

Yapay Zekaya Giriş

HAFTA 4

Makine Öğrenmesine Giriş

Dr. Öğretim Üyesi Burcu ÇARKLI YAVUZ
bcarkli@sakarya.edu.tr

Makine öğrenmesi nedir?

- Geçmişten günümüze, makine öğrenmesi üzerine birçok tanım ve açıklama yapılmıştır.
- Zaman içinde büyük veri gibi kavramların hayatımıza girmesi ve yüksek teknolojinin (daha güçlü hesaplama ve veri saklama kapasitelerinin) ulaşılabilirliğinin artması, makine öğrenmesi tanımlarını da şekillendirmiştir.

THE INTERNET IN **2023** EVERY MINUTE



Created by: eDiscovery Today & LTMG

Makine öğrenmesi nedir?

- **Arthur Samuel (1959):** "Makine öğrenimi, bilgisayarların açıkça programlanmadan öğrenme yeteneğini kazanmaları olarak tanımlanabilir."
- **Christopher Bishop (2006):** "Makine öğrenimi, sistemlerin veri örnekleri üzerinden modeller oluşturarak öğrenmelerini sağlamayı amaçlar. Bu, sistemleri, önceden belirlenmiş kurallarla programlamak yerine, verilerle **'besleyerek' öğrenmelerini sağlar.**"
- **Kevin Murphy (2012):** "Makine öğrenimi, verilere (genellikle örneklerle) dayalı olarak karmaşık modellerin otomatik olarak fit edilmesi olarak görülebilir. Bu, istatistiksel modelleme ile büyük ölçüde örtüşmektedir, ancak geleneksel istatistiksel modellemeden daha büyük bir vurguyla, büyük ölçekli veri kümesi üzerinde otomatik (yani, insan müdahalesi olmadan) çalışır."

Makine öğrenmesi nedir?

- Makine öğrenmesi, bilgisayarların veriden, genellikle deneyimden öğrenmeye dayalı olarak performanslarını artırma kabiliyetidir.
- Başka bir deyişle, bir makine öğrenmesi modeli, belirli bir görevi daha iyi yapmak için *verilere dayanarak "öğrenir"* ve bu sayede programlamaya gerek kalmadan önceden belirsiz olan senaryolarda da doğru tahminlerde veya kararlarda bulunabilir.

Makine Öğrenmesi: Deneyimlerden öğrenme



Resim: Stanford Üniversitesi Makine Öğrenmesi ders notlarından alınmıştır

Makine Öğrenmesinin farklı disiplinler ile ilişkisi

➤ Makine öğrenmesi (ML), birçok disiplinin bir araya gelmesiyle oluşan çok disiplinli bir alandır. ML'nin temellerini oluşturan ve bu alana katkı sağlayan başlıca disiplinleri şunlardır:

1.Bilgisayar Bilimi:

- Algoritma tasarımı, veri yapıları ve programlama gibi konular, ML algoritmalarının geliştirilmesi ve uygulanması için temel oluşturur.
- Yeni optimizasyon teknikleri ve hesaplama yöntemleri, ML modellerinin eğitilmesi için gereklidir.

2.Matematik:

- Lineer cebir, ML'de kullanılan vektörler, matrisler ve dönüşümler için temeldir.
- Kalkülüs, gradyan inişi gibi optimizasyon yöntemlerinin arkasındaki temel matematiksel konseptleri sağlar.
- Fonksiyonlar teorisi, aktivasyon fonksiyonları ve dönüşümler için gereklidir.

Makine Öğrenmesinin farklı disiplinler ile ilişkisi

3. İstatistik:

- ML modellerinin temelinde çoğu zaman istatistiksel yöntemler vardır.
 - Olasılık teorisi, tahmin etme ve genelleme yapma yeteneği kazandırmak için kullanılır.
 - İstatistiksel testler ve doğrulama yöntemleri, modellerin performansının ölçülmesinde kullanılır.
- Son yıllarda, istatistik ve makine öğrenmesi arasındaki sınırlar giderek daha belirsiz hale gelmektedir.
- Birçok modern veri bilimi uygulaması, istatistiksel yöntemleri ve makine öğrenmesi algoritmalarını bir araya getirerek daha iyi sonuçlar elde etmektedir.

