

Makine Öğrenimi: İşletme ve Veri Bilimi Arasındaki Köprü Araç

Son birkaç yıldır elektrik ve medya bağlantısı olmayan uzak bir adada yaşamıyorsanız, muhtemelen makine öğrenmesi konusunu duymuşsunuzdur. Kendi kendini süren arabalar, sohbet eden robotlar, AlphaGo veya tahmine dayalı analitik hakkında her konuştuğumuzda ki bu alanların herbirinde bazı makine öğrenme tekniklerinin uygulanması söz konusudur. Hikayeler ve tavsiyeler bol olsa da, makine öğrenmesi henüz iş hayatı için bir sonuç belirleyici alan olmamıştır. Halkın algısına göre, ML'de uygulanan algoritmalar bilim kurguya yakındır ve ML'nin benimsenmesi için somut bir plan ortaya koymak hala yüksek bir engeldir.

Bu nedenle, bu tanıtım belgesi vizyon belirlemek ve yapay zekaya eğilimi artırmak yerine pratik soruları cevaplamayı amaçlamaktadır. Bu çalışmada ana konu olarak veri bilimi ve alt alanlarının nasıl etkileşime girdiği, makine öğrenmesinin çözebileceği temel problemler ve bu problemlerin işletme diline nasıl çevrilebileceği ile ilgili. Ayrıca yetenek kazanma konusunda alınacak ana kararları da tartışıp ve önceden göz önünde bulundurulması gereken zorlukları tespit etmeye çalışacağız. Kısaca, bu çalışmada seyahat ve sağlık endüstrisi için yazılmış makalelerin potansiyelini ele aldığımız için bu alanlardaki makine öğrenmesi vizyonunu belirlemeye odaklanıyoruz.

-1-) Öncelikle Veri Bilimi(Data Science) kavramını ele alalım

Makine öğrenimi kavramı ilk kez 1950'lerde AI'nın öncüleri olarak tanıtıldı. 1950'de Alan Turing, bugün Turing Testi olarak bildiğimiz ünlü bir AI değerlendirme testini öneren “Bilgi İşlem Makinaları ve Zeka(Computing Machinery and Intelligence)” makalesini yayınladı. 1959'da Arthur Lee Samuel “makine öğrenmesi” terimini kullandı. Bugün kullandığımız birçok teorik keşifler o zaman yapıldı. Fakat neden bugün makine öğrenmesi ve veri bilimi hakkında bu kadar konuşuyoruz? Şimdi biraz daha açıklamaya çalışalım.

Belki de, en önemli fark, bilgisayarlardaki hesaplama güçleri ve geçmiş yıllara kıyasla toplayıp analiz edebileceğimiz veri miktarı bu durumun en önemli etkenlerinden olmalı. Bugün bir akıllı telefon, 60'larda birkaç oda büyüklüğünde olan bir bilgisayardan daha fazla veri depolayabilir ve işleyebilir. İyileştirilmiş ve küçük veri kümelerine dayanmak yerine, algoritmaları

eğitmek ve öngörülerini çizmek için binlerce parametreyle büyük ve örgütlenmemiş verileri kullanabiliriz. Verilerin miktarı ve kalitesi de modern makine öğrenme tekniklerini istatistiklerden ayıran özelliktir. İstatistikler, bir sezgi yakalamak için genellikle birkaç değişkene dayanırken, makine öğrenmesi binlerce veri özelliğiyle etkili bir şekilde kullanılabilir.

Bu bölümde Veri Biliminin çeşitli alanlarını ve bu alanlar ile olan ilişkileri inceleyelim.

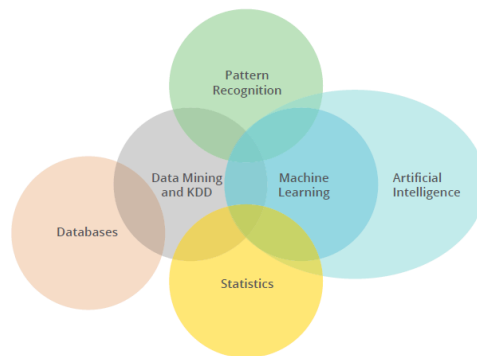
1.1 Veri Bilimi (Data Science)

Veri bilimi terimi, 1960'larda kullanılmaya başlandı. Birçok tanımı olsa da, iş merkezli olanı MailChimp Baş Veri Bilimcisi John W. Foreman tarafından dile getirildi:

***“Veri bilimi:** Matematiği ve istatistiği kullanan verilerin anlamlı içgörülere, kararlara ve ürünlere dönüştürülmesidir”*

