

# Veri duvarına çarpan yapay zeka: Akıllı sistemler neden tıkanıyor?

**Arş. Gör. Okan Şeker**

İKÜ Radyo, Televizyon ve Sinema Bölümü

Yapay zeka (AI), günümüzde birçok alanda kendine yer edinmiş, hızla gelişen bir teknoloji olarak hayatımızın ayrılmaz bir parçası haline gelmiştir. Ancak, bu ilerlemenin temelinde büyük miktarda veri ve bu verilerin işlenmesi yatmaktadır. Yapay zeka sistemleri, verileri analiz ederek öğrenir, kararlar alır ve tahminlerde bulunur. Fakat veriyle ilgili yaşanan sorunlar, bu sistemlerin verimli çalışmasını engelleyebilir ve yapay zekanın çıkma-za girmesine neden olabilir. Bu yazıda, yapay zekanın veri konusunda karşılaştığı başlıca problemler ve bu problemlerin nasıl çözülebileceği ele alınacaktır.

## Kaliteli veri

Yapay zeka sistemlerinin başarısı, kullanılan verinin kalitesine bağlıdır. Kaliteli verilerle beslenen sistemler, daha doğru ve güvenilir sonuçlar üretebilirken, düşük kaliteli veriler hatalı tahminlere ve yanlış yönlendirmelere yol açabilir. Veri kalitesi ile ilgili en büyük sorunlardan biri eksik veridir. Veri setlerinde bazı bilgilerin eksik olması, modelin doğru öğrenmesini engelleyerek yanlış genellemelere neden olabilir.

Örneğin, bir sağlık tahmin modeli, eksik hasta verileri nedeniyle hatalı teşhisler yapabilir, daha sonraki yönlendirmeleri de buna bağlı olarak yanlış devam edebilir. Bir diğer önemli kalite sorunu gürültülü veridir. Gürültülü veri, hatalı veya alakasız bilgiler içeren veri setleri anlamına gelir. Yapay zeka modelleri, bu tür verilerle eğitildiğinde yanlış sonuçlar üretebilir. Özellikle sosyal medya analizleri gibi büyük ölçekli veri işleme süreçlerinde, yanlış etiketlenmiş veya bağlam dışı veriler modele zarar verebilir. Ya da modelin sonuca ulaşmasını geciktirebilir.

Ayrıca, veri yanlılığı (bias), yapay zeka sistemlerinin adaletsiz veya taraflı kararlar vermesine neden olabilir. Eğer eğitim verisi belirli bir grup lehine veya aleyhine önyargılar içeriyorsa, model bu yanlılığı öğrenerek benzer hatalar yapabilir. Örneğin, işe alım süreçlerinde kullanılan bir yapay zeka modeli, geçmiş verilerdeki cinsiyet veya ırk yanlılıklarını devam ettirebilir.

Yapay zeka modellerinin başarılı olabilmesi için yalnızca büyük miktarda veriye değil, aynı zamanda çeşitli verilere de ihtiyaç duyulmaktadır. Sınırlı veri, yapay zeka-

nın öğrenme sürecini kısıtlayarak modelin genelleme yapmasını zorlaştırabilir. Özellikle nadir olayları tahmin etmeye çalışan sistemler için yeterli miktarda veri toplanması zor olabilir.

Örneğin, doğal afet tahmini yapan bir model, geçmişte yaşanan benzer olayların sayısı az olduğu için yeterince güçlü tahminler üretemeyebilir. Bunun yanı sıra, dengesiz veri dağılımı da önemli bir sorundur. Veri setlerinde belirli bir sınıfa ait örneklerin aşırı derecede fazla veya az olması, modelin dengesiz öğrenmesine yol açabilir. Örneğin, bir sahte işlem tespit modeli, sahte işlemlerle ilgili çok az veri içeren bir veri setiyle eğitildiğinde, gerçek

sahtekarlıkları tespit etmede başarısız olabilir. Bu nedenle, veri setleri dengeli olmalı ve tüm olası senaryoları kapsayacak şekilde çeşitlendirilmelidir.

