınıflar kısmen bilinir



Sınıflandırma ve Regresyon (Classification and Regression)

- Hem sınıflandırma hem de regresyon için;
- Önceden var olan verilerden yola çıkılarak olayın uygun modeli bulunur, ve gelecekteki kararlar için bu model kullanılır.
- **Regresyon (Regression)**
 - Sürekli değere sahip fonksiyonları modeller.
 - Bu modele göre eksik yada bilinmeyen değerleri tahmin eder.
- **Sınıflandırma (Classification)**
 - Amaç probleme ait tüm uzayın belirli sayıda sınıfa bölünmesidir.
 - Bir model oluşturur ve veriyi sınıflandırır.

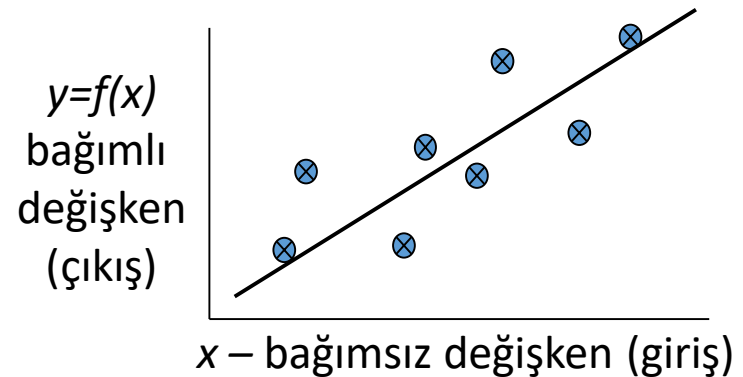
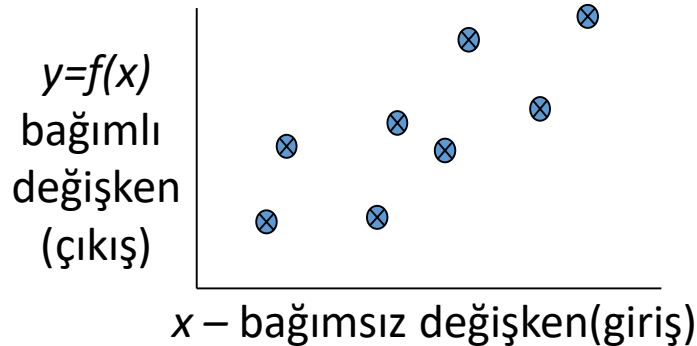
Regresyon (Regression)

- Regresyon probleminde, olayın modeli sürekli (continious) bir modeldir,
- Modellenecek değişkenlerin sürekli şekilde değişmesi gerekiyor.
- Makine öğrenme problemi, bu modellenecek değişkenlerin değişimine karar vermek için uygun modeli oluşturmaya çalışır.

Regresyon (Regression)

En basiti doğrusal regresyon (linear regression).