Makine Öğrenmesinin farklı disiplinler ile ilişkisi

Makine öğrenmesi & İstatistik benzerlikleri:

- ❖ Hem ML hem de istatistik, verilere dayalı tahminlerde bulunma ve genelleme yapma yeteneğine sahiptir. Regresyon, sınıflandırma ve kümeleme gibi birçok makine öğrenmesi yöntemi istatistiksel kavramlara dayanmaktadır.
- ❖ İki disiplinde de optimizasyon önemli bir rol oynamaktadır.



Makine öğrenmesi & İstatistik farkları:

- ❖ İstatistik, verilere dayalı çıkarımların ve modellerin temelini oluştururken, makine öğrenmesi bu modelleri geliştirme ve tahmin için kullanır.
- ❖ İstatistiksel modeller genellikle daha basit ve yorumlanabilir olma eğilimindedir. Makine öğrenmesi ise, özellikle derin öğrenme gibi yöntemlerle, daha karmaşık ve bazen yorumlanması zor modellere izin verir.
- ❖ İstatistiksel analizler genellikle bir varsayım altında (örneğin, hataların normal dağıldığı varsayımıyla) çalışır. Makine öğrenmesi, özellikle bazı yöntemlerde, bu tür varsayımlardan daha esnektir.

Makine Öğrenmesinin farklı disiplinler ile ilişkisi

4. Nörobilim:

- Derin öğrenme ve yapay sinir ağları, biyolojik sinir ağlarının basit modelleri olarak düşünülebilir.
- Nörobilim, makine öğrenimi modellerinin nasıl daha etkili ve biyolojik olarak ilham alınabileceği konusunda fikirler sunar.

5. Bilişsel Bilim:

- İnsan öğrenme süreçleri, makine öğrenimi algoritmalarının tasarımında ilham kaynağı olabilir.
- Bilişsel bilim, makine öğreniminin karmaşıklığını ve kısıtlamalarını daha iyi anlamamıza yardımcı olabilir.

6. Optimizasyon:

- Optimizasyon teorisi, ML modellerinin eğitim sürecinde kullanılan yöntemlerin temelini oluşturur.
- ML'de, modellerin performansını en iyi hale getirmek için birçok optimizasyon yöntemi kullanılır.

Makine Öğrenmesinin farklı disiplinler ile ilişkisi

7. Mühendislik:

- Yazılım mühendisliği, makine öğrenimi algoritmalarının gerçek dünya uygulamalarında etkili bir şekilde uygulanmasını sağlar.
- Elektrik ve elektronik mühendisliği, sinyal işleme ve kontrol teorisi konularında bilgi sağlayarak makine öğrenimi algoritmalarının geliştirilmesine katkıda bulunabilir.

8. Ekonomi ve Oyun Teorisi:

- Öğrenme algoritmalarının karar alma ve strateji geliştirme konularında nasıl kullanılacağına dair fikirler sağlar.

➤ Bu disiplinlerin dışında, makine öğrenimi birçok diğer alandan da ilham almaktadır ve bu alandaki yeniliklerin önemli bir kısmı farklı disiplinler arasındaki işbirliğinden doğmaktadır.