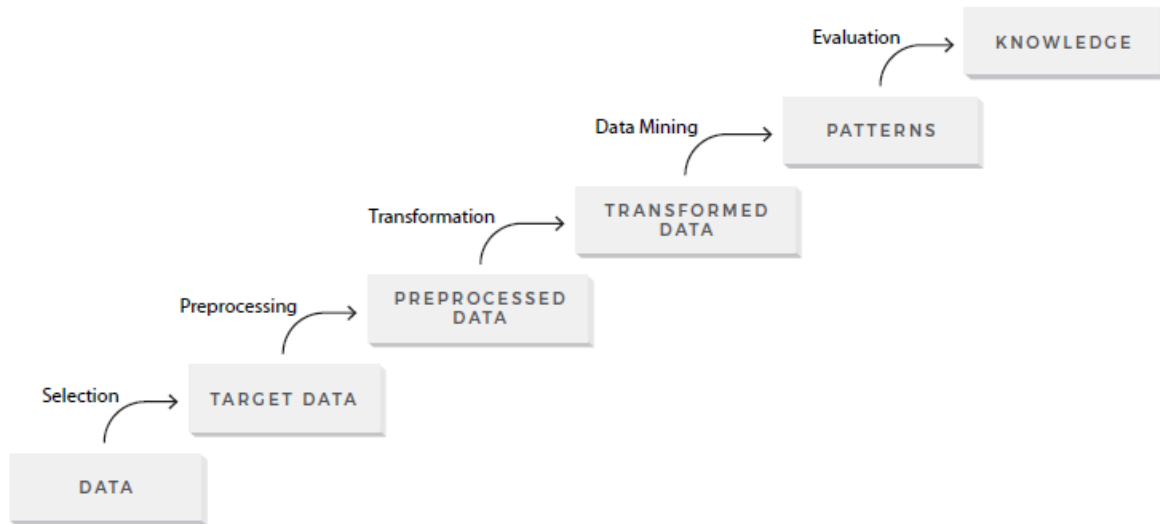
Veri bilimi zamanla geliştikçe ve yeni “araçlar” kazandıkça, temel iş hedefi, yararlı modeller bulma ve verilerden değerli bilgiler edinmeye odaklanmış olmaya devam ediyor. Günümüzde veri bilimi geniş bir endüstri yelpazesinde kullanılmaktadır ve çeşitli analitik problemlere yardımcı olmaktadır. Örneğin, pazarlamada, müşteri yaşını, cinsiyetini, yerini ve davranışını araştırmak, müşterilerin bir satın alma yapma veya yapmamaya ne kadar eğilimli olduğunu değerlendirmek için yüksek oranda hedeflenmiş kampanyalar yapılmasına olanak tanır. Bankacılıkta, müşteri olmayan kişilerin eylemlerini bulmak ve dolandırıcılık tespitinde yardımcı olmakta aynı zamanda bankacılıkta hangi müşteriye kredi çıkar hangisine çıkmaz gibi algoritmalarda yürütülmektedir. Sağlık hizmetlerinde ise hastaların tıbbi kayıtlarını analiz etmek, hangi hastalıklara sahip olma olasılığını gösterebilir, tıpta önemli bir kavram olan önleyici faaliyetler için önceden veriye dayalı tahminler yaparak önlemler alınabilir.

-Veri bilimi: farklı teknik ve araçlardan yararlanan, birbirine bağlı birden çok alanı kapsamaktadır yani multi disiplinler bir bilim alanıdır.



1.2 Veri Madenciliği ile Verilerin Keşfedilmesi

Şemada gördüğünüz gibi, tüm veri bilimi alanları, veri bilimi içindeki temel uygulama kümesini temsil ettiği için veri madenciliği ile bağlantılıdır. Veri madenciliği terimi aslında yanıltıcıdır ve kastettiği gibi değildir. Araştırma verilerinin kendisi yerine, disiplin, büyük ve muhtemelen yapılandırılmamış verilerden değerli bilgiler elde etmek için algoritmalar oluşturmaktır. Veri madenciliğinin temel sorunu, mevcut verileri haritalandırmak ve sindirilebilir(işlenebilir) kalıplara dönüştürmektir. Veri madenciliği, 1984 yılında Gregory Piatetsky-Shapiro tarafından başlatılan Veritabanlarında Bilgi Keşfi [(KDD)](<http://www.aaai.org/ojs/index.php/aimagazine/article/viewFile/1230/1131>) adı verilen daha geniş bir sürecin parçası olarak kabul edilir.



Veri madenciliği ve Veri Tabanlarında Bilgi Keşfi(KDD)'nin yalnızca veri biliminin temel sorununu ele almasına rağmen, makine öğrenmesi buna işletme verimliliğini de ekliyor.

1.3 Makine Öğrenmesi(Machine Learning)

Veri madenciliği ile popüler makine öğrenmesi arasında bir fark var. Yine de, makine öğrenmesi değerli içgörüler ortaya çıkarmak için algoritmalar oluşturmakla ilgilidir, dinamik olarak değişen ortamlarda sürekli kullanıma yoğun bir şekilde odaklanır ve önceki deneyimlere dayanan algoritmaların ayarlanması, yeniden eğitilmesi ve güncellenmesi üzerinde durur.

