

IKT484 Makine Öğrenmesi

1. Makine Öğrenmesi Kavramları (2/3)

Öğretim Görevlisi:

Bora GÜNGÖREN

bora.gungoren@atilim.edu.tr

Neden Makine Öğrenmesi

- Başka bir örnek: Müşteri sınıflandırması.
 - Bir bankayız olarak tüketici kredileri için çok sayıda talep alıyoruz.
 - Bu krediler genellikle küçük miktarlardadır, ancak bu kredileri alan birçok farklı kişi vardır.
 - İnsanları gruplara ayırmamız gerekir.

Neden Makine Öğrenmesi

- Başka bir örnek: Müşteri sınıflandırması.
 - İnsanları gruplara ayırmamız gerekir.
 - Belirlenen zamanda tüm taksitleri ödeyecek olanlar.
 - Bazı ödemeleri kaçırarak ancak sonunda telafi edecek olanlar.
 - Krediyi temerrüde düşürecek ancak borcu yeniden yapılandırabilecek olanlar.
 - Krediyi temerrüde düşürecek ve yeniden yapılandıramayacak olanlar.
 - Tarihsel veriler, birçok kişinin zaten bu gruplara atandığını gösteriyor.

Neden Makine Öğrenmesi

- Başka bir örnek: Müşteri sınıflandırması.
 - Sahip olduğumuz tarihsel verilere ve başvuru aşamasında sağlanan kişisel verilere dayanarak bir model tasarlamalıyız.
 - Başka bir deyişle, bize verilen tarihsel veri setine ve değerlendireceğimiz gelen verilere (hem içerik hem de desen olarak) dayanarak bir makine öğrenimi modeli tasarlanmalıdır.

Neden Makine Öğrenmesi

- Başka bir örnek: Müşteri sınıflandırması.
 - Makine öğrenmesi yaklaşımı tarihsel veri gerektirir.
 - Başvuran kişi hakkında özel veri.
 - “Benzer” kişiler hakkında ilişkilendirilebilir veri. (Benzerliği tanımlayın)
 - Genel nüfus hakkında daha az ilişkilendirilebilir veri.

Neden Makine Öğrenmesi

- Çoğu makine öğrenimi problemi bazı temel kategorilere girer:
 - (Yakın) gelecekteki davranışın tahmini
 - Yeni bir ögenin sınıflandırılması
 - Bir sistemi optimize etmek için parametre değerlerinin seçilmesi.
- Öncelikle problemin tahmin mi yoksa açıklama niteliğinde mi olduğunu anlamaya çalışırız.
 - Her ikisi de varsa, belki bir model yerine iki modele ihtiyacımız vardır.

Neden Makine Öğrenmesi

- Makine öğrenimi, matematiksel modeller oluştururken istatistiklerden gelen teorik temeli kullanır.
- Makine öğrenimi, bu modellerin modern, kolayca erişilebilen bilgi işlem kaynaklarında pratik kullanımını sağlamak için bilgisayar bilimlerinden gelen teorik temeli kullanır.

