

Makine Öğrenmesi

Makine öğrenimi tarihte ilk kez 1959 senesinde “Arthur Lee Samuel” tarafından IBM için geliştirilen dama oyununu tasarlarlarken kullanılmış bir terimdir.

1962 senesinde Robert Nealey, IBM tarafından üretilen IBM7094 bilgisayarına karşı dama oyununu kaybetti. Dama oyunu için tasarlanan program çok az miktarda bilgisayar belleğine sahip olduğundan tahtadaki taşların yapacağı pozisyonlar için bir puanlama işlevi tasarlanmıştı. Puanlama işlevinin amacı iki tarafında oyunda kazanma şansını ölçmeye çalışmaktı. Program sonraki hamlelerini seçerken “minimax algoritması” nı kullanarak oyun stratejisini kullanmıştı.

Oyunun daha iyi olması için Samuel sadece bununla yetinmeyip, bir dizi mekanizma tasarladı. İsmi “ezberci öğrenme” olarak adlandırdığı öğrenme metoduyla, programa bütün pozisyonları kaydetti ve hatırlamasını sağladı. Bunu ödül fonksiyonunun değerleriyle birleştirdi. Bu çalışması sayesinde Samuel tarafından kullanılmaya başlanan “Makine Öğrenmesi” terimi literatürde kendine yer bulmaya başladı.

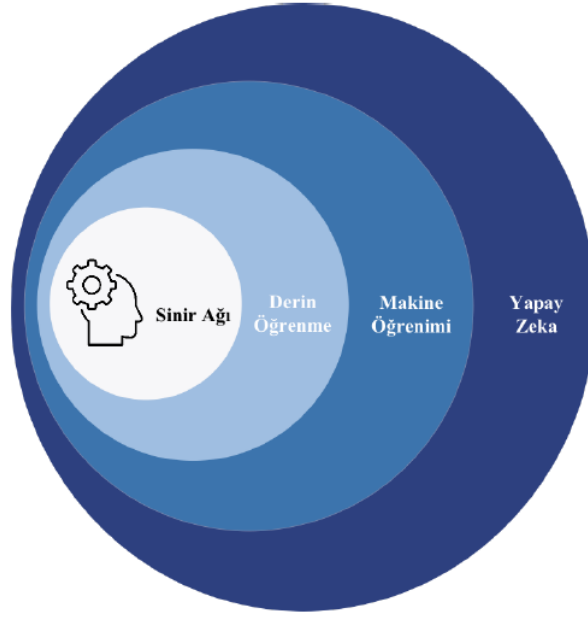
Kaybedilen oyunu günümüz teknolojik imkânları dahilinde değerlendirdiğimizde, bilgisayar programının insan zekâsına karşı olan en basit zaferi olarak adlandırabiliriz. Bizim için basit görünen bu zafer ise geleceğin yapay zekâsına temel oluşturmuştur.

Makine öğrenmesi; verilerin ayrıştırılması için algoritma kullanılması, onu öğrenme ve sonra dünyadaki bir şey hakkında bir belirleme yapılması ya da tahmin yapma uygulaması olarak ifade edilebilmektedir. Makine öğrenmesinde; manuel olarak programlamak yerine model mevcut bir veri seti ile eğitilir ve ardından yeni veriler üzerinde öğrenilmiş görevler yerine getirilir. Verilerden öğrenmeyi etkin bir şekilde gerçekleştirebilmek için örnek vakaların ve ilgili girdi parametrelerinin toplanması gerekmektedir. Makine öğrenmesinin bu özelliği modeli eğitmek için kullanılacak verinin yeterli boyut ve kalitede olmasını gerektirir.

Makine öğrenme teknikleri, öğrenme görevlerindeki farklılıklardan bağımsız olarak, çoğu zaman insanların varoluşundan beri biriktirdiği bilgi ve deneyime dayanarak doğayı taklit etmeye çalışır. Makine öğrenme teknikleri insan beyni işlevinden, insan evrimini ele alan süreçlerden, insan bilgi edinme ve akıl yürütme teorisinden ve insan davranışlarının arkasındaki sosyolojik teori tarafından harekete geçirilmiştir.

Makine öğrenmesi, günümüzde verilerin giderek çoğalması ve bu verilerin bütününden yeni çıkarımlar elde edilmesini sağlar. Mevcut büyük verilerle, ileride tahmin edilebilecek sonuçların üretilmesine zemin oluşturur. Günümüzde veriler karmaşık ve büyük hacimli olmaları nedeniyle, insanoğlunun bunları kısa zamanda yorumlama kapasitesinin kat be kat üzerindedir. Bu noktadan itibaren, makine öğrenmesi devreye girer karmaşık verilerden ileriye dönük doğru analizler çıkarılmasını sağlar.

