

Open in app ↗



Search

✦ **Jump-start your best year yet:** Become a member and get 25% off the first year

Adım Adım Makine Öğrenmesi Bölüm 3 : Denetimsiz Öğrenme Nedir?



Hatice Candan · Follow

Published in Machine Learning Türkiye

5 min read · Dec 26, 2021



Share



More

Machine Learning Step-by-Step Part 3: What is Unsupervised Learning?

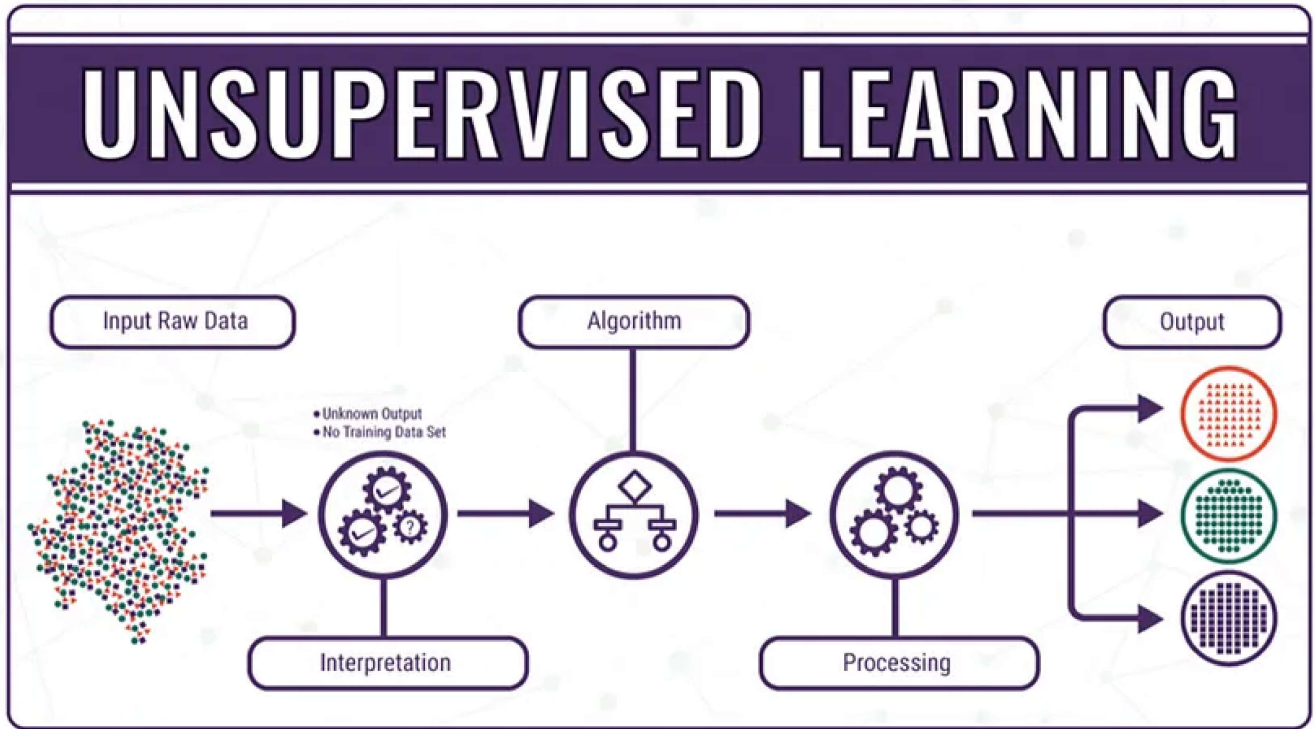
Photo by [Corinne Kutz](#) on Unsplash

Merhabalar efendim merhabalaarr, Adım Adım Makine Öğrenmesi serimizin üçüncü yazısı ile karşınızdayım. Serinin ikinci yazısında Denetimli Öğrenmeyi incelemiştik okumayanları şöyle alalım tık tık... Bu yazımda ise makine öğrenmesi türlerinden ikincisi olan Denetimsiz Öğrenmeyi inceliyoruz. Şimdiden hepinize keyifli ve verimli okumalar :)

Denetimsiz Öğrenme (Unsupervised Learning) Nedir?