Makine öğreniminin amacı sürekli olarak yeni verilere uyum sağlamak ve içindeki yeni kalıpları veya kuralları keşfetmektir. Bazen insan rehberliği ve açıkça yeniden programlama olmadan dahi bu işlemler gerçekleştirilebilir.

Makine öğrenmesi, günümüzde teorik ve teknolojik gelişmelere bağlı olarak en dinamik gelişen veri bilimi alanıdır. Doğal dil işleme, görüntü tanıma ve hatta makineler tarafından yeni görüntüler, müzik ve metinler üretilmesine dahi yol açtılar. Makine öğrenmesi yapay zekanın oluşturulmasında ana “araç” olmaya devam etmektedir.

1.4 Yapay Zeka(Artificial Intelligence)

Yapay zeka (AI) belki de en belirsizce anlaşılan veri bilimi alanıdır. Ayrıca diğerlerinden ayrı olarak durur çünkü Yapay Zeka bir çatı ise, Makine Öğrenmesi, Derin Öğrenme , Veri Bilimi ise onun alt bileşenleridir. AI inşa etmenin arkasındaki ana fikir, insanlar gibi düşünebilecek ve akıl yürütebilecek bir aracı (veya bu kabiliyete yaklaşmak) oluşturmak için örüntü tanıma ve makine öğrenmeyi kullanmaktır. Ancak, bu kadar yaygın kullanılan bu terim ile, AI'yi yorumlamak konusunda henüz bir fikir birliğine varamadığını gözlemlemekteyiz. Bilgi biçimlendirmek zordur ve belirleme yolları çoktur. İş dilinde AI, yeni sorunları çözme yeteneği olarak yorumlanabilir. Etkili olarak, yeni problemleri çözmek algının, genellemenin, muhakemenin ve değerlendirmenin sonucudur.

İnsanlara göre(İşin Uzmanı olmayan), AI genellikle makinelerin birçok bilgi alanıyla ilgili sorunları çözme yeteneği olarak algılanır. Bu onları insanlara biraz benzer yapar. Bununla birlikte, yapay genel zeka kavramı (AGI), bilim kurgu alanında kalır ve henüz mevcut olan en son gelişmelere uymuyor gibi değerlendirilir. Son zamanlarda Texas Hold'em'de insanları yenen AlphaGo, IBM Watson veya Libratus gibi ünlü sistemler yapay dar zekanın (ANI Artificial Narrow Intelligence) temsilcileridir. Bir alanda uzmanlaşırlar ve verileri işlemek için benzer tekniklere dayalı görevler yapabilirler. Dolayısıyla, ANI'den AGI'ye ölçeklendirme, veri biliminin henüz başaramadığı bir çabadır, ancak bu atılımın birkaç on yılda gerçekleşmesi muhtemeldir. Çoğu işi yapabilen ve yapabilecek olan makinelerin insanlar tarafından artan korkuları kısmen makul olsada, bu makinaların dünyaya hakim olma senaryosu doğru değildir.

1.5 Büyük Veri(Big Data)

Büyük veri aynı zamanda aşırı abartılmış ve yanlış anlaşılmış bir kavramdır. İşletmelerde dijital dönüşümün büyümesi, müşteriler, çalışanlar ve kurumsal varlıklar hakkında çeşitli ve

genellikle yapılandırılmamış kayıtlar içeren gittikçe daha geniş veri kümelerinin toplanmasına olanak sağlamıştır. Bunlar demografik bilgiler, etkileşimler ve davranışlar, uç nokta cihazları ve kelimenin tam anlamıyla dijital yollarla izlenebilecek veya manuel olarak girilebilecek her şeyle ilgilidir. Ancak, bu yapılandırılmamış veri kümeleri henüz büyük veri değildir.

“Bilgi toplama o bilgiyi keşfetmek demek değildir” (Sean McClure, Ph.D. Space-Time Insight'ta Veri Bilimi Direktörü)

Bugün çok fazla veri toplamak mümkün olsa da, bu verilerin içgörülü kalıpların içinde keşfedilebileceği anlamına gelmez. Büyük veri kavramı, veri madenciliği ve makine öğrenmesi tekniklerini kullanarak büyük veri setlerinde kalıpları keşfetmeyi gerektirir. Neden bugün büyük verilere bu kadar önem veriyoruz? Teknoloji meraklıları arasında büyük verilerin popülerliği, hesaplama gücündeki son gelişmelerden kaynaklanmaktadır. Sonuçları tüm konu alanına keşfetmek ve tahmin etmek için sınırlı veri alt kümeleri kullanmak yerine, tüm ham verileri işleyebilir, daha yüksek doğruluk elde edebilir ve veriler arasında daha fazla gizli bağımlılık bulabiliriz. Bu, giderek daha büyük miktarda yapılandırılmamış veri kümesini hesaplayabilen, veri toplama araçları ve uzmanlığı ile verileri uygun şekilde görselleştirmek ve bu konuda bilgi edinmek için yüksek teknoloji altyapısı oluşturmayı gerektirir.