Neden Makine Öğrenmesi

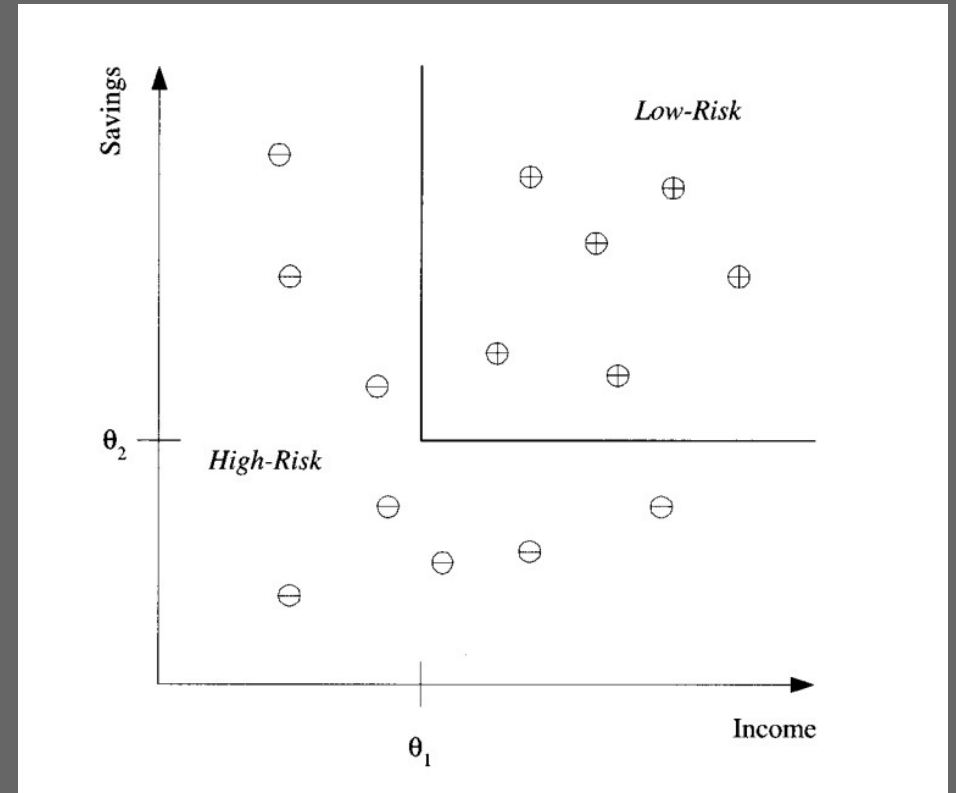
- İstatistikten teorik temel, tahmin doğruluğu gibi niteliklere odaklanır.
- Bilgisayar biliminden teorik temel, algoritmik karmaşıklık, paralel hesaplama ve bellek kullanımına odaklanır.
- Bilgisayar bilimi, hem modeli eğitmede hem de pratikte kullanmada önemlidir.
- Bazı modeller her zaman eğitindedir, bazıları değildir.

Tipik Makine Öğrenmesi Görevleri

- Öğrenme İlişkileri
 - Koşullu olasılık $P(Y | X)$ biçiminde bir ilişki kuralı bulma, burada Y , girdinin mevcut durumu olan X üzerinde koşullandırmak istediğimiz durumdur.
 - Tipik Örnek: $P(\text{süt satın alındı} | \text{yumurta satın alındı}) = 0,75$. O zaman birisi yumurta satın aldıysa süt satın alabileceğini varsayabiliriz. Ama bizden satın alacaklar mı?
 - Rafları daha yakın olacak şekilde yeniden düzenleyin.
 - Satın alınanları satın alanlar için bir ödül programı sağlayın.
 - Sepette yumurta olan müşterilere süt satın almalarını önerin.

Tipik Makine Öğrenmesi Görevleri

- Sınıflandırma
 - Eşikleri, kütle merkezlerini, mesafeleri, yarıçapları vb. belirleyerek fenomenleri (olaylar, insanlar vb.) sınıflandırmak için kategoriler tanımlayabiliriz.



Tipik Makine Öğrenmesi Görevleri

- Sınıflandırma
 - Ardından gelen bir öğeyi kolayca (yani zor hesaplama kaynakları olmadan) ve hızlı bir şekilde sınıflandırılacak şekilde değerlendirin.
 - Bu parametrelerin (eşik değerleri, kütle merkezleri vb.) belirlenmesi makine öğrenimi modelimizin eğitim aşamasıdır.
 - Model eğitildikten sonra onu birkaç ardışık soru basitliğiyle kullanırız.
 - Bu durumda değerleri olan sorulara ayırıcı denir ve sonuca tahmin denir.
 - Model ara sıra yeniden eğitilir. Neden?

Tipik Makine Öğrenmesi Görevleri

- Desen Tanıma (Kalıp Eşleme)
 - Desen tanıma, verilerdeki desenleri ve düzenlilikleri otomatik olarak tanımak için makine öğrenimi algoritmalarını kullanan bir veri analizi yöntemidir.
 - Desen tanıma sistemleri, tanıdık desenleri hızlı ve doğru bir şekilde tanımalıdır.

Tipik Makine Öğrenmesi Görevleri

- Desen Tanıma
 - Desenle ilgili veriler, metin ve resimlerden seslere veya diğer tanımlanabilir niteliklere kadar her şey olabilir.
 - Tarayıcılar ve OCR yazılımları görüntüleri kullanır
 - Yüz tanıma görüntüleri kullanır
 - Finansal dolandırıcılık tespiti finansal işlem verilerini kullanır
 - Tıbbi teşhis yardımcıları, laboratuvar sonuçlarından gelen tıbbi verilerin yanı sıra doktorlardan gelen metinsel girdileri kullanır
 - Konuşmadan metne yazılım, konuşmacıların ifadelerindeki ses desenlerini kullanır.