Denetimsiz Öğrenme etiketsiz veriler arasındaki gizli kalmış yapıyı/örüntüyü bulmaya çalışan bir makine öğrenmesi yöntemidir. Burada önemli olan elimizdeki verilerin etiketi hakkında bir bilginin olmadığı durumlarda denetimsiz öğrenme kullanılmaktadır. Elimizde sadece veriler vardır ve elimizdeki bu verilerden sonuç çıkarılmaya çalışılmaktadır. Aslında her işi kendi kendine öğreniyor.

Veriler etiketsiz olduğu için de Denetimsiz Öğrenme algoritması çıktısına kesinlikle doğrudur gözüyle bakmamak gerekir.



Şekil 1 Denetimsiz Öğrenme Algoritması

Aslında gün içinde hepimizin sıkça denk geldiğini küçük bir örnek verelim. Diyelim ki online bir alışveriş sitesinde geziyorsunuz ve sepetinize bazı ürünler eklediniz. Sepete eklemiş olduğunuz bu ürünler göz önüne alınarak sitenin size alabileceğiniz başka ürünleri de tavsiye etmesi denetimsiz öğrenmenin bir örneğidir.

Denetimsiz öğrenmede veriler yorumlanarak aralarında bir ilişki bulunur ki buna da İlişkilendirme (Association) problemleri denir veya kümeleştirme işlemi Kümeleme (Clustering) yapılır. Zaten ilk yazımızda da denetimsiz öğrenmenin Kümeleme (Clustering) ve İlişkilendirme (Association) problemlerine uygulandığını söylemiştik.

Gelin İlişkilendirme (Association) ve Kümeleme (Clustering) problemlerine yakından bakalım.

1) İlişkilendirme (Association) :

*İlişkilendirme algoritması, verileriniz arasındaki birlikteliği ortaya çıkarır. Bu algoritma genelde **Birliktelik Kuralı Madenciliği (Association Rule Mining)** olarak bilinir. İlişkisel veritabanları, işlem veritabanları ve diğer depo biçimleri gibi çeşitli veritabanlarında bulunan veri kümelerinden sık sık meydana gelen kalıpları, bağıntıları veya ilişkileri gözlemlemeyi amaçlayan bir prosedürdür.*

Genellikle pazar sepeti analizi (market basket analysis) için yani hangi ürünler birlikte satın alınır, insanların birlikte ziyaret etme eğiliminde olduğu mağazalar/siteler, ürün çeşitliliği kararları, çapraz satış gibi alanlarda kullanıldığından alışveriş siteleri, mağazalar gibi satış yapılan alanlarda sıklıkla kullanılır.

Yukarıda vermiş olduğumuz örnekte olduğu gibi bir ürünü alan kullanıcıların yanında aldıkları ürünlerin diğer kullanıcılara da tavsiye edilmesi İlişkilendirme algoritmasının en yaygın kullanım alanıdır.



Şekil 2 Birliktelik Kuralı Madenciliğinin (Association Rule Mining) Pazar Sepeti Analizinde Kullanılması

1.1) Apriori algoritması

Birliktelik kuralları oluşturmak için kullanılan en eski algoritmalarından biri olan Apriori Algoritması, öğeler arasındaki bağımlılığı hesaplamaya dayanır. Kısacası “tüketiciler hangi ürünleri birlikte alıyor” sorusunun cevabını vermeye çalışmaktadır. Priori kelimesi önce anlamına gelmektedir ve kendinden önceki ürünlerin durumuna göre sonuçlar çıkarır ve bu sonuçları başka bir tüketici aynı ürünü aldığı anda ona tavsiye edilecek ürünleri gösterir.