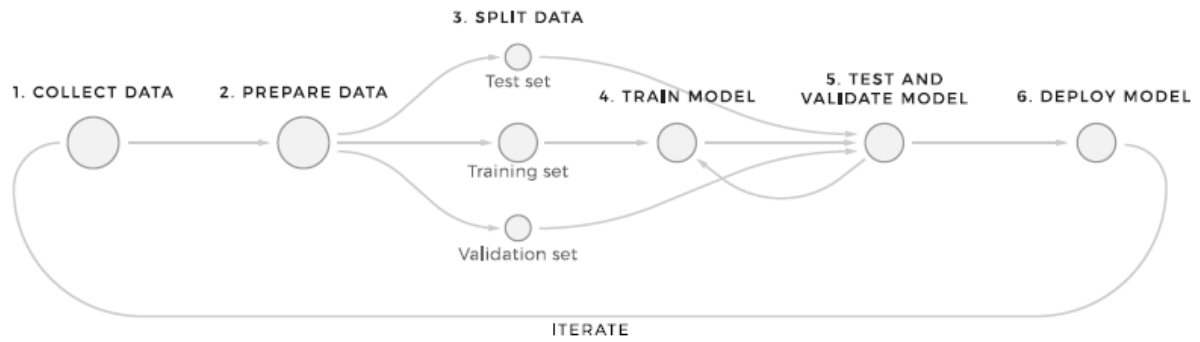
2. Makine Öğrenmesi Adımları

- Öyleyse algoritmaların veride faydalı örüntüler bulmasını nasıl sağlarız? Makine öğrenimi ve geleneksel olarak programlanmış algoritmalar arasındaki temel fark, açıkça programlanmadan verileri işleme yeteneğidir. Bu aslında bir mühendisin, bir makineye her bir veri kaydının nasıl ele alınacağına ilişkin ayrıntılı talimatlar sağlaması gerekmediği anlamına gelir. Bunun yerine, bir makine bu kuralların kendisini girdi verilerine dayanarak tanımlamaktadır.

- Belirli bir makine öğrenim uygulamasından bağımsız olarak, genel iş akışı aynı kalır ve sonuçlar bir kez belirlendiğinde ya da daha yüksek bir doğruluk istediğinde tekrar eder. Bu bölüm, makine öğrenmesi iş akışını oluşturan temel kavramlarına odaklanmıştır.

- Herhangi bir makine öğrenmesi uygulamasının temel amacı, bir algoritmanın, bir tarihsel veri alt kümesiyle eğitildikten sonra yeni verileri nasıl işlediğini açıklayan matematiksel bir modeldir. Eğitimin amacı, her veri nesnesinin bilinmeyen bir değeri olan bir hedef değer (özellik) formüle edebilen bir model geliştirmektir. Kulağa karmaşık gelse de, aslında değil çünkü anladıkça daha kolay olduğunu göreceksiniz.

Örneğin, e-Ticaret Mağazanızın müşterilerinin bir satın alma işlemi yapıp yapmayacağını tahmin etmeniz gerekir. Bu tahminler **satın alır** veya **satın almaz** olarak aradığımız hedef niteliklerdir. Bu tür tahminleri yapmada bir model geliştirmek için, farklı müşteri davranışlarının kayıtlarını (sonuçları almış veya satın almamış olsun) depolayan bir veri kümesiyle bir algoritmayı “eğitiriz” algoritmamız bu tarihsel verilerden öğrenerek, bir model oluşturarak, gelecekteki veriler hakkında tahminlerde bulunabilecek hale gelir işte bu şekilde makine öğrenmesi gerçekleşmiş olur.



Genel olarak, iş akışı şu basit adımlardan oluşur:

-Veri topla :**** Dijital altyapınızı ve diğer kaynaklarınızı mümkün olduğunca çok sayıda yararlı kayıt toplamak ve bunları bir veri kümesinde birleştirmek için kullanın.

-Veri hazırla :**** Verilerinizi mümkün olan en iyi işlenecek şekilde hazırlayın. Veri ön işleme ve temizleme prosedürleri oldukça karmaşık olabilir, ancak genellikle, bir sütundaki aynı değerlerin farklı gösterimleri gibi, eksik değerleri doldurmayı ve verilerdeki diğer hataları düzeltmeyi hedefler (örneğin, 14 Aralık 2016 ve 12.14.2016 algoritması tarafından aynı şekilde ele alınmalıdır).

-Verileri böl :**** Bir modeli eğitmek için veri alt kümelerini ayırın ve yeni verilere karşı nasıl performans gösterdiğini daha fazla değerlendirin.

-Bir model eğitin :**** Algoritmanın içindeki örüntüleri tanımasını sağlamak için bir tarihsel veri alt kümesi kullanın.

-Bir modeli test edin ve doğrulayın :**** Tarihsel verilerin test ve doğrulama alt kümelerini kullanarak bir modelin performansını değerlendirin ve tahminin ne kadar doğru olduğunu anlayın.