Tipik Makine Öğrenmesi Görevleri

- Bilgi Keşfi / Bilgi Çıkarımı
 - Yapılandırılmış veya yapılandırılmamış veri kümelerinden örtük, daha önce bilinmeyen ve potansiyel olarak yararlı bilgilerin önemsiz olmayan bir şekilde çıkarılması.
 - Çıkarma sonucu, yapılandırılmış bilginin oluşturulmasının veya ilişkisel bir şemaya dönüştürülmesinin ötesine geçer.

Tipik Makine Öğrenmesi Görevleri

- Bilgi Keşfi / Bilgi Çıkarımı
 - Örnek. Adlandırılmış varlık tanıma (NER), yapılandırılmamış metinde belirtilen adlandırılmış varlıkları kişi adları, kuruluşlar, konumlar, tıbbi kodlar, zaman ifadeleri, miktarlar, parasal değerler, yüzdeler vb. gibi önceden tanımlanmış kategorilere yerleştirmeyi ve sınıflandırmayı amaçlar.

Tipik Makine Öğrenmesi Görevleri

- Bilgi Keşfi / Bilgi Çıkarımı
 - Örnek. Moleküler madencilik. Moleküller moleküler grafiklerle temsil edilebildiğinden, bu grafikler belirli bir etki için potansiyel olarak yararlı molekülleri aramak için kullanılabilir. Moleküler madenciliğin varyasyonları ilaç keşfinde kullanılır.

Tipik Makine Öğrenmesi Görevleri

- Aykırı Değer Tespiti / Anomali Tespiti
 - Kurallara uymayan ve istisna olarak değerlendirilmesi gereken öğeleri bulma.
 - İstisnalar birçok şey ifade edebilir.
 - Yüksek performans gösteren çalışanlar gibi olağanüstü iyi örnekler.
 - Aşırı borçlu hesaplar gibi olağanüstü kötü örnekler.

Tipik Makine Öğrenmesi Görevleri

- Regresyon
 - Çoğu tahmin problemi, kullanılan matematiksel teknik ne olursa olsun, regresyon olarak adlandırılır.
 - Genel olarak, hareketli bir ortalama bile bir regresyon modelidir.
 - Regresyon, optimizasyon problemlerinde de kullanılır.
 - Örnek. Çamaşır makinesi çevrim süresi optimizasyonu.
 - Regresyon, birden fazla modeli birleştirerek dengelerini optimize etmeye çalışmak için de kullanılır.

Tipik Makine Öğrenmesi Görevleri

- Gözetimli Öğrenme
 - Kümeleme ve Regresyon tabanlı modeller gözetimli öğrenme modelleridir.
 - Amaç, bir denetleyicinin yardımıyla bir girdiden bir çıktıya eşlemeyi (veya işlevi) keşfetmek/öğrenmek/optimize etmektir.
 - Gözetmen bize keşif sürecimizin başarı seviyesini doğru/yanlış değeri veya gerçeğe olan bir tür mesafe olarak söyler.
 - Bu modda, modelimizi gözetimle eğitiyoruz.
 - Bu tür eğitim genellikle geçmişe dayalı ve iyi hazırlanmış bir eğitim veri kümesiyle yapılır.

Tipik Makine Öğrenmesi Görevleri

- Gözetimli Öğrenme
 - Gözetimli öğrenmenin genel modeli $y = g(x | \Theta)$ biçimindedir; burada y sonuç, x girdiler ve g modeldir. Θ model parametrelerini temsil eder.
 - Sınıflandırma modellerinde y , belirli bir sınıflandırmaya atamayı gösteren ayrı bir sonuçtur.
 - Tahmin modellerinde y , tahmin edilen değerdir.
 - Makine öğrenimi, hataların eğitim veri kümesine (sonuçları, y değerlerini bildiğimiz yer) göre en aza indirilmesi için model parametrelerini optimize eder.

Tipik Makine Öğrenmesi Görevleri

- Gözetimli Öğrenme
 - Parametre optimizasyonu genellikle bir noktada bir tür regresyon modeline dayandığından, çoğu gözetimli öğrenme görevi regresyon kullanır.
 - Bu nedenle yaygın bir söz vardır: "makine öğreniminin yüzde ... regresyondur".