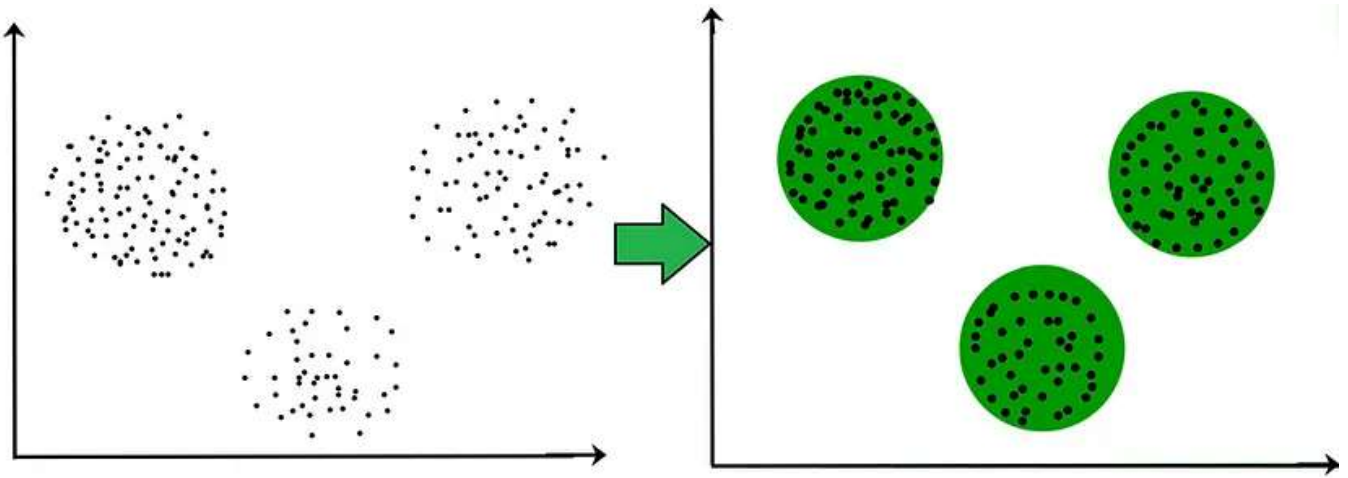
Diyelimki bir e-ticaret sayfanız var ve oradan çeşitli ürünler satmaktasınız. Bir müşteri sepetine filtre kahve eklediğinde Apriori algoritması o müşteriye çeşitli kahve kupaları tavsiye edecektir. Böylece satışlarınız artarken karınız da artacaktır.

2) Kümeleme (Clustering)

Kümelenme denetimsiz öğrenmede önemli bir kavramdır. Kümeleme, aynı gruplardaki veri noktalarının aynı gruptaki diğer veri noktalarına daha benzer ve diğer gruplardaki veri noktalarından farklı olacak şekilde birkaç gruba ayırma algoritmasıdır. Temel olarak, verileri aralarındaki benzerlik ve farklılık temelinde analiz ederek gruplara ayırır.

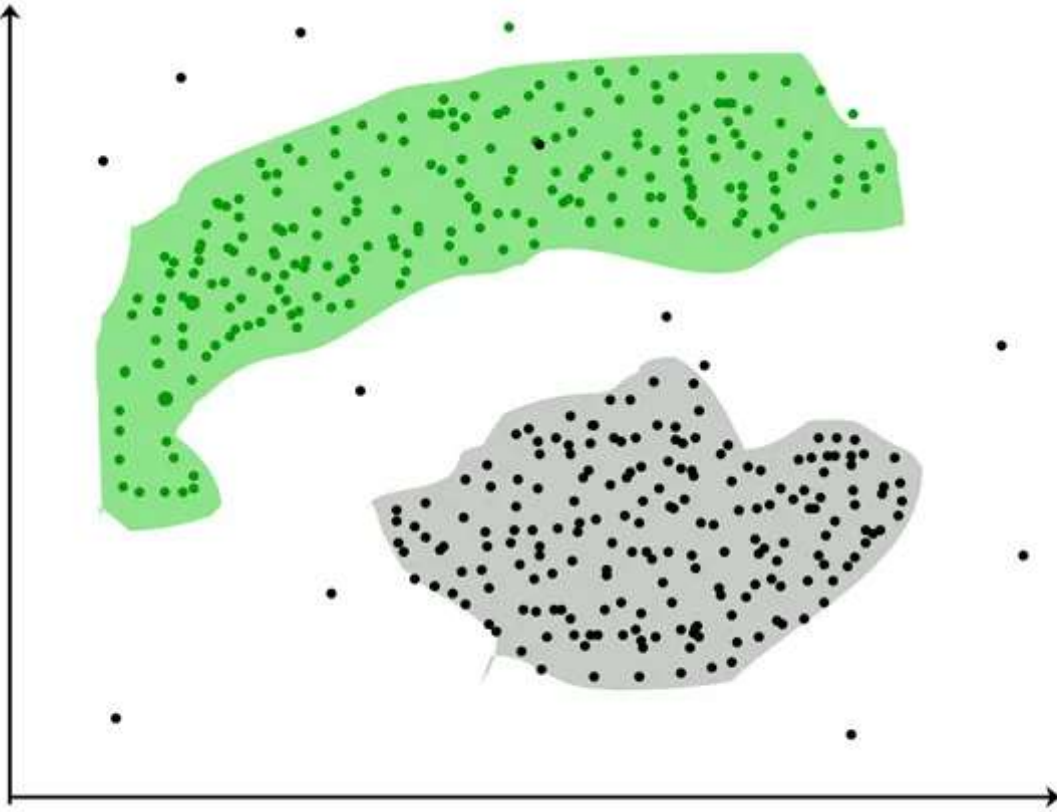
Kümeleme algoritmaları verilerinizi işler ve verilerde varsa doğal kümeleri (grupları) bulur. Ayrıca algoritmalarınızın kaç kümeyi tanımlaması gerektiğini ayarlayabilirsiniz böylece grupların ayrıntı düzeyini ayarlamış olursunuz.