****Bir model oluřturun :**** Test edilen modeli, bir analitik çözümünün bir parçası olarak karar alma çerçevenize yerleřtirin veya kullanıcıların yeteneklerinden yararlanmasına izin verin (örneğin, ürün önerilerinizi daha iyi hedefleyin).

****Uygulayın :**** Modeli kullandıktan sonra adım adım iyileřtirmek için yeni veriler toplayin.

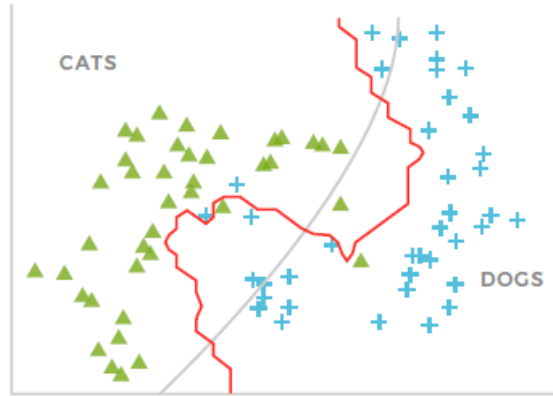
. 3. Makine Öğrenmesinin Çözebileceđi 5 Görev

- İş açısından, makine öğrenmesi geniş bir görev yelpazesine hitap eder, ancak daha yüksek seviyelerde algoritmaların çözdüğü görevler beř ana gruba ayrılır:

****Sınıflandırma, Kümeleme analizi, Regresyon, Sıralama ve Üretim**.**

3.1 Sınıflandırma(Classification)

Sınıflandırma algoritmaları, veri kümesindeki nesnelerin hangi kategoriye ait olduğunu tanımlar. Bu nedenle, kategoriler genellikle sınıflarla ilişkilidir. Sınıflandırma problemlerini çözerek çeřitli sorunları çözebilirsiniz:



- İkili sınıflandırma problemleri:

- Bu aday deđiřecek mi, deđiřmeyecek mı?
- Bu e-posta spam mı, deđil mi?
- Bu işlem sahte mi deđil mi?

- Ve ayrıca, çoklu sınıflandırma problemleri:

- Bu daire New York'ta mı, San Francisco'da mı, Boston'da mı?
- Resimde neler var: Bir kedi, bir köpek veya bir kuş?
- Bu müşterinin hangi ürünü alması daha muhtemeldir: bir dizüstü bilgisayar, masaüstü veya akıllı telefon?

Bir başka oldukça spesifik sınıflandırma görevi tipi anomali tespittir. Genellikle bir tür sınıflandırma olarak kabul edilir, çünkü anormallik tespitinin amacı, normal dağılımında görünmeyen verilerdeki olağandışı nesnelerin aykırı değerlerini bulmaktır. Ne tür problemleri çözebilir:

- Veri setimizde farklı müşteriler var mı?
- Banka müşterilerimiz arasındaki olağandışı davranışları tespit edebilir miyiz?
- Kayıtlara göre bu hasta diğerlerinden farklı mı?



3.2 Kümeleme Analizi(Clustering Analysis)

Düzenli sınıflandırma ve kümelenme arasındaki temel fark, algoritmanın, önceden tanımlanmış sınıflar olmadan kümelerdeki gruplara uygulanmasıdır. Başka bir deyişle, bölünmenin ilkelerine insan rehberliği olmadan karar vermelidir. Küme analizi genellikle, bir dakika içinde konuşacağımız denetimsiz öğrenme tarzında gerçekleştirilir. Kümeleme aşağıdaki sorunları çözebilir:

- Demografi ve davranışlarını göz önünde bulundurarak müşterilerimizin ana segmentleri nelerdir?
- Bazı banka müşterilerinin temerrüt riskleri ile davranışları arasında bir ilişki var mı?
- İnsanların web sitemize ulaşmak için kullandıkları anahtar kelimeleri nasıl sınıflandırabiliriz?



3.3 Gerileme(Regresyon)

Regresyon algoritmaları, sınıflar yerine sayısal hedef değerleri tanımlar. Sayısal değişkenleri tahmin ederek, bu algoritmalar ürün talebini, satış rakamlarını, pazarlama iadelerini vb. tahmin etmede güçlü algoritmalar.

Örneğin:

- Bu üründen önümüzdeki ay kaç tane satabileceğiz?
- Uçak ile x yerinin uçuş ücreti ne kadar olacak?
- Bir aracın çalışma ömrünü sürdürmesi için en yüksek hız limiti ne olmalıdır?

3.4. Sıralama(Ranking)

Sıralama algoritmaları, diğer nesnelerle bağlantılı olarak nesnelerin (veya öğelerin) göreceli önemine karar verir. En iyi bilinen örnek, Google tarafından yoğun olarak arama motoru so-

nuçları sayfasındaki sayfaları sıralamak için kullanılan PageRank'tir. Facebook'ta bir haber akışında hangi yayınların diğerlerine göre daha fazla ilgi çekici olduğunu tanımlamak için sıralama algoritmaları da uygulanır. Başka hangi problemler sıralanabilir?