Yapay zeka sistemleri, büyük miktarda kişisel ve hassas veri kullanmaktadır. Ancak, veri ihlalleri ve gizlilik endişeleri, bu teknolojilerin kullanımını sınırlandırabilir. Veri ihlalleri, kötü niyetli kişilerin sistemlerden hassas bilgileri ele geçirmesi anlamına gelir ve bu durum hem bireyler hem de kurumlar için büyük riskler oluşturur.

Ayrıca, kullanıcılar yapay zekaya veri sağlama konusunda giderek daha fazla temkinli davranmaktadır. Gizlilik endişeleri, kişisel verilerin kötüye kullanılabilmesi veya izinsiz paylaşılabilmesi korkusuyla giderek artmaktadır. Bu tür endişeler, yapay zeka sistemlerinin veri toplama süreçlerini zorlaştırabilir ve kullanıcıların güvenini sarsabilir.

## Farklı kaynak verileri

Farklı kaynaklardan elde edilen verilerin bir araya getirilerek işlenmesi, yapay zeka sistemlerinin başarısı için kritik bir unsurdur. Ancak, farklı veri formatları, veri entegrasyonunu zorlaştırabilir. Örneğin, bir hastane sisteminde hasta bilgileri farklı departmanlarda farklı formatlarda saklanıyor olabilir. Bu durum, verilerin tek bir yapay zeka modeli tarafından işlenmesini zorlaştırır. Bununla birlikte, veri uyumsuzluğu da yapay zeka sistemlerini zorlayan bir diğer faktördür. Aynı bilginin farklı kaynaklarda farklı biçimlerde temsil edilmesi, modelin doğru sonuçlar üretmesi-

ni engelleyebilir. Örneğin, müşteri bilgileri içeren bir veri setinde, aynı müşterinin bilgileri farklı sistemlerde

farklı formatlarda saklanabilir ve bu da analiz süreçlerinde hatalara neden olabilir.

Yapay zeka sistemleri, büyük miktarda veriyi işlemek ve depolamak için güçlü bir altyapıya ihtiyaç duyar. Veri depolama sorunları, yapay zeka sistemlerinin verileri etkili bir şekilde işlemesini engelleyebilir. Büyük veri kümelerinin depolanması ve yönetilmesi, güçlü donanım ve yazılım gerektirir. Ayrıca, veri işleme kapasitesi, yapay zeka sistemlerinin performansını doğrudan etkileyen bir faktördür. Veri işleme kapasitesinin yetersiz olması, sistemlerin yavaş çalışmasına veya büyük veri kümeleriyle başa çıkamamasına neden olabilir.

## Hatalı veriler temizlenmeli

Yapay zekanın veri kaynaklı çıkmazlarını aşmak veri kalitesini artırmak, eksik ve hatalı verilerin temizlenmesini içerir. Yanlılıkları tespit etmek ve gidermek için özel algoritmalar geliştirilebilir. Veri çeşitliliğini artırmak, modellerin daha geniş veri setleriyle eğitilmesini sağlayarak genelleme yeteneklerini artırabilir.

Özellikle dengesiz veri dağılımı olan durumlarda, sentetik veri üretme teknikleri kullanılabilir. Veri gizliliği ve güvenliği için güçlü şifreleme ve anonimleştirme yöntemleri uygulanmalıdır. Kullanıcı verilerinin korunması için sıkı güvenlik önlemleri alınmalı ve düzenleyici çerçevelere uygun hareket edilmelidir. Veri entegrasyonu ve uyumluluk sorunu ise farklı veri kaynaklarının standart formatlara dönüştürülmesi ile sağlanabilir. Veri işleme altyapısı ise, büyük veri analitiği platformları kullanılarak güçlendirilebilir.

Yapay zeka, büyük miktarda veriye dayalı olarak çalışan bir teknolojidir. Ancak, veri kalitesi, çeşitliliği, gizliliği, entegrasyonu ve yönetimi konusunda yaşanan sorunlar, yapay zekanın gelişimini sınırlayabilir. Bu sorunların üstesinden gelmek için veri kalitesini iyileştirmek, çeşitliliği artırmak, güvenlik önlemlerini güçlendirmek ve altyapıyı geliştirmek gerekmektedir. Bu adımlar uygulandığında, yapay zeka sistemleri daha güvenilir ve etkili hale gelecektir.