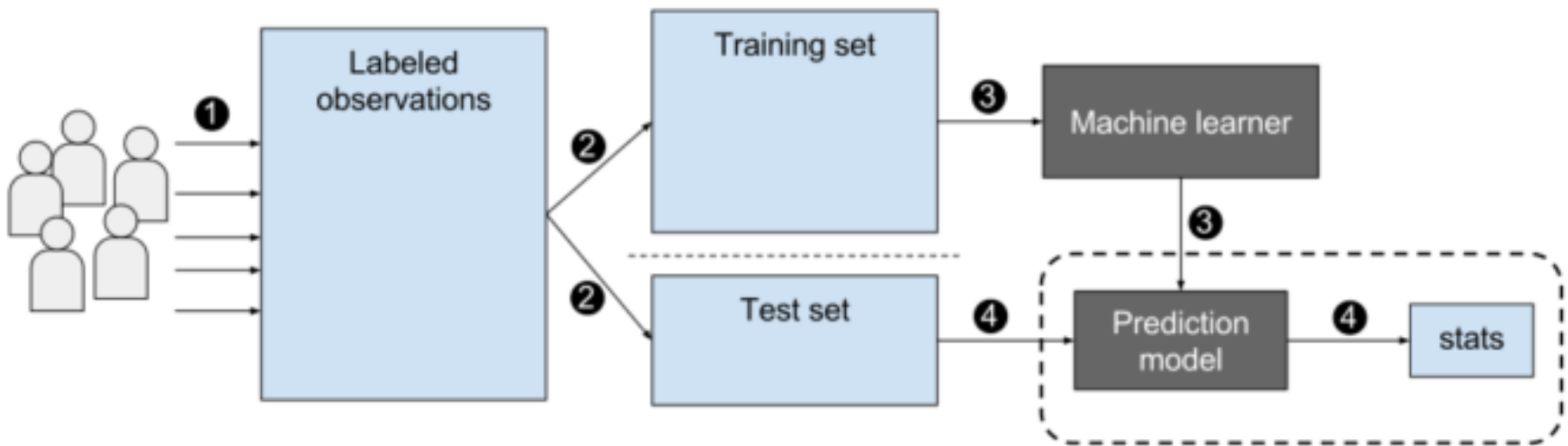
Doğrusal regresyon modelinde, neden ve sonuç arasında doğrusal bir ilişki olduğu varsayılmaktadır.



Sınıflandırma (Classification)

- Sınıflandırmada, olayın modeli sürekli bir model değildir, yani ayırık (discrete) modeldir.
- Modellenecek değişkenlerin birkaç sınıfta bulunabilmesi gerekir.
- Sınıflandırma işlemi üç aşamadan oluşur:
 1. **Model oluşturma** (Eğitim kümesi üzerinden öğrenme)
 2. **Model değerlendirme** (Öğrenilen değerlerle test kümesi üzerinde sınıflandırma)
 3. **Modeli kullanma**

Sınıflandırma (Classification)



Sınıflandırma (Classification)

- **1. Model Oluşturma:**

- Hangi sınıfa ait olduğu bilinen örnekler (Eğitim kümesi: training set) ile bir model oluşturulur.
- Eğitim kümesi (Training set) $x = (x_1, x_2, \dots, x_n)$
 - Ayırık örneklerden oluşur.
- Eğitim kümesi özniteliklerden oluşur, özniteliklerden biri sınıf bilgisidir (sınıf etiketi-class label). Öznitelikler de örneklerden oluşmuştur.

Sınıflandırma (Classification)

- **2. Model değerlendirme**

- Oluşturulan modelin başarımı belirlenir.
- Oluşturulan model eğitim kümesinde yer almayan örnekler (test-sınama veri kümesi: test set) ile denenerek başarısı ölçülür.
- Amaç : Eğitim veri kümesinde yer almayan örnekleri mümkün olan en yüksek başarımla doğru sınıflara atamak.
- **Modelin başarımı** : Doğru sınıflandırılmış test veri kümesi örneklerinin toplam test veri kümesi örneklerine oranıdır.
- Test veri kümesi model öğrenirken kullanılmaz. Test veri kümesi ile eğitim veri kümesi bağımsız olmalı.

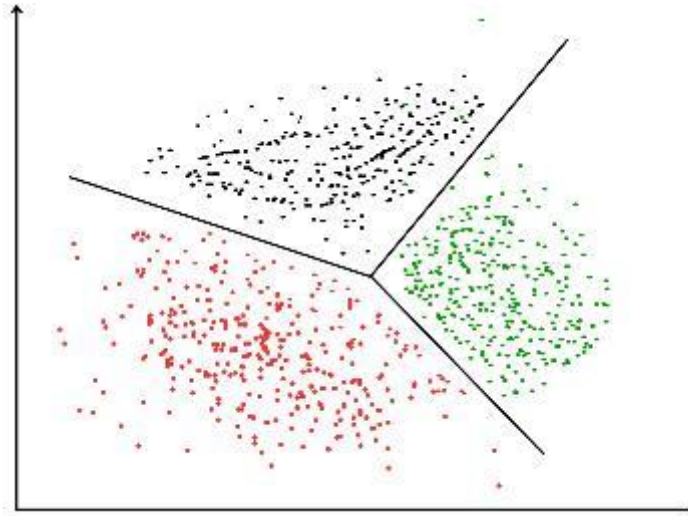
Sınıflandırma (Classification)