- Bu kullanıcının en çok hangi filmlerden hoşlanacağı?
- Bu müşteri için önerilen otellerin ilk listesi ne olacak?
- Ürünleri bir e-ticaret mağazasının arama sayfasına nasıl sıralamalıyız?

3.5. Generasyon(Üretken Ağlar)

Metin, resim veya müzik oluşturmak için Generation algoritmaları uygulanır. Bugün, fotoğrafları sanat tarzı görüntülere dönüştüren Prisma, insan konuşmasını taklit edebilen veya müzik besteleri yaratabilen WaveNet gibi uygulamalarda kullanılıyorlar. Üretken görevler, tahmine dayalı analitik çözümlerinden ziyade toplu tüketici uygulamaları için daha yaygındır. Bu nedenle, bu tür bir makine öğreniminin eğlence yazılımı için büyük bir potansiyeli var. Üretken algoritmaların olası görevleri nelerdir?

- Fotoğraflarınızı özel stil resimlerine ekleyin.
- Mobil sesli asistanlar için (örneğin Google asistanı) metinden konuşmaya uygulamalar oluşturun.
- Bir stille veya belirli bir müzisyeni hatırlatan müzik örnekleri oluşturun.

Bu görevleri yerine getirmek için farklı model eğitim yaklaşımları (veya eğitim stilleri) kullanılır. Eğitim, tarihsel verilerdeki değerler arasındaki bağımlılıklara göre uyarlanmış belirli bir matematiksel model geliştirme prosedürüdür. Eğitimli bir model gelecekteki verilerdeki bu bağımlılıkları tanıyabilecek ve aradığınız değerleri tahmin edebilecektir. Bu yüzden, üç model eğitim tarzı vardır.

4. Model eğitim türleri

Eğitim stillerini seçmek, bulunması gereken hedef değerleri bilmenize bağlı olup olmadığı ile ilgilidir. Başka bir deyişle, hedef değerlerin zaten eşlendiği eğitim veri kümelerine sahip olabilirsiniz ve algoritmanın gelecekteki verilerde bu kesin değerleri tahmin etmesini istiyorsunuz.

Veya amacınız değerler arasındaki gizli bağlantıları bulmak olabilir. İkinci durumda, hem geçmiş veriler hem de gelecekteki veriler için hedef değerler bilinmemektedir. Amaca yönelik olan durumlar arasındaki bu fark, eğitim stili seçimini etkiler ve seçtiğiniz algoritmaları tanımlar.

4.1. Denetimli öğrenme(Supervised Learning)

Denetimli öğrenme algoritmaları, zaten hedef değerleri olan tarihsel verilerle çalışır. Eğitim veri setlerinde bu hedef değerlerin haritalanması etiketleme olarak adlandırılır. Başka bir deyişle, insanlar algoritmaya hangi değerleri arayacağınızı ve hangi kararların doğru ya da yanlış olduğunu söyler. Bir etikete başarılı bir tahmin örneği olarak bakarak, algoritma bu hedef değerleri gelecekteki verilerde bulmayı öğrenir. Günümüzde, denetlenen makine öğrenmesi, genel olarak eğitim veri setlerinde zaten hedef değerler mevcut olduğundan, sınıflandırma ve regresyon problemleri ile aktif olarak kullanılmaktadır.

Bu, denetimli öğrenmeyi iş dünyasında kullanılan en popüler yaklaşım haline getirir. Örneğin, potansiyel müşterinin dönüşüm olasılığını tahmin etmek için ikili sınıflandırmayı seçerseniz, hangisi müşteriyi tekrar kazanacağız hangi müşteri artık bizi tercih etmez bunları tahmin edebiliriz. Hedef değerleri etiketleyebilir (dönüştürülmüş / dönüştürülmemiş değil veya 0/1) ve daha fazla model eğitebiliriz. Denetimli öğrenme algoritmaları, resimlerdeki nesnelerin tanınmasında, sosyal medya paylaşımlarının ruh halinin tanımlanmasında ve sayısal değerleri sıcaklık, fiyatlar vb. Olarak tahmin etmede kullanılır.

4.2. Denetimsiz öğrenme (Unsupervised Learning)

Denetimsiz öğrenme, hedef değerleri etiketli olmayan verileri organize etmeyi amaçlar. Makine öğreniminin amacı, bu durumda, değerlerdeki kalıpları tanımlamak ve nesneleri benzerlik veya farklılıklara göre yapılandırmaktır. Sınıflandırma görevleri alanında, denetimsiz öğrenme genellikle kümeleme algoritmaları, anomali tespiti ve üretken görevlerle uygulanır. Bu modeller öğeler arasında gizli ilişkilerin bulunmasında, segmentasyon problemlerinin çözülmesinde kullanılır.