Şekil 3'te yer alan grafikte kümelenmiş veri noktalarının 3 farklı küme şeklinde tespit edildiğini görüyorsunuz.



Şekil 3 Kümeleme (Clustering) örneği — 1

Kümeleme algoritması çıktısını genelde dairesel olarak hayal edebiliriz fakat kümelerin şekli dairesel olmak gibi bir zorunluluğu yoktur. Zaten düşününce de böylesi mükemmel bir veri seti bulmak gerçek hayatta pek de mümkün değildir. Kısaca yazar burada der ki “*Ey ahali kümeleme diyince aklınıza Şekil 4 de gelsin :)*”.



Şekil 4 Kümeleme (Clustering) örneği — 2

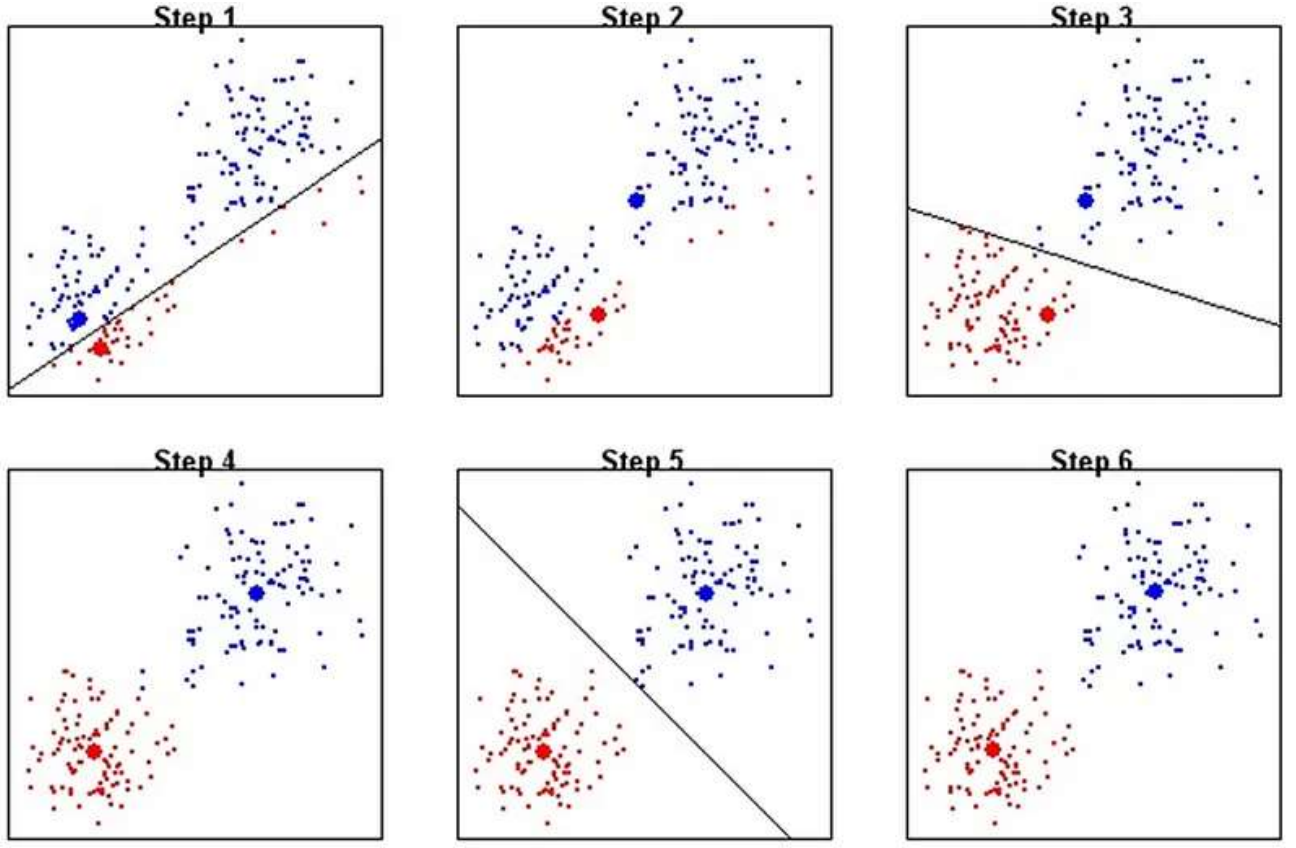
Kümeleme sadece dairesel olmayacağı gibi birden fazla da kümeleme algoritması bulunmaktadır. Şimdi sırasıyla bu algoritmaları inceleyeceğiz.