Veri madenciliği-Makine öğrenmesi ilişkisi

- **Veri madenciliği**, büyük veri kümelerindeki örüntüleri, ilişkileri ve düzenlilikleri keşfetmek için kullanılan süreçler ve yöntemler bütünüdür.
- Ana hedefi, veri içindeki bilgiyi ortaya çıkarmaktır. Bu, sınıflandırma, kümeleme, ilişkilendirme kuralları ve sekans analizi gibi çeşitli yöntemlerle yapılır.
- Veri madenciliği, genellikle veri içindeki örüntüleri ve ilişkileri **keşfetmeye** odaklanırken, makine öğrenmesi bu örüntüleri ve ilişkileri kullanarak **tahminlerde bulunmaya** odaklanır.
- *Veri madenciliği süreçlerinde sıklıkla makine öğrenmesi algoritmaları kullanılır.* Örneğin, bir veri kümesindeki müşteri segmentasyonunu belirlemek için kümeleme algoritmaları ya da bir müşterinin bir ürünü satın alıp almayacağını tahmin etmek için sınıflandırma algoritmaları kullanılabilir.
- Hem veri madenciliği hem de makine öğrenmesi için **veri ön işleme adımları** kritik öneme sahiptir. Verinin temizlenmesi, dönüştürülmesi ve özellik mühendisliği, her iki alanda da yaygın olarak gerçekleştirilen adımlardır.

Makine öğrenmesi aşamaları

- 1.Problem Tanımı**
- 2.Verİ Toplama**
- 3.Verİ Ön İşleme**
- 4.Özellik Mühendisliği, Seçimi ve Çıkarımı**
- 5.Model Seçimi**
- 6.Model Eğitimi**
- 7.Model Değerlendirmesi**
- 8.Hiperparametre Ayarı**
- 9.Modelin Üretim Ortamına Alınması**
- 10.İzleme ve Bakım**

Makine öğrenmesi aşamaları



Makine öğrenmesi aşamaları

1. Problem Tanımı: Başlamadan önce neyi çözmeye çalıştığınızı net bir şekilde tanımlamak önemlidir.

Regresyon, sınıflandırma, kümeleme gibi spesifik bir makine öğrenmesi görevi ile karşılaşabiliriz.

Bir e-ticaret sitesi için ürün önerisi yapmak istiyorsanız bu bir "öneri" problemi olabilir.

2. Veri Toplama: Modeli eğitmek için ihtiyaç duyulan veriyi toplama sürecidir. Bu, mevcut veri kaynaklarından veri çekmek veya yeni veri toplamak anlamına gelebilir.

Modelinizin eğitilmesi için yeterli miktarda ve kalitede veriye ihtiyacınız vardır.

E-ticaret sitesi örneği için kullanıcıların hangi ürünleri satın aldığına dair geçmiş veri gerekli olabilir.

Makine öğrenmesi aşamaları

- 3. Veri Ön İşleme:** Toplanan verinin genellikle ham bir formda olması nedeniyle, modeli eğitmeye hazır hale getirmek için çeşitli ön işleme adımları gereklidir.

Bu adımlar arasında eksik veri doldurma, aykırı değerlerin ele alınması, ölçeklendirme, kategorik değişkenlerin sayısallaştırılması ve daha fazlası bulunabilir.

Kullanılabilecek bazı Yöntemler:

- **Eksik Veri İşleme:** Ortalama, medyan doldurma veya ileri/geri doldurma gibi yöntemlerle.
- **Aykırı Değer Tespiti:** IQR, Z-Skoru veya izolasyon ormanı gibi yöntemlerle.
- **Veri Normalizasyonu ve Standartlaştırma:** Min-Max ölçekleme, Z-skoru normalizasyonu.
- **Kategorik Veri Dönüşümü:** One-hot kodlama, etiket kodlama.

Makine öğrenmesi aşamaları

- 4. Özellik Mühendisliği, Seçimi ve Çıkarımı:** Veriden, modelin eğitimi için yararlı olabilecek yeni özelliklerin türetilmesi sürecidir.

Veriden yeni özellikler türetmek veya mevcut özellikleri dönüştürmek modelin performansını artırabilir.

Yöntemler:

- **Özellik Türetme:** Mevcut özelliklerden yeni özellikler oluşturma.
- **Özellik Seçimi:** En önemli özellikleri seçme. RFE (Recursive Feature Elimination), özellik önemi, korelasyon analizi.
- **Boyut Azaltma:** PCA (Principal Component Analysis), LDA (Linear Discriminant Analysis).