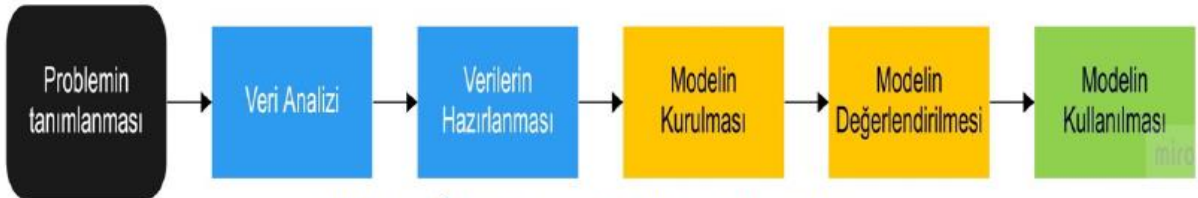
Makine öğrenimi yapay zekânın bir alt katmanında olup, derin öğrenmenin bir üst kümesinde yer almaktadır. Sinir ağları – derin öğrenme – makine öğrenmesi – yapay zekâ olarak sıralanır. Makine öğreniminde algoritmalar, büyük veri kümelerindeki kalıpları ve korelasyonları bulmak ve bu analizlere dayalı olarak en iyi kararları ve tahminleri yapmak için eğitilirler. Makine öğrenimi uygulamaları, kullanımlarıyla birlikte gelişir ne kadar fazla veriye sahip olurlarsa o kadar doğru sonuçlar üretirler.



Makine öğrenimi şeması

Şekilde de görüldüğü gibi, makine öğrenmesi aslında derin öğrenme ve yapay zekâyla ilişkilidir. Yapay zekânın oluşabilmesi için makine öğrenmenin olması gerekmektedir. Veri kümesini modele dönüştürmek için makine öğrenimi en uygun algoritma becerisidir.

Geleneksel kullanılan programlamada yazar tarafından bilgisayara girilen veriler, çıktılarına nasıl dönüştürüleceğini sisteme girilen bir dizi talimatlarla belirlenir. Bu talimatlara göre program buna göre bir eylem gerçekleştirir. Makine öğrenimde ise, insan girdisi minimal hatta yok denecek kadar azdır. Bu durumda var olan sorunu çözümlenmesi ve geçmiş gözlemler değerlendirilerek harekete geçmesi sağlanan otomatize edilmiş bir süreçtir. Makine öğreniminin çalışma prensibi şöyle ifade edebiliriz.



Makine Öğreniminin Çalışma Prensibi

Makine öğrenimi ve yapay zekâ bu nedenle birbiriyle bu süreçte karıştırılır. Yapay zekâ ve makine öğrenimi bambaşka süreçlerdir. Yapay zekâda makinalar kararları kendi verir bundan yeni beceriler öğrenir ve sorunları insanların çözdüğü gibi çözmeye çalışır.

Makine öğrenimi ise verilerden beslenerek yeni şeyler öğrenme sistemidir. Algoritmaları programlamak yerine kalıpları otomatik olarak tanımlayarak etiket verileri örnekleri beslenebilir. Makinalara “Elma veya armudu” etiketlemeyi öğretirsek, sonraki süreçte hiçbir yardım almadan, elma ve armudu kendi başına etiketlemeye başlayacaktır.

DeneySEL verilere dayalı davranışlara göre algoritma tasarlama ve geliştirme, Makine Öğrenimi olarak bilinir. Yapay zeka, makine öğrenimine ek olarak bilgi sunumu, doğal dil işleme, planlama, robotik gibi diğer konuları da kapsar.

Makine Öğrenmesi Tanımları

- Makine öğrenimi, deneyimle otomatik olarak öğrenmek ve gelişmek için sistem programlamayla ilgilenen bir bilgisayar bilimi dalıdır. Örneğin: Robotlar, sensörlerden topladıkları verilere göre görevi yerine getirebilecek şekilde programlanmıştır. Verilerden programları otomatik olarak öğrenir.
- Veri yığınının tahmin etmeye ya da karar vermeye yönelik otonom davranış paternleri geliştiren algoritmalar ve matematiksel modellerin oluşturulmasıdır.
- Veri yığınının kendi kendine öğrenen matematiksel modeller ve algoritmalar ile insandan bağımsız otonom davranış geliştirilmesidir.
- Yapısal işlev olarak veri yığınının öğrenebilen ve veriler üzerinden karar vermeye yönelik tahmin yapabilen algoritmaların çalışma ve inşalarını araştıran bir sistemdir.