2.1) K-Means Kümeleme (K-means Clustering)

K-means kümeleme yinelemeli bir kümeleme algoritmasıdır. “K” değeri verileri kaç kümeye ayıracağımızı belirleyen bir faktördür. K değeri her yinelemde en yüksek değeri bulmaya yardım eder. Başlangıçta K değeri bizim tarafımızdan belirlenebilir. Sonrasında seçilen K değerine göre veriler K tane merkez noktada kümelenir. Veriler arasındaki benzerlik bulundukça diğer veriler de yakınındaki noktalara göre kümelenir ve veriler arasındaki benzerlik en verimli haline ulaşana kadar algoritma çalışmaya devam eder.

Birazcık uzun ve karmaşık gelmiş olabilir ama aslında çok eğlenceli.

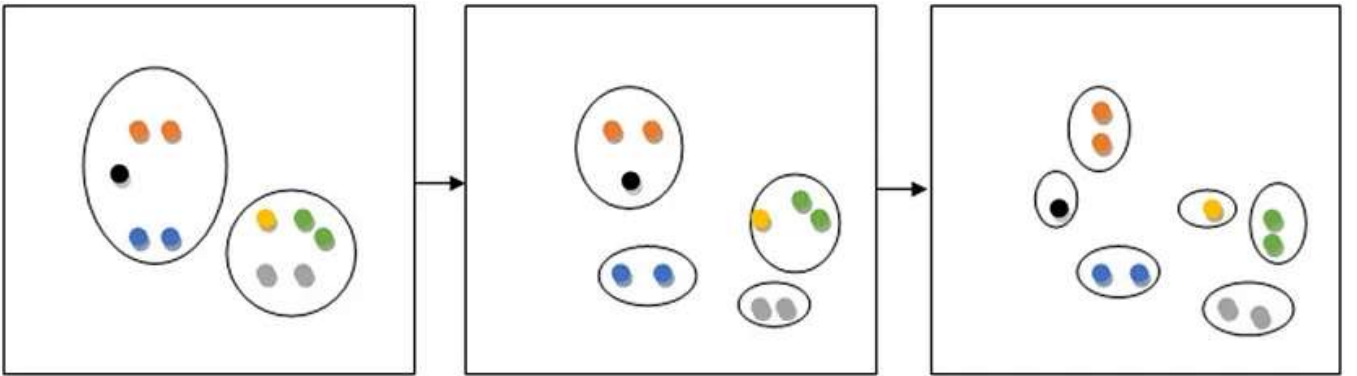
Şekil 5 üzerinden örnekle anlatmaya çalışalım. Başlangıç K=2 olsun. Step1’e bakıldığında elimizdeki verilerden 2 ortak nokta bulunmuş. Bu ortak noktaları mavi ve siyah renklerle görebilirsiniz. Ortak noktalar bulunduktan sonra diğer veriler birbirlerine yakınlığa göre yeniden kümelenmiş ve ortak noktaların yeri değişmiş (Step 2). Bu durum veriler arasındaki en verimli benzerlik yakalanana kadar devam etmiş (Step 6).



Şekil 5 K-Means Kümeleme Örneği

2.2) Hiyerarşik Kümeleme (Hierarchical Clustering)

Hiyerarşik kümeleme benzer nesneleri küme adı verilen gruplara ayıran bir algoritmadır. Bu algorithma veri kümesi gözden geçirilerek veriler arasındaki benzerliğe dayalı olarak benzer bulunan verileri gruplayarak kümeler oluşturulur. Her küme birbirinden farklıdır ve her kümedeki veriler büyük ölçüde birbirine benzer. Yani kümelerin her biri için bir hiyerarşi oluşturulur (Şekil 6).