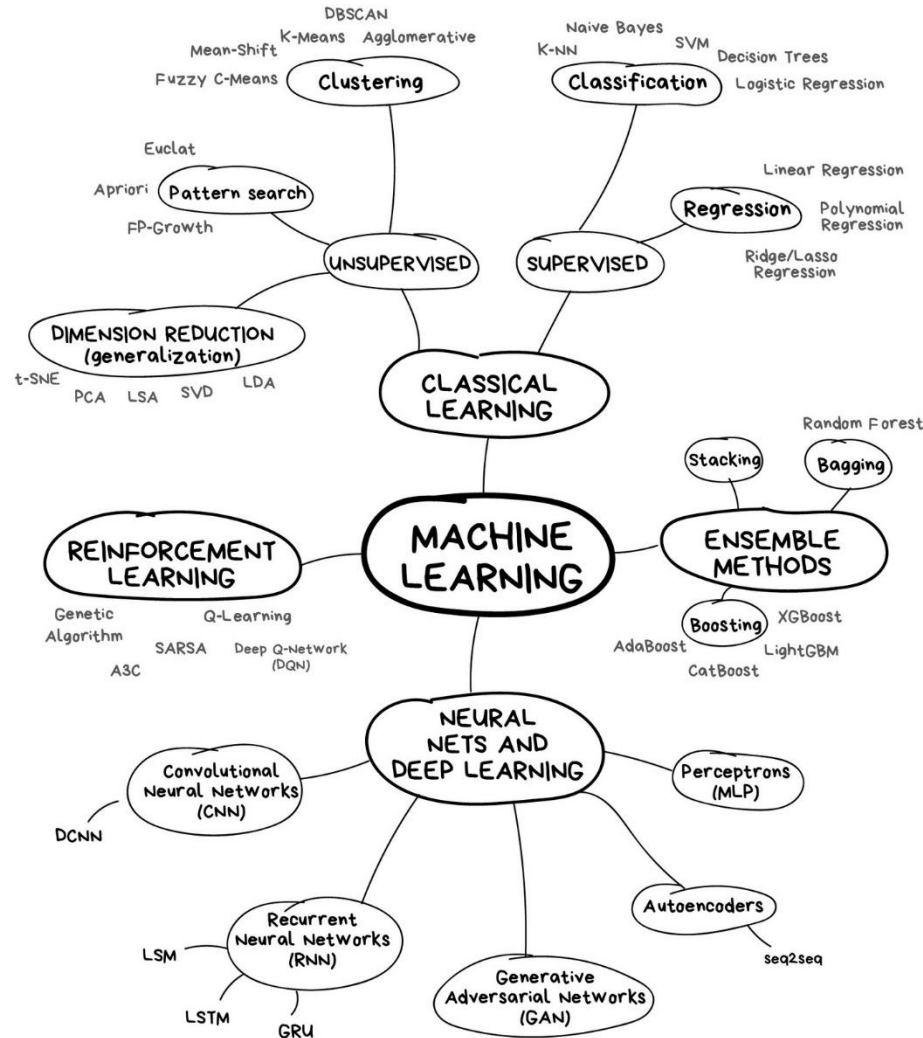
Makine öğrenmesi aşamaları

- 5. Model Seçimi:** Uygun algoritmanın veya algoritmaların seçilmesi. Kullanılabilecek birçok farklı makine öğrenmesi algoritması vardır ve problem tipine göre en uygun olanını seçmek önemlidir.

Yöntemler:

- **Sınıflandırma Algoritmaları:** Lojistik Regresyon, Karar Ağaçları, Rastgele Orman, Destek Vektör Makineleri.
- **Regresyon Algoritmaları:** Lineer Regresyon, Polinom Regresyon, Ridge/Lasso Regresyon.
- **Kümeleme Algoritmaları:** K-Means, DBSCAN, Hiyerarşik Kümeleme.

Makine öğrenmesi türleri



Kaynak: <https://dltr.asmaamir.com/faydali-gorseller>

Makine öğrenmesi aşamaları

6. Model Eğitimi: Seçilen algoritmayı, toplanan veriyle eğitme sürecidir.

Yöntemler: Veri setini eğitim ve test setine ayırma, çapraz doğrulama (cross-validation).