- **3. Modeli kullanma**

- Model daha önce görülmemiş ve sınıfı belli olmayan örnekleri sınıflandırmak için kullanılır.
- Örneklerin sınıf etiketlerini tahmin etme
- Bir öz niteliğin değerini tahmin etme

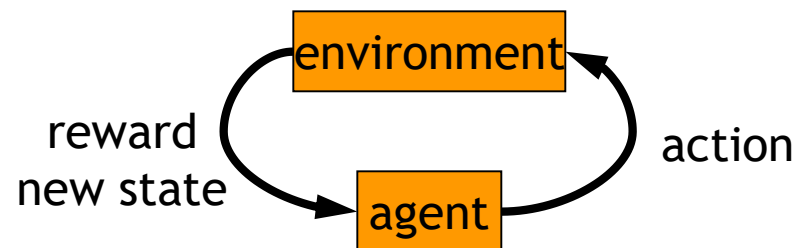
Kümeleme (Clustering)

- Kümeleme, mevcut etiketlenmemiş veriler arasındaki içsel gruplandırmayı belirleme işlemidir.
- Gruplamayı, temel olarak verilerin aralarındaki benzerlik ve farklılıklara dayanarak gerçekleştirir.
- Benzerlik ve farklılık, verilerin birbirlerine olan uzaklıkları veya veri uzayında yoğunlaştıkları noktalar gibi yaklaşımlarla belirlenir.
- K-means kümeleme algoritması, kümeleme problemini çözen en basit ve yaygın olarak tercih edilen denetimsiz öğrenme algoritmasıdır.



Pekiştirmeli öğrenme (Reinforcement learning)

