**MAKİNE ÖĞRENMESİNE GİRİŞ**

**1 Makine Öğrenmesi nedir?**

‘Big Data’ çağındayız. Geçmişte sadece şirketler datalara sahipti. Buralar dataların saklandığı ve işlendiği yerlerdi. İlk önce kişisel bilgisayarların ortaya çıkışı ve sonrasında kablosuz iletişimin yaygınlaşması ile hepimiz birer data üreticisi olduk. Her aldığımız ürünle, her kiraladığımız arabayla, ziyaret ettiğimiz internet siteleriyle, blog yazılarımızla yada sosyal medyada paylaştığımız verilerle, hatta sadece yolda yürürken veya araba sürerken bile data üretiyoruz.

Ve bizler sadece onu üretenler değiliz, aynı zamanda onun tüketicileriyiz. Bize özel ürünler ve hizmetler istiyoruz. İsteklerimizin anlaşılmasını ve tahmin edilmesini istiyoruz.

Örnek vermek gerekirse, bir dev şirket internet sitesi üzerinden binlerce hatta milyonlarca müşterisine, 100 ün üzerindeki farklı ülkelerdeki depolarından ürün satıyor. Burdaki sihir şu ki burda yapılan her işlem saklanır ; Tarihler, müşteri id’leri, alınan mallar ve onların maliyetleri, toplam harcanan para ve daha birçok şey. Bu türden birçok bilgi günü gününe saklanır. Şirketin bunu yapmaktaki amacı hangi müşterisinin en çok hangi ürününü satın aldığını görmek, satışlarını en üst düzeye çıkarmanın yollarını aramak ve karı arttırmak. Basitçe her müşteri de kolayca ihtiyacı olan ürünleri bulmak ister.

Ancak bu iş belirgin değildir. Tam olarak hangi kişilerin çilekli dondurma sipariş ettiğini, bu yeni filmi görüp görmediğini, bu şehre daha önce gelip gelmediğini yada bu linki tıklayıp tıklamadığını bilemeyiz. Tüketici davranışları zamana ve coğrafi konuma göre değişir, ama biliriz ki bu tamamen random bir şekilde olmaz. İnsanlar süpermarketlere gidip rastgele bir şeyler satın almazlar. Bir bira aldıklarında, cips aldıklarında; yazın dondurma veya kış için mont aldıklarında… Burda belli davranış kalıpları söz konusudur.

Bu problemi çözmek için de bilgisayar algoritmalarına ihtiyaç duyarız. Algoritmalar inputları outputlara dönüştürecek bir talimatlar sırası olmalıdır. Örnek, bir sınıflandırma algoritması tasarlayalım. İnputlarımız bir sayı dizisi ve çıkış ise o sayılarla ilişkili sıralı bir liste olsun. Aynı işi yapmak için belki de birçok algoritma var ve biz en etkilisini arıyoruz, yani en az sayı yada hafıza alanı isteyen yada ikisi birden.

Ancak bazı işler için –her nedense- algoritmalara ihtiyaç duymayız. Müşteri davranışlarını tahmin etmek bunlardan biridir; bir başkası bir emailin spam oluğ olmadığını söylemek gibi. Girişin ne olduğunu biliyoruz, bir başlık bile birkaç karakterden oluşan bir metin. Çıkış da 2 tane; evet yada hayır. Ancak bu inputlardan output almayı bilmiyoruz. Bir mesajın spam olup olmadığını gösteren belirtiler zamandan zamana ve kişiden kişiye değişir.

Bilgi konusunda nerde eksiğimiz varsa, orda data içinde \*bir şeyler\* yaparız. Spam olduklarını bildiğimiz binlerce ve olmadığını bilmediğimiz binlerce emaili inceleriz ve burdan ne tür mesajların spam olabileceğini öğrenmeye çalışırız. Başka bir deyişle, bilgisayarın (makinenin) algoritmalarla bunu bu işi bizim yerimize otomatikçe yapmasını isteriz. Bir sayı dizisini sıralamayı öğrenmeye ihtiyacımız yok, bunun için hali hazırda birsürü algoritmamız var. Ancak bazı konular var ki nasıl yapılacağını bilmiyoruz, ama onlarla ilgili elimizde çokça data var.

Belki işin nasıl yapılacağını net bir şekilde açıklayamayız, ama çok isabetli ve tahminler yapabilecek sistemler inşa edebileceğimize inanıyoruz. Evet, tahminler her şeyi açıklamaz, ama dataların kayda değer bir kısmını hesaplamamızı sağlayabilir. İşin tam olarak nasıl yapılacağını tanımlayamamıza rağmen biz inanıyoruz ki hala belli kalıpları veya düzensizlikleri tespit edebiliriz. İşte makine öğrenmesi budur. Elimizdeki çokça data işin nasıl yapılacağını anlamamıza yardımcı olabilir, yada biz burdan elde ettiğimiz bilgileri tahminler yürütmede kullanabiliriz : Geleceği tahmin edelim, hatta ona yakın gelecek diyelim, dataların toplanmaya başladığı geçmişten çok da farklı olmayacak, ancak bu datalardan yapılan tahminlerin ise daha isabetli olmaya başlamalarını bekleyebiliriz.