Makine öğrenmesi aşamaları

7. **Model Değerlendirmesi:** Eğitilen modelin performansını değerlendirme sürecidir.

Bu, doğruluk, hassasiyet, hatırlama, F1 puanı gibi metriklerle gerçekleştirilebilir. Aynı zamanda, modelin genelleştirme kabiliyetini test etmek için ayrı bir test veri kümesi kullanılır.

Yöntemler:

- **Sınıflandırma İçin:** Doğruluk (accuracy), F1-Skoru, ROC Eğrisi, AUC.
- **Regresyon İçin:** Ortalama Kare Hatası (MSE), Ortalama Mutlak Hata (MAE), R^2 skoru.
- **Kümeleme İçin:** Siluet Skoru, Davies-Bouldin İndeksi.

Makine öğrenmesi aşamaları

8. Hiperparametre Ayarı (Model optimizasyonu): Modelin performansını optimize etmek için modelin hiperparametrelerinin ayarlanması sürecidir.

Bir Karar Ağacı modelinde ağacın derinliği, bir Derin Öğrenme modelinde öğrenme hızı gibi parametrelerin ayarlanması.

Yöntemler: Izgara arama (Grid Search), rastgele arama (Random Search), Bayes optimizasyonu.

9. Modelin Üretim Ortamına Alınması: Eğitilen ve optimize edilen modelin gerçek dünyada kullanılmak üzere bir uygulama veya hizmete entegre edilmesi sürecidir.

10. İzleme ve Bakım: Modelin üretimdeki performansının sürekli izlenmesi ve gerektiğinde modelin güncellenmesi veya yeniden eğitilmesi.

Makine öğrenmesi nerelerde kullanılır?

- Makine öğrenmesi, geniş bir uygulama yelpazesine sahip bir teknolojidir ve birçok sektörde farklı amaçlarla kullanılmaktadır.
- Makine öğrenmesinin kullanıldığı bazı alanlar ve örnek uygulamalar:
 - **Sağlık ve Tıp:**
 1. Hastalık teşhisi: Tıbbi görüntülerin ya da kan değerlerinin analiziyle farklı hastalıkların teşhis edilmesi.
 2. İlaç keşfi: Yeni ilaç moleküllerinin hızla tanımlanması ve test edilmesi.
 - **Finans:**
 1. Algoritmali ticaret: Piyasa verilerini analiz ederek otomatik alım-satım kararları almak.
 2. Kredi skorlama: Müşterilerin kredi riskini değerlendirme.
 - **Perakende ve E-ticaret:**
 1. Ürün önerileri: Kullanıcının geçmiş alışveriş deneyimlerine dayanarak ürün önerisi yapma.
 2. Talep tahmini: Stok yönetimi için gelecekteki ürün talebini tahmin etme.

Makine öğrenmesi nerelerde kullanılır?

- **Eğlence ve Medya:**

1. Kişiselleştirilmiş içerik önerileri: Netflix veya Spotify gibi platformlarda kullanıcıya içerik önerisi.
2. Oyunlarda yapay zeka: Oyun karakterlerinin otomatik olarak hareket etmesi için.

- **Otomotiv:**

1. Otonom araçlar: Araçların çevreyi algılaması ve kararlar alması.
2. Sürücü yardımcı sistemleri: Sürücüye yardımcı olacak şekilde tehlikeleri algılama.

- **Enerji:**

1. Talep tahmini: Enerji talebinin ne zaman artacağını veya azalacağını tahmin etme.
2. Optimal enerji dağıtımı: Kaynakların en etkili şekilde nasıl kullanılacağını belirleme.

Makine öğrenmesi nerelerde kullanılır?

- **Tarım:**

1. Hasat tahmini: Hasat zamanının ve miktarının tahmini.
2. Zararlı tespiti: Uydu görüntüleri veya dronlar aracılığıyla zararlıları tespit etme.

- **Emlak:**

1. Fiyat tahmini: Bir mülkün piyasa değerini tahmin etme.
2. Optimal konum belirleme: Yeni bir işletme veya konut için en iyi konumu bulma.