Makine öğrenmesi algoritmalarından oluşur. Açık bir şekilde tahminler veya kararlar vermek için "öğrenme verileri" olarak bilinen verilere dayalı kendi kendine öğrenen bir matematiksel model oluşturulmasıdır. Makine öğrenmesi, bilgisayarları açıkça programlanmaksızın görevleri nasıl gerçekleştirebileceklerini keşfetmeyi içerir. Belirli görevleri yerine getirmeleri için verilerden öğrenen algoritmaları içerir.

Atanan basit görevler için, makineyi eldeki sorunu çözmek için gereken tüm adımların nasıl yürütüleceğini söyleyen algoritmaları programlamak mümkündür; bilgisayar tarafında öğrenmeye gerek yoktur. Daha gelişmiş görevler için, bir insanın gerekli algoritmaları manuel olarak oluşturması zor olabilir. Makine Öğrenmesi uygulamasında, programcıların gereken her adımı belirtmesinden ziyade makinenin kendisinin algoritmaları geliştirmesine yardımcı olmaktadır.

Makine öğrenmesinin omurgasını oluşturan disiplinler

- İstatistiksel hesaplama ile yakından ilgilidir.
- Matematiksel optimizasyon çalışması, makine öğrenmesi alanına yöntemler, teori ve uygulama alanları sağlar.
- Veri madenciliği, denetimsiz öğrenim yoluyla keşifsel veri analizine odaklanan ilgili bir çalışma alanıdır.
- Uygulamalı Matematik
- Bilgisayar sistemleri ve yazılımlar

Makine öğrenimi karmaşık olarak algılanan problemlerin analizi için mükemmel bir yöntemdir. Genel anlamda makine öğreniminin görevi sorunları tespit etmek, gelecekte olabilecek senaryoları önceden tahmin etmek ve büyük veri yığınının genel olarak kalıplarını keşfedebilme yeteneğine sahiptir ancak her türlü teknolojide olduğu gibi makine öğrenimi de tamamen mükemmel değildir.

Hata yapma olasılığını arttıran nedenler; verileri elde etmenin zorluğu gerekli hesaplama gücü getirilen karmaşıklık, bazı algoritmaların uzun süren eğitim süreleri vb. Bu nedenler makine öğrenimine geçişin engeli değil erteleyicisidir.

Makine Öğreniminin Türleri

Makine öğrenmesi türleri, yaklaşımlarına, girdikleri ve çıktıkları veri türlerine ve çözmeleri amaçlanan görev veya sorun türlerine göre farklılık gösterirler.

- Denetimli Öğrenme (Supervised Learning Algorithms)
- Denetimsiz Öğrenme (Unsupervised Learning Algorithms)
- Yarı denetimli Öğrenme
- Takviye Öğrenme
- Transdüksiyon
- Öğrenmeyi öğrenmek

Genellikle, makine öğrenimi yöntemleri iki aşamaya ayrılır:

1) Eğitim modeli: Verilerin bir koleksiyonundan öğrenilir.

2) Uygulama modeli: Yeni test verileri hakkında kararlar almak için kullanılır.

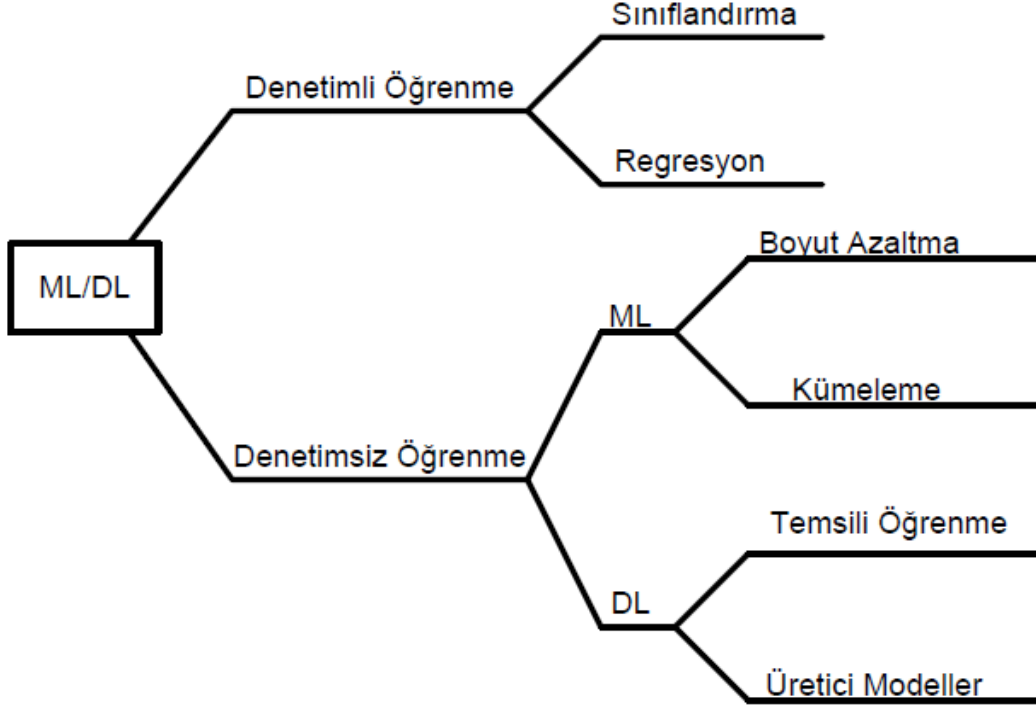
Makine öğrenimi türlerinden bazıları şunlardır:

