

Makine Öğrenmesi Nedir?

Makine Öğrenmesi, verilen bir problemi probleme ait ortamdan edinilen veriye göre modelleyen bilgisayar algoritmalarının genel adıdır.

Çok büyük miktarlardaki verinin elle işlenmesi ve analizinin yapılması mümkün değildir. Amaç geçmişteki verileri kullanarak gelecek için tahminlerde bulunmaktır. Bu problemleri çözmek için Makine Öğrenmesi yöntemleri geliştirilmiştir. Makine öğrenmesi yöntemleri, geçmişteki veriyi kullanarak yeni veri için en uygun modeli bulmaya çalışır. Verinin incelenip, içerisinden ise yarayan bilginin çıkarılmasına da Veri Madenciliği adı verilir.

Makine Öğrenmesi Tarihi

Makine öğrenmesi tarihi 1950 yılına dayanır. 1950 yılında Alan Turing tir, Turing testinin insan zekasını geçip geçemeyeceği üzerine bir algoritma tasarlar. Bu tasarımda bir bilgisayar botu insan zekasını geçmeyi başarmıştır. 1967’de en yakın komşuları algoritması ile ilk makine öğrenmesi oluşturuldu. 1980 ve 1990 yılında ise bu makine öğrenmesi uygulamalarının çok daha ileri taşınması için adımlar atılmıştır. Son 20 yılda ise verilerin artışıyla birlikte çok daha farklı alanlarda kullanılabilen, donanımlı ve gelişmiş algoritmaları bulunan makine öğrenmesi ürünleri tasarlandı. Bu dijital dönüşüm sayesinde gelecekteki yapay zekâ uygulamalarının önü açıldı.

Yapay Zekâ

Yapay zekâ (AI), basit bir ifadeyle, görevleri yerine getirmek için insan zekâsını taklit edebilen ve toplanan bilgilere dayanarak kendi performansını tekrar tekrar iyileştirebilen bir sistemdir. Yapay zekânın birçok farklı biçimi vardır.

- Sohbet robotları
- Akıllı Asistanlar
- Öneri motoru

Yapay Zekânın kalbi belirli bir biçim veya işlev değil, çok güçlü bir zihniyet ve veri analizi gücünü destekleyen bir süreç ve işlevdir. AI, dünyayı ele geçirecek yüksek performanslı insan benzeri bir robotun imajını yaratır, ancak insanların yerini alamaz. AI, insan yeteneklerini ve katkılarını büyük ölçüde geliştirmek için tasarlanmıştır. Bu nedenle çok önemli bir ticari varlıktır.

Makine Öğrenmesi Terimleri:

Tahmin: Veriden öğrenen modellerde sistem çıkışının nicel olması durumunda kullanılan yöntemlerin ürettiği değerlerdir.

Sınıflandırma: Giriş verisine ait çıkışların nitel olduğu durumlarda kullanılan yöntemlerin her veri örneğinin hangi sınıfa ait olduğunu belirlemesidir. Amaç probleme ait tüm uzayın belirli

sayıda sınıfa bölünmesidir. Sağdaki resimde her renk bir sınıfı göstermektedir. Sınıflandırma teknikleri sayesinde hiçbir veri örneğinin olmadığı bölgeler de renklendirilebilir.

Sınıflandırma örnekleri

Kümeleme: Geçmişteki verilerin sınıfları/etiketleri verilmediği/bilinmediği durumlarda verilerin birbirlerine yakın benzerliklerinin yer aldığı kümelerin bulunmasıdır.

Regresyon: Geçmiş bilgilere ait sınıflar yerine sürekli bilginin yer aldığı problemlerdir.

Özellik (Öznitelik) Vektörleri: Makine öğreniminde, özellik vektörleri, bir nesnenin özellikleri olarak adlandırılan sayısal veya sembolik özelliklerini matematiksel, kolay analiz edilebilir bir şekilde temsil etmek için kullanılır.

Özellik vektörleri, doğrusal regresyon gibi istatistiksel prosedürlerde kullanılan açıklayıcı değişkenlerin vektörlerinin eşdeğeridir.

Aşına olabileceğiniz bir özellik vektörü örneği, RGB (kırmızı-yeşil-mavi) renk açıklamalarıdır. Bir renk, içinde ne kadar kırmızı, mavi ve yeşil olduğu ile tanımlanabilir. Bunun için bir özellik vektörü $color = [R, G, B]$ olacaktır.

Entropi: Entropi, bilgisayar bilim dalında daha çok veri analizi, veri madenciliği, makine öğrenmesi ve sınıflandırma, yapay tahmin mekanizmalarında kullanılmaktadır. Özellikle bir veri setinde veri gruplarının bilgi kazancını hesaplanmasında entropi kullanılmaktadır

Özellik Vektörleri: Bir nesnenin özellikler adı verilen sayısal veya sembolik özelliklerini matematiksel olarak kolayca analiz edilebilir bir şekilde temsil etmek için kullanılır

Veri Madenciliği: Veri madenciliği, veri analizi yoluyla iş sorunlarının çözümüne yardımcı olabilecek kalıpları ve ilişkileri belirlemek için büyük veri kümelerini sıralama sürecidir. Veri madenciliği teknikleri ve araçları, işletmelerin gelecekteki eğilimleri tahmin etmesini ve daha iyi iş kararları almasını sağlar.

Veri madenciliği, işletmelerdeki başarılı analitik girişimlerinin temel bir bileşenidir. Çıktıları, geçmiş verilere bakan iş zekası (BI) ve gelişmiş analitik uygulamalarında ve oluşturulmakta veya toplanmakta olan akış verilerine bakan gerçek zamanlı analitik uygulamalarında kullanılabilir.