



## Veri – 21.Yüzyılın Petrolü



Veriler birçok yapay zeka çözümü ve uygulaması için çok merkezi ve önemli bir rol oynamaktadır.



Doğru verilerle makine doğru ve anlamlı şeyler yapmayı öğrenir.



Yanlış verilerle makine yanlış şeyler yapar.

# Makine öğrenmesi algoritmaları için eğitim verileri

| 2



Bu eğitim verileri doğru ve tutarlı olmalıdır. Alakalı ve anlamlı olmalıdır. Ne kasıtlı ne de kasıtsız olarak manipüle edilmeli, herhangi bir önyargı içermemelidir.



Bu algoritmaların sonuçlarının doğruluğu, kesinliği ve kullanışlılığı, doğrudan eğitildikleri verinin alaka düzeyine ve kalitesine bağlıdır.



Veriler kritik minimum miktarda mevcut olmalıdır. Ancak o zaman makine ve derin öğrenme algoritmaları tam etkinliğini ve gücünü ortaya çıkarabilir.



## Yapılandırılmış veri

- Bilgi ve Teknoloji sistemlerinde ve şirketin temel veri tabanlarında yapılandırılmış bir biçimde saklanan verilerdir.
- Veriler, tablonun satırları ve sütunları gibi yapılandırılmış formlarda görüntülenebilir. Örneğin, sayılar içeren bir excel sayfası.



## Yapılandırılmamış veri

- Kasıtlı olarak planlanmamış ve sistematize edilmemiştir.
- Tablo satırı ve sütunu gibi yapılandırılmış formlarda bile görüntülenemeyen veriler olabilir. Örneğin; e-postalar, videolar, resimler, ses kayıtları ve benzerleri.

## 1

### Dahili veri

- **Şirket içinde** bulunan, yani **şirketin kendi güvenlik duvarları içinde** bulunan veriler.
- Örneğin; müşterilerin verileri, ürünleri, süreçleri ve işlemleri.

## 2

### Harici veri

- Şirketler, **kendi güvenlik duvarlarının dışında**, çok sayıda ek dış verilere de erişebilir.
- Örneğin; sosyal medya ile ilgili veriler.



## Basit korelasyon

Gerçek bir açıklanabilir nedensel ilişki olmadan iki değişken arasında ilişki anlamına gelen basit bir istatistiksel korelasyon



## Gerçek nedensel ilişki

Nedenin etkiden kısmen sorumlu olduğu ve etkinin kısmen nedene bağlı olduğu iki değişken (neden ve etki) arasındaki ilişki

**Önyargı**, genellikle kapalı görüşlü, önyargılı veya haksız bir şekilde bir fikir veya şey lehine veya aleyhine orantısız ağırlıktır. Önyargı doğuştan ya da öğrenilmiş olabilir. İnsanlar bireye, gruba ya da inanca karşı ya da ona karşı önyargı geliştirebilirler. Bilim ve mühendislikte bir yanlılık, sistematik bir hatadır. **İstatistiksel önyargı**, bir popülasyonun haksız örneklenmesinden veya ortalama olarak doğru sonuçlar vermeyen bir tahmin sürecinden kaynaklanır.

Önyargılı verilerle bir algoritmayı **kasıtlı** veya **kasıtsız** olarak beslerseniz ve eğiterseniz, yapay zeka yanlış istatistiksel modelleri tanıyacak ve daha sonra **yanlış sonuçlara** varacaktır.



## Veri Bilimi

Verileri analiz etme, yapılandırma, temizleme ve işleme, aynı zamanda değerlendirme ve yorumlama sanatıdır.