Tipik Makine Öğrenmesi Görevleri

- Gözetimsiz Öğrenme
 - Bir gözetmen olmadığında, doğruyu yanlıştan ayırmanın bir yolu yoktur ve genellikle doğru veya yanlışın bir tanımı yoktur.
 - Biz sadece girdi veri kümesindeki yapıyı veya kalıpları anlamak istiyoruz.
 - Bu, gözetimsiz öğrenmenin bir örneğidir.
 - Kümeleme algoritmaları gözetimsiz öğrenmenin tipik örnekleridir.

Tipik Makine Öğrenmesi Görevleri

- Güçlendirilmiş Öğrenme
 - Geçmiş yürütmenin çıktılarının sisteme şimdilik girdi olarak verildiği bir sistemle uğraştığımızda, sistem kendini güçlendiriyor demektir.
 - Önceki fiyat belirleme örneği, güçlendirici bir sistemin böyle bir örneğidir.
 - Fiyatları rakiplerin fiyatlarına göre yükseltiriz.
 - Rakipler, bizim yükselttiğimiz fiyatı girdi olarak kullanarak benzer (veya aynı) bir model kullanır ve fiyatları yükseltir.
 - Daha sonra rakipler de yükselttiği için biz de fiyatları yükseltiriz.
 - Bu, müşterilerimiz hiç satın almamaya karar verene kadar devam eder.

Tipik Makine Öğrenmesi Görevleri

- Güçlendirilmiş Öğrenme
 - Güçlendirilmiş öğrenme birçok durumda oyun teorisiyle ilişkilidir, fark şu ki oyuncuların kesin karar modelini bilmemize gerek yoktur.
 - Ortak olarak paylaşılan tek teori, bir dizi eylemin (oyun sırasındaki dönüşler) olduğu ve iyi eylemler dizisinin genellikle tek bir iyi eylemden daha önemli olduğudur.

Tipik Makine Öğrenmesi Görevleri

- Güçlendirilmiş Öğrenme
 - Güçlendirilmiş öğrenme kolay değildir çünkü oyun oynamayı içeren sistemler yalnızca kısmen gözlemlenebilir ve kararlar eldeki verilere dayanmalıdır.
 - Ajanlar bir dereceye kadar işbirliği yaptığında, güçlendirme öğrenimi tasarlamak daha kolaydır.
 - Yönlendiriciler arasındaki etkileşimle ağ yönlendirmesi için makine öğrenimi durumu, ajan işbirliğinin böyle bir örneğidir. Yönlendiriciler, sistemin genel durumu hakkında bilgilerini artırmak için kasıtlı olarak durumları sorabilir veya durumları duyurabilir.

Makine Öğrenmesinin İstatistiksel Anlayışı

- Terminolojideki Değişim
 - Belirli deneyimlerden genel açıklamalar oluşturmak istatistiksel çıkarım olarak açıklanır.
 - Öğrenme, değişen ortama uyum sağlama ve yanıt verme ile ilgilidir. Bu, modeli belirli girdilerle çalıştırmayı ve bir sonuç elde etmeyi içerir (genellikle bir regresyon modeli içerir). Bu nedenle öğrenme genellikle tahmin olarak açıklanır.
 - Sınıflandırmaya ayırıcı analiz denir.

Makine Öğrenmesinin İstatistiksel Anlayışı

- Terminolojideki Değişiklik
 - Çoğu sinir ağı modelinin teorik temeli istatistiktir.
 - Geri yayılım algoritmasının versiyonlarına dayanırlar.
 - Ve genellikle sinir ağlarını tahmin edici veya sınıflandırıcı olarak kullanırız.
- Şunu unutmayın:
 - Bir tahmin edici, regresyon algoritmasından bulunan bir tahmin edicidir.
 - Bir sınıflandırıcı, bir sınıflandırma algoritmasından bulunan bir tahmin edicidir.

Makine Öğrenmesinin İstatistiksel Anlayışı

- Terminolojide Değişiklik
 - Çoğu sinir ağı modeli doğru bir şekilde parametrik olmayan regresyon modelleri olarak sınıflandırılır.
 - Parametrik regresyon modellerine göre avantajları olabilir.
 - Ancak bunların da birkaç uygulama sorunu vardır.
 - Bu sorunların üstesinden gelmek için çok fazla araştırma yapılmış ve teknikler geliştirilmiştir.
 - Bu derste yalnızca sinir ağlarına odaklanmasak da, bu tekniklerden bazıları dersimizin konusu içinde yer alacaktır.

Sorular?

İLETİŞİM:

bora.gungoren@atilim.edu.tr

License: Creative Commons Attribution Non-Commercial Share Alike 4.0 International (CC BY-NC-SA 4.0)