- Eğitim verilerinin doğru yanıtlarla etiklendiği denetimli öğrenim. En yaygın iki denetimli öğrenme türü, sınıflandırma ve regresyondur.
- Analiz etmek ve keşfetmek istediğimiz kalıpları, etiketlenmemiş verilerden oluşan bir koleksiyondan öğrenen denetimsiz öğrenme. En önemli iki örnek, boyut küçültme ve kümelenmedir.
- Robot veya kontrolör gibi bir temsilcinin geçmişteki eylemlerin sonuçlarına dayalı olarak uygun eylemleri öğrenmeye çalıştığı pekiştirmeli öğrenme.
- Eğitim verilerinin yalnızca bir alt kümesinin etiklendiği yarı denetimli öğrenme.
- Mali piyasalarda olduğu gibi zaman serisi tahmini
- Fabrikalarda ve gözetimde arıza tespiti için kullanılanlar gibi anormallik tespiti.
- Verilerin elde edilmesinin pahalı olduğu aktif öğrenme

Bu nedenle bir algoritmanın hangi eğitim verilerinden elde edileceğini ve diğerlerini belirlenmesi gerekir.

Makine öğrenimi modeli nedir?

Bir makine öğrenimi modeli, makine öğrenimiyle ilgili görevleri işleme koyan bir soru ya da yanıtlama sistemidir. İş sonuçlarını iyileştirmek için değerli iç görüler toplamak, etkin olan parametreleri belirlemek için verileri kullanmayı amaçlar.



Denetimli Öğrenme

Denetimli öğrenme, etiketli eğitim verilerinden bir fonksiyon oluşturulması ile ilgili makine öğreniminin bir dalıdır. Denetimli öğrenmede, eğitim verileri bir dizi giriş ve hedef çiftinden oluşur; burada giriş, özelliklerin bir vektörü (Öznitelik Vektörü) olabilir ve hedef, işlevin çıktısı için ne istediğimizi belirtir. Hedef, sınıfın veya değer etiketinin tahmin edilmesidir.

Hedefin türüne bağlı olarak, denetimli öğrenimi kabaca iki kategoriye ayırır: Sınıflandırma ve regresyon. (Kategori: Aralarında herhangi yönden benzerlik, bağ ya da ilgi bulunması)

– Sınıflandırma, aralarında herhangi yönden benzerlik, bağ ya da ilgi bulunan hedefleri içerir; Görüntü sınıflandırması gibi bazı basit durumlardan makine çevirileri ve resim yazısı gibi bazı gelişmiş konulara kadar değişen örnekler.

– Regresyon, nicel (sayısal) değişkenler arasındaki ilişkilerin belirlendiği hedefleri içerir. Uygulamaların tümü bu kategoriye girer. Örneğin, stok tahmini, görüntü maskeleyme ve diğerlerini içerir.

Denetimli Öğrenmenin işlevleri:

- Sınıflandırmalar
- Regresyon

Denetimli öğrenmeye yönelik standart yaklaşım, örnek setini eğitim seti ve teste bölmektir. Makine öğrenimi gibi bilgi biliminin çeşitli alanlarında, "Eğitim Seti" olarak bilinen potansiyel olarak öngörücü ilişkiyi keşfetmek için bir dizi veri kullanılır. Eğitim seti öğrenen algoritmaya verilen bir örnektir, Test seti ise öğrenci tarafından oluşturulan hipotezlerin doğruluğunu test etmek için kullanılır ve öğrenen algoritmalarından saklanan bir örnek setidir. Eğitim seti, test setinden farklıdır.