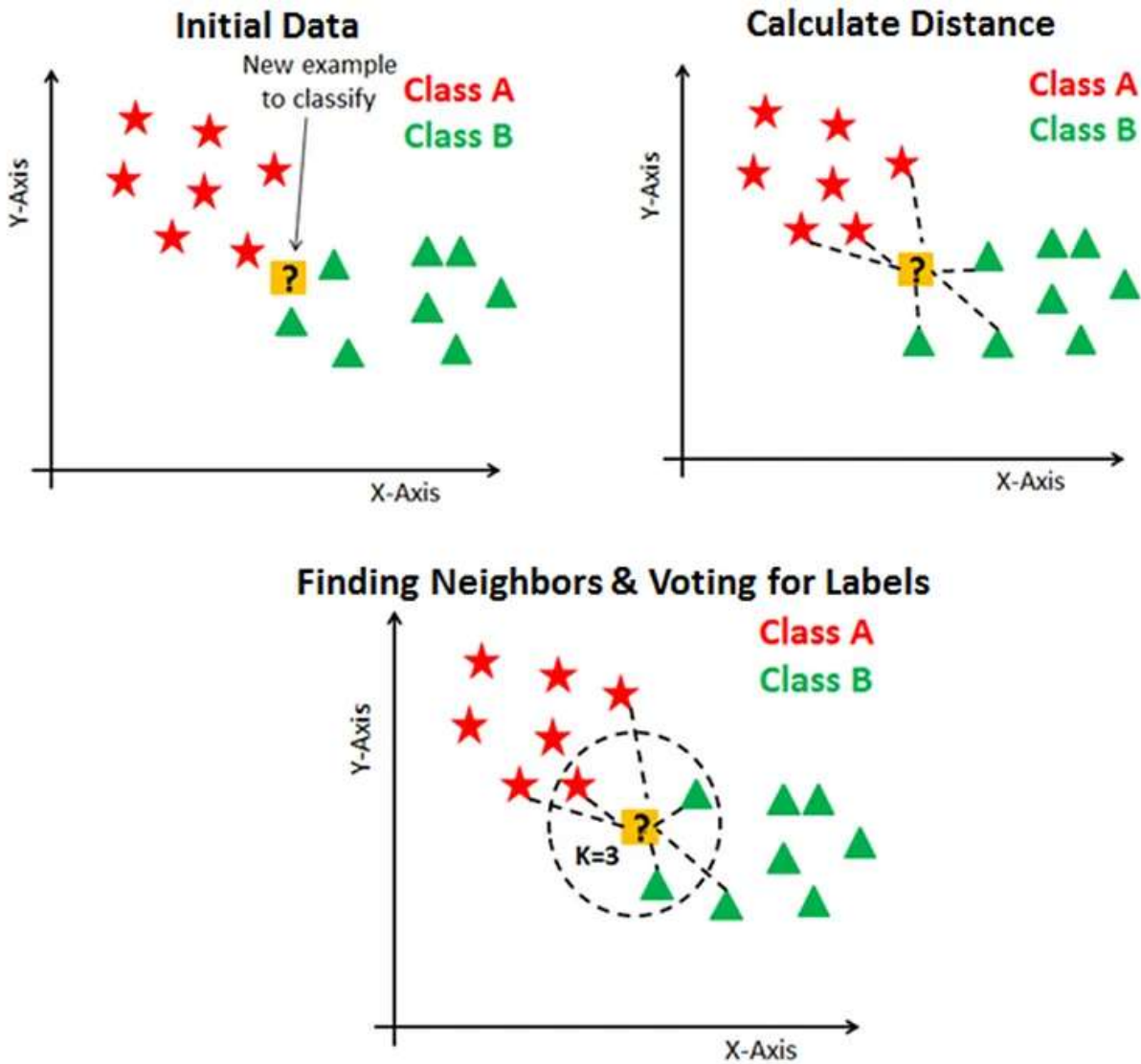
Şekil 6 Hiyerarşik Kümeleme (Hierarchical Clustering)

2.3) K-NN Kümeleme (K-Nearest Neighbors Clustering)

K-NN Makine Öğrenmesinde en sık kullanılan algoritmalarından birisi olabilir. K-NN algoritması önceden elde edilen veriler ile yenileri arasındaki mesafeye dayalı olarak sınıflandırma yapan bir algoritmadır. “K”, komşuluk sayısını temsil etmektedir. K sayısı kadar komşuluklarla olan mesafe incelenir ve bunun sonucunda veri, en olası sınıfa dahil olur.