Örneğin, bir banka müşterileri çoklu gruplara ayırmak için denetimsiz öğrenmeyi kullanabilir. Bu, her bir grupta ilgilenmek için özel talimatlar geliştirilmesine yardımcı olacaktır. Denetimsiz öğrenme teknikleri, bireyselleştirilmiş öneriler sunmak için sıralama algoritmalarında da kullanılır.

4.3. Takviye öğrenme (Reinforcement Learning)

Takviye öğrenme belki de oyun teorisi ve davranış psikolojisinden ilham alan en karmaşık makine öğrenim tarzıdır. Bir ajan (bir algoritma) girdi verilerine dayanarak kararlar vermek ve daha sonra bu kararların ne kadar başarılı olduğuna bağlı olarak “ödüllendirilmeli” veya “cezalandırılmalıdır”. Temsilci, “ödülleri” ve “cezalar” ile yüzleşerek kararlarını değiştirir ve kademeli olarak daha iyi sonuçlar almayı öğrenir.

Takviye öğrenme teknikleri aktif olarak robotik ve AI gelişiminde kullanılır. DeepMind tarafından iyi bilinen bir AlphaGo algoritması, tüm olası tahta kombinasyonlarını numaralandırmak yerine, eski Go oyununda en verimli hareketleri tahmin etmek için takviye öğrenimini kullandı. İddiaya göre, güçlendirilmiş öğrenme Tesla otopilotu tarafından denetimli öğrenme teknikleri ile birlikte kullanılıyor. Stil, otomatik pilot açıkken ve bir sürücü kararlarını düzelttiğinde kullanılır.

Bununla birlikte, iş alanında, çoğu algoritmanın yalnızca değişmeyen kurallar, hedefler ve dünya koşulları çerçevesinde başarılı bir şekilde öğrenebileceği için, takviye öğrenmenin uygulanması zordur. Bu nedenle, günümüzde yapılan modern takviye öğrenme gelişmelerinin birçoğunun, bu üç parametrenin kararlı olduğu Go veya eski Atari oyunları gibi oyunlara bağlı olmasının nedeni budur. Takviye öğrenmenin diğer bir sorunu öğrenme döngülerinin uzunluğudur. Oyunlarda, ilk karar ile elde edilen puanlar arasındaki süre nispeten kısadır, gerçek hayatta ise kararın ne kadar başarılı olduğunu tahmin etme zamanı haftalar alabilir.

5. Makine öğrenmeye nereden başlamalıyız

Tahmini analizler ve makine öğrenmesi çoğu işletme için hâlâ bilinmemektedir. Makine öğrenmesi araçlarının gelişimi etkileyici görünse de, işletme değerini yakalamak yine de bir zorluktur. Şirketler, yetenek edinme engelleri, iç liderlik güçlükleri ve son olarak, en önemlisi aşırı düzenlenmiş kurum kültürünün sertliği konusunda tam olarak ne yapacaklarını bilememektedirler. Birçok medya kanalında aktif olan büyük verilerin büyük potansiyeli hakkında teori yapmak nispeten kolaydır, ancak gerçek şu ki, büyük veriye yatırım yapmayı planlayan şirketlerin sayısının 2016'da yüzde 31'den 25'e düştüğünü göstermektedir. Büyük verilere yapılan yatırım genellikle büyük oyuncular sayesinde gerçekleşir. Bu, rekabet açığının yalnızca daha küçük veya daha az esnek işletmeler için arttığı anlamına gelir.

Bu bölümde, bu engellerin üstesinden gelmek ve rekabete uyum sağlamak için yönetici düzeyinde yapılması gereken en kritik kararlardan bahsedeceğiz.

5.1. “İşletme Yöneticisi” ve Vizyon sahibi Patron

Doğru analitik ve veri bilimi liderliği, veri odaklı kültür elde etmenin önündeki en büyük engeldir. McKinsey Global Institute anketine göre, şirketlerin yüzde 45'i veri ve makine öğrenimi için doğru vizyon ve stratejiyi belirlemeye çalışıyor. Bunu göz önünde bulundurun -bu alanda yetenek kazanmanın zorlukları iyi bilinmelidir: veri bilimi çalışanı hem tazminat hem de elde tutma açısından az ve pahalıdır. Bir veri bilimcisi bulmak zor olsa da, ankete göre bir analitik lider bulmak daha da zordur. İronik olarak, bu rol verimli veri süreçlerini yönetmek için kritik öneme sahiptir. Bazı makine öğrenme uygulamalarını tanıtabilseniz ve analitik bir lider olmadan birkaç eksik bağlantıyı telafi edebilseniz de, bu yaklaşım proaktif olmak yerine duyarlı olmaya yöneliktir.

Bu “işletmenin tercümanı” veya baş analitik sorumlusu (CAO) becerisi, işletme değerleri ve veri bilimi kabiliyetleri arasında köprü kurmak için çok disiplinlidir. Kişi liderlik etmeli ve bilgi teknolojisi departmanı, veri bilimi, pazarlama ve paydaşların bir veri stratejisi oluşturma ve geliştirme çabalarını uzlaştırmalıdır.