Denetimli öğrenmenin ne olduğunu anlamak için, bir çocuğa içinde on aslan, on maymun, on fil ve diğerleri gibi her türden on hayvanın bulunduğu 100 oyuncak hayvan verdiğimizizi düşünelim. Daha sonra, çocuğa bir hayvanın farklı özelliklerine (özelliklerine) dayalı olarak farklı hayvan türlerini tanımayı öğretiyoruz. Rengi turuncuysa, o zaman bir aslan olabilir. Gövdesi büyük bir hayvansa, fil olabilir. Çocuğa hayvanları nasıl ayırt edeceğini öğretiriz, bu denetimli öğrenmeye bir örnektir. Artık çocuğa farklı hayvanlar verdiğimizizde, onları uygun bir hayvan grubuna ayırabilmelidir.

Aynı bilgisayarlar için de geçerlidir. Onlara gerçek etiketli değerleri ile binlerce veri noktası sağlıyoruz (Etiketli veriler, özellik değerleriyle birlikte farklı gruplara sınıflandırılır). Daha sonra eğitim döneminde farklı özelliklerinden ders çıkarır. Eğitim dönemi bittikten sonra eğitilmiş modelimizi tahmin yapmak için kullanabiliriz. Makineyi zaten etiketli verilerle beslediğimizi, bu nedenle tahmin algoritmasının denetimli öğrenmeye dayandığını unutmayın.

Denetimli öğrenme algoritmaları

- K-En Yakın Komşular
- Doğrusal Regresyon
- Lojistik regresyon
- Rastgele Orman
- Gradyan Güçlendirilmiş Ağaçlar
- Destek Vektör Makineleri (SVM)
- Naive Bayes
- Nöral ağlar
- Karar ağaçları

Denetimli öğrenme algoritmaları, hem girdileri hem de istenen çıktıları içeren bir veri kümesinin matematiksel bir modelini oluşturur. Veriler, öğrenen verileri olarak bilinir ve bir dizi öğrenme örneğinden oluşur. Her öğrenme örneğinde, denetim sinyali olarak da bilinen bir veya daha fazla giriş ve istenen çıkış bulunur.

Matematiksel modelde, her öğrenme örneği bazen özellik vektörü olarak adlandırılan bir dizi veya vektörle temsil edilir ve öğrenme verileri bir matrisle temsil edilir. Nesnel bir fonksiyonun yinelemeli optimizasyonu yoluyla, denetimli öğrenme algoritmaları yeni girdilerle ilişkili çıktıyı tahmin etmek için kullanılacak bir işlevi öğrenir.

Optimal bir fonksiyon, algoritmanın egzersiz verilerinin bir parçası olmayan girişler için çıkışı doğru bir şekilde belirlemesine izin verecektir. Zaman içinde çıktıların veya tahminlerinin doğruluğunu arttıran bir algoritmanın bu görevi yerine getirmeyi öğrendiği söylenir.

Denetimli öğrenme algoritması türleri sınıflandırma ve regresyonu içerir. Sınıflandırma algoritmaları, çıktılar sınırlı bir değer kümesiyle sınırlandırıldığında ve regresyon algoritmaları, çıktılar bir aralık içinde herhangi bir sayısal değere sahip olduğunda kullanılır. Örnek olarak, e-postaları filtreleyen bir sınıflandırma algoritması için girdi gelen bir e-posta olur ve çıktı da e-postanın dosyalanacağı klasörün adı olur.

Denetimsiz Öğrenme

Öğrenme algoritmasına hiçbir etiket verilmez ve girdisinde yapı bulmak için tek başına bırakılır. Denetimsiz öğrenme kendi içinde bir hedef (verilerdeki gizli kalıpları keşfetme) veya bir sona doğru bir araç (özellik öğrenme) olabilir.

- k-means clustering
- t-SNE (t-Distributed Stochastic Neighbor Embedding)
- PCA (Principal Component Analysis)
- Association rule

Denetimsiz öğrenme algoritmaları, yalnızca girdileri içeren bir veri yığını alır ve verileri yapısal olarak gruplar veya kümeler. Bu nedenle algoritmalar, etiketlenmemiş, sınıflandırılmamış veya kategorize edilmemiş test verilerinden öğrenilir. Geri bildirim yanıt vermek yerine, denetimsiz öğrenme algoritmaları verilerdeki ortaklıkları tanımlar ve her yeni veri parçasında bu tür ortak özelliklerin varlığına veya yokluğuna bağlı olarak tepki verir.

Denetimsiz öğrenmenin merkezi bir uygulaması, olasılık yoğunluk işlevini bulmak gibi istatistiklerde yoğunluk tahmini alanındadır. Denetimsiz öğrenme, veri özelliklerini özetleme ve açıklamayı içeren diğer alanları kapsar.