Şekil 7 , K-NN’i anlamak açısından gayet açıklayıcı olacaktır.

Büyük veri kümelerinin öğrenme süresi zaman alacağından eğer elinizde daha küçük veri kümesi varsa K-NN daha iyi sonuç verecektir.



Şekil 7 K-NN Kümeleme (K-Nearest Neighbors Clustering)

Son Söz ve Kapanış

Denetimli öğrenmeden sonra denetimsiz öğrenmeyi de öğrendiğimizi düşünüyorum. Fazla detaya girmeden kısaca ve açıklayıcı şekilde anlatmaya çalıştım. Bir sonraki yazımda **denetimli öğrenme ve denetimsiz öğrenme**

arasındaki farklılıklara değiniyor olacağım. Okuduğunuz için teşekkür ederim.
Şimdilik hoşçakalın....

Kendinize cici bakın :)



Photo by [Jan Tinneberg](#) on Unsplash

Makine Öğrenmesi

Denetimsiz Öğrenme

Kümeleme

Association Rule Mining



Follow

Written by Hatice Candan

156 Followers · Writer for Machine Learning Türkiye

AI Engineer Intern at Mia Teknoloji || ex - Core Team Lead at Google Developer Student Clups ML and AI Team

More from Hatice Candan and Machine Learning Türkiye

 Hatice Candan in Machine Learning Türkiye

Adım Adım Makine Öğrenmesi Bölüm 2: Denetimli Öğrenme Nedir?

Machine Learning Step-by-Step Part 2: What is Supervised Learning?

4 min read · Dec 18, 2021





Merve KÜÇÜKDOĞRU in Machine Learning Türkiye

SQL Server Tarih Fonksiyonları

1)GETDATE

4 min read · Dec 18, 2020



--



ABDULLAH ATCILI in Machine Learning Türkiye

Clustering—Kumeleme

Clustering nedir?

11 min read · Feb 23, 2022



--



...

 Hatice Candan in Machine Learning Türkiye

Adım Adım Makine Öğrenmesi Bölüm 4: Denetimli Öğrenme ve Denetimsiz Öğrenme Arasındaki Fark

Machine-Learning Step-by-Step Part 4: Differences Between Supervised and Unsupervised Learning

3 min read · Jan 8, 2022



--



1



...

See all from Hatice Candan

See all from Machine Learning Türkiye

Recommended from Medium

 James Presbitero Jr. in Practice in Public

These Words Make it Obvious That Your Text is Written By AI

These 7 words are painfully obvious. They make me cringe. They will make your reader cringe.

4 min read · Jan 1

 --  646



Unbecoming

10 Seconds That Ended My 20 Year Marriage

It's August in Northern Virginia, hot and humid. I still haven't showered from my morning trail run. I'm wearing my stay-at-home mom...



• 4 min read • Feb 16, 2022



--



1046



Lists

Staff Picks

563 stories • 662 saves

Stories to Help You Level-Up at Work

19 stories • 430 saves

Self-Improvement 101

20 stories • 1242 saves

Productivity 101

20 stories • 1140 saves



Gowtham Oleti

Apps I Use And Why You Should Too.

Let's skip past the usual suspects like YouTube, WhatsApp and Instagram. I want to share with you some less familiar apps that have become...

10 min read · Nov 14, 2023



--



257



Ignacio de Gregorio in Towards AI

Is Mamba the End of ChatGPT As We Know It?

The Great New Question

✦ • 8 min read • Jan 11

👏 -- 💬 61

🔖 + ⋮

👤 Sheila Teo in Towards Data Science

How I Won Singapore's GPT-4 Prompt Engineering Competition

A deep dive into the strategies I learned for harnessing the power of Large Language Models (LLMs)

✦ • 24 min read • Dec 29, 2023

👏 -- 💬 119

🔖 + ⋮



Scott Galloway

2024 Predictions

Each year, we review/make predictions re the past/coming year. Most years, we hit more than we miss. But we do miss—if we made 10...

11 min read • Jan 6



--



138

[See more recommendations](#)