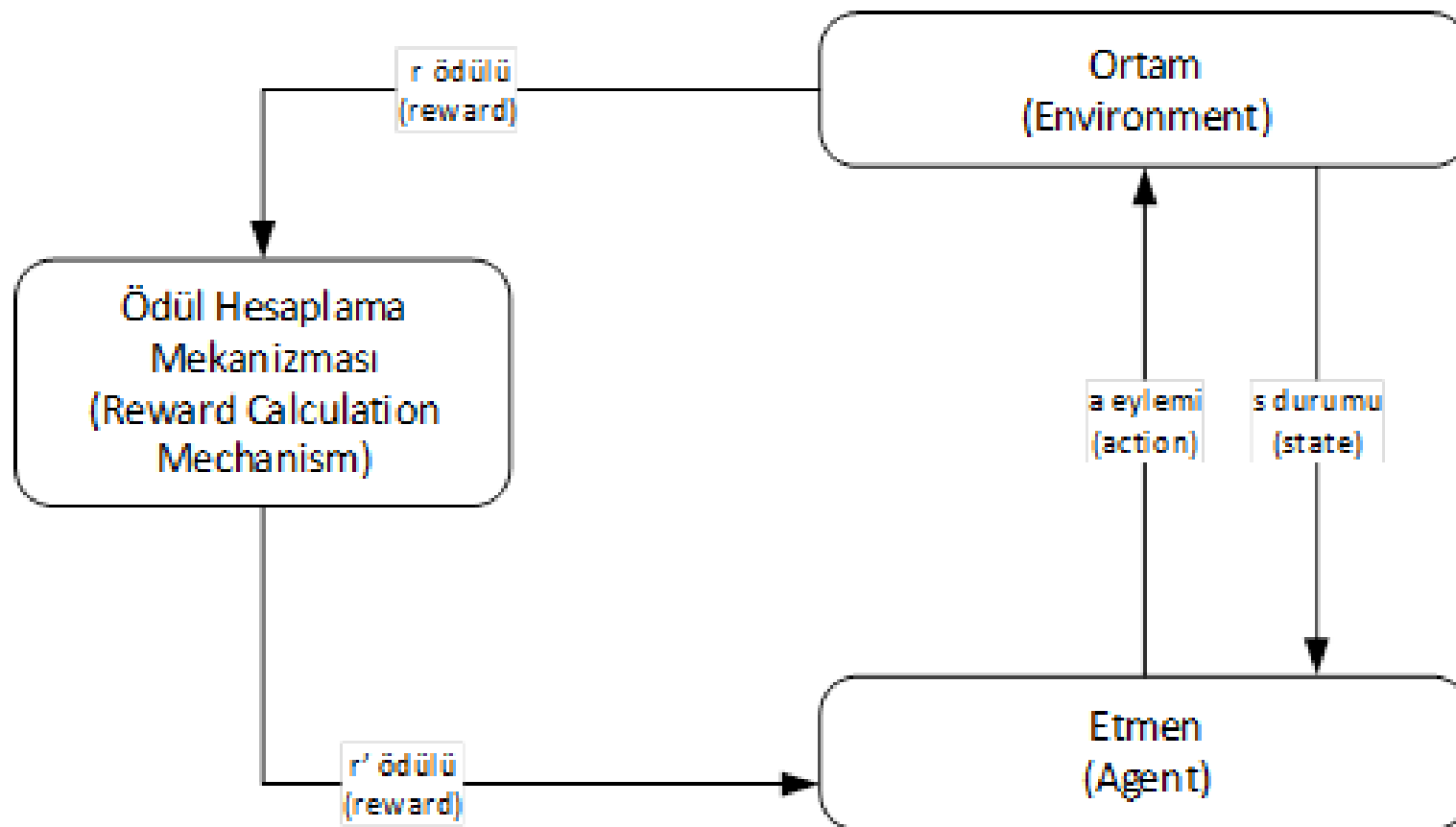
- Pekiştirmeli öğrenme, denetimli öğrenme (supervised learning) ve denetimsiz öğrenme (unsupervised learning) metodlarından farklıdır.
- Pekiştirmeli öğrenme algoritmalarında genelde önceden hazırlanmış bir veri kümesi kullanılmaz.
- Pekiştirmeli öğrenme, gerçekleşen bir durumun ödüllendirilmesi veya cezalandırılmasını temel alarak öğrenmeyi sağlayan bir algoritmadır. Bir hedefe ulaşmak için çevre ile etkileşimden öğrenmeyi gerçekleştirir.
- Eğitilen unsur, etmen (agent) olarak isimlendirilir etmenin eğitim sırasında verileri elde ettiği ve sürekli olarak ilişki halinde bulunduğu bir ortamı (environment) mevcuttur.



Pekiştirmeli öğrenme (Reinforcement learning)

- Pekiştirmeli öğrenme sırasında etmen ortam ile etkileşime girer ve ortamdan aldığı verilere göre eylem üretir.
- Ortam ise bu eylemin hedefe yakınlık derecesine bakarak etmeni ödüllendirir veya cezalandırır.
- Etmen tüm eylemleri dener ve karşılığında aldığı ödül puanları sayesinde hangi eylemlerin hedefe yönelik olduğunu deneme-yanılma yoluyla öğrenir.

Pekiştirmeli öğrenme (Reinforcement learning)



Pekiştirmeli öğrenme (Reinforcement learning)

- Pekiştirmeli öğrenme algoritmasının çalışma mantığı deneme-yanılma yöntemine dayalıdır ve deneme sayısı ile beraber etmenin performansı artar.
- En iyi sonuca ulaşmak için yeterli sayıda deneme işlemini gerçekleştirmek gerekir.
- Pekiştirmeli öğrenme yönteminde ideal sayılabilecek durum, deneme sayısının sonsuz olduğu durumdur bu pratikte mümkün olmayacağı için, algoritmanın kullanıldığı duruma özel olarak belirlenen yeterli sayıda deneme yapılır.
- Fakat, algoritmanın kullanıldığı alanlara ve deneme sayısına bağlı olarak da bazı sorunlar ortaya çıkabilmektedir (zaman ve enerji tüketiminin fazla olması vb.).