Denetimsiz Öğrenmenin işlevleri:

- Veri kümeleri bulunur,
- Verilerin düşük boyutlu temsilleri bulunur,
- Verilerde ilginç yönler bulunur,
- İlginç koordinatlar ve korelasyonlar elde edilir,
- Yeni gözlemler ya da veri tabanı elde edilir.

Elmalardan oluşan torbayı taşıyan robotun taşıdığı torba parçalanır ve tüm elmalar (Çürük, büyük, küçük, ort boy, ...) birbirine karışır. Robot, topladığı elmaları önceden etiketlenmediğinden hemen ayırtamaz. Tek başına farklılıkları öğrenerek ayırtama işleme yapar. Denetimli öğrenmenin aksine, denetimsiz öğrenme, verilerdeki gizli yapıları tanımlayan bir işlev olan etiketlenmemiş verilerden kaynaklanır.

Denetimsiz öğrenme algoritmalarına örnekler:

- Boyut Küçültme
- Yoğunluk Tahmini
- Pazar Sepeti Analizi
- Üretken düşmanlık ağları (GAN'lar)
- Kümeleme

Yarı denetimli öğrenme

Yarı denetimli öğrenme, denetimsiz öğrenme (herhangi bir etiketlenmiş öğrenme verisi olmadan) ve denetimli öğrenme (tamamen etiketlenmiş öğrenme verisi ile) arasındadır. Bazı öğrenme örnekleri öğrenme etiketlerinin eksik olmasına rağmen, birçok makine öğrenmesi araştırmacısı, etiketlenmemiş verilerin, az miktarda etiketlenmiş verilerle birlikte kullanıldığında, öğrenme doğruluğunda önemli bir gelişme sağlayabildiğini bulmuştur.

Zayıf denetimli öğrenmede, öğrenme etiketleri gürültülü, sınırlı veya kesin değildir; bununla birlikte, bu etiketlerin elde edilmesi genellikle daha ucuzdur, bu da daha büyük etkili öğrenme setleriyle sonuçlanır.

Takviyeli Öğrenme

Takviye öğrenmesi, yazılım temsilcilerinin kümülatif ödül kavramını en üst düzeye çıkarmak için bir ortamda nasıl işlem yapmaları gerektiğiyle ilgili bir makine öğrenmesi alanıdır. Genelliği nedeniyle, oyun teorisi, kontrol teorisi, yöneylem araştırması, bilgi teorisi, simülasyon tabanlı optimizasyon, çok etmenli sistemler, sürü zekası, istatistikler ve genetik algoritmalar gibi birçok disiplinde çalışılmaktadır.

Makine öğrenmesinde, ortam tipik olarak bir Markov Karar Süreci (MKS) olarak temsil edilir. Pek çok takviye öğrenme algoritması dinamik programlama teknikleri kullanır. Takviye öğrenme algoritmaları, MKS'nin kesin bir matematiksel modeli hakkında bilgi sahibi değildir ve kesin modeller mümkün olmadığında kullanılır. Takviye öğrenme algoritmaları otonom araçlarda veya bir insan rakibe karşı bir oyun oynamayı öğrenmek için kullanılır.

- Q-Learning
- Temporal Difference (TD)
- Monte-Carlo Tree Search (MCTS)
- Asynchronous Actor-Critic Agents (A3C)

Pekiştirmeli öğrenme

Pekiştirmeli öğrenme, davranışçılıktan esinlenen, öznelere bir ortamda en yüksek ödül miktarına ulaşabilmesi için hangi eylemleri yapması gerektiğiyle ilgilenen bir makine öğrenmesi yaklaşımıdır. Bu problem, genellikle ötürü oyun kuramı, kontrol kuramı, yöneylem araştırması, bilgi kuramı, benzetim tabanlı eniyileme ve istatistik gibi birçok diğer dalda da çalışılmaktadır.

Makine öğrenmesinde, ortam genellikle bir Markov karar süreci (MKS) olarak modellenir, bu bağlamda birçok pekiştirmeli öğrenme algoritması dinamik programlama tekniklerini kullanır.

Pekiştirmeli öğrenme algoritmalarının klasik tekniklerden farkı, MKS hakkında ön bilgiye ihtiyaç duymamaları ve kesin yöntemlerin verimsiz kaldığı büyük MKS'ler için kullanılmalarıdır.

Pekiştirmeli öğrenme, doğru girdi/çıkı eşleşmelerinin verilmemesi ve optimal olmayan eylemlerin dışarıdan düzeltilmemesi yönleriyle gözetimli öğrenmeden ayrışır. Dahası, pekiştirmeli öğrenmede bilinmeyen uzayda keşif ile mevcut bilgiden istifade arasında bir denge kurma söz konusudur.