- **Eğitim:**

1. Öğrenci performansı tahmini: Öğrencilerin gelecekteki başarısını veya başarısızlığını tahmin etme.
2. Kişiselleştirilmiş öğrenme yolları: Her öğrencinin öğrenme stiline uygun içerik önerisi.

Makine öğrenmesi nerelerde kullanılır?

- **Güvenlik:**

1. Sahtecilik tespiti: Banka işlemleri sırasında sahtecilik aktivitelerini tespit etme.
2. Ağ güvenliği: Saldırıları ve güvenlik ihlallerini gerçek zamanlı olarak tespit etme.

- **Havacılık ve Uzay:**

1. Uçuş yolu optimizasyonu: En verimli ve güvenli uçuş rotalarını planlama.
2. Uydu görüntü analizi: Coğrafi değişiklikleri, doğal afetleri veya tarım alanlarını izleme.

- **İnsan Kaynakları:**

1. Yetenek avcılığı: Adayların CV'lerini analiz ederek en uygun adayları otomatik olarak belirleme.
2. İşten ayrılma tahmini: Hangi çalışanların işten ayrılma olasılığının yüksek olduğunu tahmin etme.

Makine öğrenmesi nerelerde kullanılır?

- **Sanat ve Tasarım:**

1. Otomatik içerik oluşturma: Müzik, metin veya görsel sanat üretimi.
2. Stil transferi: Bir resmin stilini başka bir resme aktarma.

- **Sosyal Medya ve Dijital Pazarlama:**

1. Duygu analizi: Kullanıcı yorumlarını analiz ederek marka veya ürün hakkında genel duygu durumunu belirleme.
2. Hedefleme reklamcılığı: Kullanıcıların ilgi alanlarına göre reklam gösterimi.

- **Akıllı Evler ve IoT (Nesnelerin İnterneti):**

1. Enerji tasarrufu: Evde enerji kullanımını optimize eden akıllı termostatlar.
2. Güvenlik izleme: Hareket sensörleri ve kameralarla anormal aktiviteleri tespit etme.

Makine öğrenmesi nerelerde kullanılır?

- **Telekomünikasyon:**

1. Ağ trafiği optimizasyonu: Veri trafiğini analiz ederek ağ performansını optimize etme.
2. Müşteri hizmetleri otomasyonu: Otomatik sesli yanıt sistemleri ve chatbotlar.

- **Spor ve Eğlence:**

1. Oyun içi strateji optimizasyonu: Takımların oyun stratejilerini analiz ve optimizasyon.
2. Fiziksel performans analizi: Sporcuların hareketlerini analiz ederek performans ve yaralanma riskini değerlendirme.

- **Lojistik ve Tedarik Zinciri:**

1. Rota planlama: En verimli teslimat rotalarını belirleme.
2. Stok optimizasyonu: Stok seviyelerini ve talebi tahmin ederek tedarik zinciri verimliliğini artırma.

Makine öğrenmesi nerelerde kullanılır?

- **Yasal ve Adli:**

1. Belgelerin otomatik incelemesi: Büyük hukuki belge havuzlarında bilgi arama.
2. Adli tıp analizi: DNA dizileri, parmak izleri veya diğer adli tıp verilerinin analizi.

- **Çevre Bilimi:**

1. Vahşi yaşam izleme: Hayvan hareketlerini ve göçlerini takip etme.
2. İklim değişikliği tahminleri: İklim değişikliğinin etkilerini modelleme ve tahmin etme.

- Bu örnekler, makine öğrenmesinin potansiyel uygulamalarının sadece bir kısmını temsil etmektedir.
- Teknoloji sürekli olarak gelişmekte olup, yeni kullanım alanları ve uygulamalar her geçen gün ortaya çıkmaktadır.

Önerilen Kaynaklar

- <https://ufukcolak.medium.com/makine-%C3%B6%C4%9Frenmesi-veri-%C3%B6n-i%CC%87%C5%9Fleme-5-58e1ce73c1fb>
- <https://miuul.com/not-defteri/ml101-makine-ogrenmesi-nedir>
- <http://muratsakal.com/?p=230>