Bir analitik liderin bir diğer önemli görevi vizyoner rolünü yerine getirmesi gerektiğidir. Yeni veri bilimi araştırma çalışmalarında yaygın olarak kabul edilmeden önce iş uygulama potansiyelinin öngörülmesini gerektirir. Son zamanlarda iş taleplerini karşılayan makine öğrenme tekniklerinin çoğu, veri biliminde on yıllardır bilinmektedir. Ve çoğu - takviye öğrenmesi gibi - uygulamalarını DeepMind gibi önde gelen laboratuvarların ötesinde henüz yeni şeyler bulamadı. Bu gelişmeleri erken yakalayarak ve bunları işletmeye dönüştürmenin yollarını bularak bu alanda organizasyonunuzu rekabetin önünde tutabilirsiniz.

Ancak, analitik uzmanı kazanımı basit olmayacaktır. Üst düzey analitik pozisyonlarına olan talep ile yetenek arzı arasındaki mevcut uyumsuzluk 5:1'dir. Yani mevcut durumda ilgili pozisyonlar için çalıştırılabilecek aday sayısı çok fazla değil. Ve eğer işe alım başarısız olursa, yani bir veri bilimci istihdam edemezseniz bunun için ilgili bölümlerde dahili olarak bir analitik uzmanı bulmak ve eğitmek anlamına gelmektedir. Bu durumda en uygun fırsat, hem teknik hem de etki alanı iş geçmişine sahip bir kişiyi istihdam etmektir. Bazen, bu rol, kuruluşun büyüklüğüne bağlı olarak bir baş teknoloji görevlisi, yönetime giren bir veri bilimcisi, hatta bir baş yürütme görevlisi tarafından elde edilebilir.

5.2. Veri odaklı İşletmeler

Tek başına bir veri bilimcisi yalnızca verimli şirket ortamında etkili olabilir. Bir makine öğrenimi girişiminin tanıtılması tüm örgütsel seviyelerde desteklenmeli ve anlaşılmalıdır. Her yeni teknolojinin gelmesiyle birlikte, sadece eğitim gerekli değil, aynı zamanda değişimin değişmesi için de çok çaba sarfedilmesi gerekiyor. Makine öğrenimini karar vermeye destek olarak veya önemli kararlar vermek için kaldıraç olarak kullanmayı planlıyorsanız, büyük olasılıkla bu düşünme biçimi makul bir dirençle karşı karşıya kalacaktır. İnsanlar sezgilerine ve deneyimlerine dayanarak karar vermeye alışmakta ve bu da onları yalnızca profesyonel yapmıştır. Bir analiz lideri (veya CAO) ve diğer C düzeyi yöneticilerinin rolü, çalışanları eğitmek ve yeniliği teşvik etmektir. Bu, iletişim ve sunum becerilerinin bir veri bilimcisi için tercih edilen nitelikler olmasının nedenidir.

****Silinen veri:**** Departmanların yapısı gereği, veri odaklı bir organizasyon oluşturmaya önündeki başka bir engeldir. Verilere erişim aşırı düzenlenebilir veya bölümlerin kendileri topladıkları verileri içeride tutmak isteyebilir. Bu davranışla mücadele ederek, daha faydalı veriler elde etmede çok daha iyi sonuçlar elde edebilirsiniz.

****Anonimleştirilmiş veri.**** Bazen yasal olarak bankacılık veya sigorta gibi işletmelere yasal düzenlemeler yapılır ve veriler kolayca paylaşılamaz. Bu durumda, verilerdeki tüm değerler veri hazırlama aşamasında anonimleştirilmiş sayılara dönüştürülebilir. Böylece hassas iş veya müşteri detayları ortaya çıkmayacak ve veriler korunmuş olacaktır.

Sonuç olarak;

Bu makale ayrıntılı değildir ve yeni makine öğrenim inisiyatifiniz için bir uygulama rehberi olarak düşünülmemelidir. Keşfedilecek çok şey olsa da, bu yazıyı stratejinizi değerlendirmek için bir rehber olarak kullanabilirsiniz.

Günümüzde iş dünyası için temel sorun, bu stratejinin makine öğreniminin ve öngörücü analizlerin sağlayabileceği değişim hızına ayak uydurmak için nasıl ve ne zaman gerçekleştirileceğini anlamaktır. Modern iş kararları dönemi, topladıkları verileri kullanabilecekleri rekabeti en üst seviyeye çıkaracaktır.

Yenilikleri takip etmek ve ivedilikle hayata dahil etmek bize büyük kazanımlar sağlayacaktır.

Not: Bu Yazı ekteki linkten çeviri yapılarak oluşturulmuştur.

<https://www.altexsoft.com/whitepapers/machine-learning-bridging-between-business-and-data-science/>