Özellik Öğrenme

Temsili öğrenme algoritmaları olarak da adlandırılan özellik öğrenme algoritmaları, genellikle girdilerindeki bilgileri korumaya çalışır, ancak sınıflandırma veya tahminler gerçekleştirmeden önce genellikle bir ön işleme adımı olarak yararlı hale getirecek şekilde dönüştürür.

Bu teknik, bilinmeyen veri üreten dağıtımdan gelen girdilerin yeniden yapılandırılmasına izin verirken, bu dağıtım altında mantıksız olan yapılandırmalara mutlaka sadık kalmaz. Bu, manuel özellik mühendisliğinin yerini alır ve bir makinenin hem özellikleri öğrenmesini hem de belirli bir görevi gerçekleştirmek için kullanmasını sağlar.

Özellik öğrenmesi denetimli veya denetimsiz olabilir. Denetimli özellik öğrenmede, özellikler etiketli giriş verileri kullanılarak öğrenilir. Örnekler arasında yapay sinir ağları, çok katmanlı algılayıcılar ve denetimli sözlük öğrenmesi sayılabilir.

Denetimsiz özellik öğrenmede, özellikler etiketlenmemiş girdi verileriyle öğrenilir. Örnekler arasında sözlük öğrenmesi, bağımsız bileşen analizi, otomatik kodlayıcılar, matris çarpanlarına ayırma ve çeşitli kümeleme biçimleri bulunmaktadır.

Özellik öğrenme, sınıflandırma gibi makine öğrenmesi görevlerinin genellikle matematiksel ve hesaplamaya uygun olarak işlenmesi için girdi gerektirmesi gerçeğiyle motive edilir. Bununla birlikte, görüntüler, video ve duyuşal veriler gibi gerçek dünya verileri, belirli özellikleri algoritmik olarak tanımlama girişimlerine yol açmamıştır.

Bir alternatif, açık algoritmalara dayanmadan, bu özellikleri veya gösterimleri muayene yoluyla keşfetmektir.

Makine Öğrenmesi Algoritmaları

Makine öğrenmesi algoritmalarının temel bileşenleri:

- Hesaplamalar için çeşitli kütüphaneler kullanıldığından gerekli kütüphaneler aktarılır.
- Veritabanı dosyaları okunur.
- Etiketli verilerin ağırlıklı ortalama, standart sapma gibi istatistiksel analiz özellikleri belirlenir ve yorumlanır.
- Verilerin grafiği çizilir.
- Değerler tahmin edilirken dikkate almak istenilen özellikler ve katsayılar seçilir.
- Bir modelin doğruluğunu kontrol etmek için, veriler eğitim ve test veri setlerine ayrılır.
- Model eğitilir
- Test veri seti için bir tahmin fonksiyonu bulunur.
- Test verilerinin doğruluğu kontrol edilir: Gerçek değerleri veri setindeki tahmin edilen değerlerle karşılaştırılarak bir modelin doğruluğu kontrol edilebilir.

Makine Öğreniminin beş popüler algoritması:

- Karar ağaçları
- Sinir Ağları (geri yayılım)
- Olasılık ağları
- En yakın komşu
- Vektör makineleri desteklemek

Makine Öğrenmesi Algoritmaları

1- Regresyon (Tahmin) Sürekli değerleri tahmin etmek için regresyon algoritmaları kullanılır.

- Doğrusal Regresyon
- Polinom Regresyon
- Üstel Regresyon
- Lojistik regresyon
- Logaritmik Regresyon

2-Sınıflandırma Bir dizi ögenin sınıfını veya kategorisini tahmin etmek için sınıflandırma algoritmaları kullanılır.

- K-En Yakın Komşular
- Karar ağaçları
- Rastgele Orman
- Destek Vektör Makinesi
- Naive Bayes

3- Kümeleme Özetlemek veya verileri yapılandırmak için kümeleme algoritmaları kullanılır.

- K-means
- DBSCAN
- Mean Shift
- Hierarchical

4- İlişkilendirme Birlikte meydana gelen öğeleri veya olayları ilişkilendirmek için ilişkilendirme algoritmaları kullanılır.

- Apriori

5- Anomali (Sapma) Algılama: Anormal etkinlikleri ve dolandırıcılık tespiti gibi olağandışı durumları keşfetmek için anormallik algılama kullanılır.