Makine öğrenmesi yöntemlerinin büyük veritabanlarına uygulanmasına ‘Veri Madenciliği’ denir. Bi benzetme yapmak gerekirse Dünya’da bol miktarda maden var ve bir kısmı ham olarak çıkartılmış, bunlar işlendiklerinde az miktarda değerli maden ortaya çıkarır. Benzer şekilde veri madenciliğinde büyük miktarda veriden küçük ama kullanışlı bir model çıkartmaya çalışırız, yani yüksek başarılı bir tahmin modeli. Bunun sıklıkla uygulama alanları : perakende satışlar, bankaların geçmiş datalarına bakarak müşteriye verilecek krediyi gözden geçirmesi, dolandırıclık tespiti, market stoğu kontrol vs. Üretimde bu modeller optimizasyon, kontrol ve hata yakalama amaçlı kullanılır. Sağlık alanında programlar erken teşhis için kullanılır. Telekominikasyonda dataya göre analizle optimizasyon ve görüşme kalitesini maksimum yapmak. Bilimde; fizikte, astronomide, biyolojide bolca bulunan veri ancak hızlı makineler olan bilgisayarlarla analiz edilebilir. İnternet kocaman; ve durmaksızın büyüyor ve aranan bilgilerin manuel olarak yapılması yeterince efektif değil.

Ancak makine öğrenmesi sadece bir veritabanı problemi değildir, aynı zamanda yapay zekanın bir parçasıdır. Zeki olmak için, sistem öğrenebilecek şekilde değilebilir olmalıdır. Eğer sistem öğrenebilir ve değişimlere adapte olabilirse, bu sistemin tasarımcısı bundan sonraki olası her problemin çözümünü tek tek açıklamak zorunda değildir. Sistem her olası durum için çözümü öngörür.

Makine Öğrenmesi ayrıca bize görüş, dil işleme ve robotik konularındaki sorunları çözmemize yardımcı olur. Örnek olarak yüz tanımlamayı ele alalım: İnsanlar bu işi çok basitçe yapabilir, hergün ailemizin ve arkadaşlarımızın yüzünü gerçekte, fotoğrafta, farklı ışıklar altında, farklı saç stilinde ve farklı açıdan tanıyabiliriz. Ancak bunu nasıl yaptığımızı açıklayamayız. Çünkü becerilerimizi açıklayabilme yeteneğimiz yok, onları bir bilgisayar programı koduna dönüştüremeyiz. Aynı zamanda şunu da biliyoruz ki bir yüz rastgele piksellerden oluşmaz, bir yüzün belli bir yapısı vardır. Simetriktir; gözler, ağız, burun, yüzde belli bir bölgede konumlanır. Her insanın yüzü bu kalıplar dahilinde farklı kombinasyonlardır. Bir insan fotoğrafını analiz ederken, o kişiye özgü kalıbı yakalar ve tanımlar.

Makine öğrenmesi, bilgisayarların performansını geçmiş tecrübe ve birikmiş dataları kullanarak optimize eder. Bazı parametrelerle bir model tanımlarız ve öğrenme programı ‘öğrenme seti’ndeki yada geçmiş deneyimleri göz önüne alarak verilerin parametreleri optimize etmeye çalışır. Bu modeller geleceğe dair tahminler yapmaya veya datalardan bilgi edinmeye, yada ikisi için birden kullanılacak.

Makine öğrenmesi matematiksel modeller oluştururken istatistik biliminin teorilerini kullanır, çünkü temel görev bir örnekten çıkarım yapabilmektir. Burda bilgisayar biliminin rolü 2 kat daha çoktur, birincisi programın eğitimi için; optizimasyon problemlerini çözmek için verimli algoritmalara en az koca datalara sahip olmak kadar ihtiyacımız var. İkincisi, model bir kere öğrendiğinde olabildiğince verimli ve doğru sonuç vermelidir. Bazı uygulamalarda ise öğrenme etkinliği veya çıkarım algoritması, yani zaman ve alan kısıtlılığı, tahminin doğruluğu kadar önemlidir.

Şimdi makine öğrenmesinin tipleri ve uygulamaları hakkında daha detaylı bilgi edinmek için bazı uygulamaları üzerine tartışalım.

**1.2 Makine Öğrenmesinin Uygulama Örnekleri**

**1.2.1 Learning Associations**

Perakende satışlarda –örneğin bir süpermarket zincirinde- makine öğrenmesinin kullanımlarından biri sepet analizidir, insanların aldığı ürünler arasındaki ilişkileri yakalamaktır. Örneğin ‘X ürününü alan müşteriler, Y ürününü de aldı’. X ürününü alan bir müşteri Y ürününü almadıysa bile o artık bir potansiyel Y ürünü alıcısıdır. Bir kere böyle müşteriler bulduğumuz zaman, onlar artık bizim eşleştirmeli satış hedefimizdir(Y ürününü satmayı hedefleriz).

Bu tarz, ürünler arasındaki gibi bir ilişki yakaladığımızda, koşullu olasılık devreye girer ( P(Y|X) ). Bu ifadenin anlamı ‘X koşulu sağlandığında Y koşulunun sağlanma olasılığı’dır. Y ürünü X’e koşullu olduğu ilişkisi kurulduğunda, Y ürününü almış müşterilerin X ürününü çoktan aldığını tahmin edebiliriz.

Örnek olarak datalarımızdan yola çıkarak şunu hesaplamış olalım: P(cips | bira) = 0.7. Burdan şu kural çıkacaktır :

“Bira alan müşterilerin %70’i cips de aldı.”

Müşterilerimiz arasında buna göre ayrım yapmak isteyebiliriz, P(cips | bira, A) diyelim. A burda herhangi bir müşteri değişkeni (yaş, cinsiyet, medeni durum vb.) olsun. Eğer konu süpermarket yerine kitapçı olsaydı, ürünler kitaplar veya yazarlar olabilirdi. Bir web sitesi olsaydı ürün karşılığı linkler olurdu, insanların hangi linkleri daha çok tıkladığı ve bu bilgi kullanılarak çok kullanılan sayfaların erişimini hızlandırmayı hedeflerdik.