Veri madenciliğinde, aykırı tespit olarak da bilinen anomali tespiti, verilerin çoğundan önemli ölçüde farklılık göstererek şüphe uyandıran nadir maddelerin, olayların veya gözlemlerin tanımlanmasıdır. Tipik olarak, anormal kalemler banka sahtekarlığı, yapısal bir kusur, tıbbi sorunlar veya bir metindeki hatalar gibi bir konuyu temsil eder. Anomalilere aykırı değerler, yenilikler, gürültü, sapmalar ve istisnalar denir.

Özellikle, kötüye kullanım ve ağ izinsiz giriş tespiti bağlamında, ilginç nesnelere genellikle nadir nesnelere değil, faaliyette beklenmedik patlamalardır. Bu model, bir aykırı değer nadir bir nesne olarak ortak istatistiksel tanımına uymaz ve uygun şekilde toplanmadığı sürece birçok aykırı algılama yöntemi (özellikle denetimsiz algoritmalar) bu verilerde başarısız olur. Bunun yerine, bir küme analiz algoritması bu örüntüler tarafından oluşturulan mikro kümeleri tespit edebilir.

Üç geniş anomali tespit tekniği kategorisi bulunmaktadır.

Gözetimsiz anomali tespit teknikleri, veri kümesindeki örneklerin çoğunun normal olduğu varsayımıyla etiketlenmemiş bir test veri kümesindeki anormallikleri, veri kümesinin geri kalan kısmına en az uyan görünen örnekleri arayarak tespit eder.

Denetimli anomali algılama teknikleri, "normal" ve "anormal" olarak etiketlenmiş ve bir sınıflandırıcıyı (diğer birçok istatistiksel sınıflandırma problemi için temel fark, aykırı algılamanın doğasında dengesiz doğasıdır) içeren bir veri seti gerektirir.

Yarı denetimli anomali tespit teknikleri, belirli bir normal eğitim veri setinden normal davranışı temsil eden bir model oluşturur ve daha sonra model tarafından bir test örneğinin üretilme olasılığını test eder.

6- Sıra Desen Madenciliği Örüntü - Pattern: Bir nesnenin ya da olayın iki veya üç boyutlu, uzaysal ve geometrik davranışının matematiksel ifadesidir. Diğer bir ifadeyle, nesnenin davranışı ile ilgili uzayda gözlenebilir veya ölçülebilir geometrik bilgilerdir.

Örüntü Tanımanın kullanıldığı alanlar:

- Örüntü Tanıma şu alanlarda kullanılabilir:
- Bilgisayar görüşü
- Konuşma tanıma
- Veri madenciliği
- İstatistik
- Gayri Resmi Erişim
- Biyo-Bilişim

7- Boyut Küçültme Makine Öğrenimi ve istatistikte boyut küçültme, dikkate alınan rastgele değişkenlerin sayısını azaltma işlemidir ve özellik seçimi ve özellik çıkarımı olarak ikiye ayrılabilir.

Bir veri kümesinden yalnızca yararlı özellikleri çıkarmak için verilerin boyutunu küçültmek için boyut azaltma kullanılır. Boyut azaltma, denimsiz bir öğrenme tekniğidir.

Veri biliminde, boyut indirgeme, bir verinin yüksek boyutlu bir uzaydan, düşük boyutlu bir uzaya, anlamını kaybetmeyecek şekilde dönüştürülmesidir. Yüksek boyutlu bir veriyi işlemek daha fazla işlem yükü gerektirir. Bu yüzden, yüksek sayıda gözlemin ve değişkenin incelendiği sinyal işleme, konuşma tanıma, nöroinformatik, biyoinformatik gibi alanlarda boyut indirmesi sıkça kullanılır.

Boyut indirgeme yaklaşımları doğrusal ve doğrusal olmayan olarak ikiye ayrılır.

Boyut indirgeme var olan özneliklerin bir alt kümesini seçerek ya da yeni öznelikler çıkararak yapılabilir. Boyut indirgemesi gürültü filtreleme, veri görselleştirme ya da kümeleme analizi amacıyla kullanılabilir gibi, diğer makine öğrenimi yöntemlerinin ön adımı olarak uygulanabilir.

8- Önerilerden eğilim ya da yön bulma Öneri motorları oluşturmak için öneri algoritmalarını kullanılır.

- Netflix öneri sistemi.
- Bir kitap tavsiye sistemi.
- Amazon'da bir ürün öneri sistemi.

Kaynaklar

Gökalp, Ö. M. Makine Öğrenmesi.

Aylak, B. L., Oral, O., & Yazıcı, K. (2021). Yapay Zeka ve Makine Öğrenmesi Tekniklerinin Lojistik Sektöründe Kullanımı. *El-Cezeri*, 8(1), 74-93.

Karakuş, C., Makine Öğrenmesi Temelleri Ders Notu