

# 1 Algoritmik Önyargı ve Yapay Zekâ

Algoritmik Önyargı ve Yapay Zekâ sistemlerin toplumsal etkileri, kaynaklarda hem olumlu potansiyelleri hem de beraberinde getirdikleri ciddi riskler açısından geniş bir yelpazede tartışılmaktadır. Yapay zekâ (YZ) teknolojileri, hayatımızı iyileştirme yetenekleri sunarken, özellikle **algoritmik önyargı** ve bunun yol açtığı sorunlar önemli endişe kaynaklarıdır.

Bu makaleyi Spotify'da sesli olarak dinlemek için podcast'ine bu linkten ulaşabilirsiniz.

## 1.1 Algoritmik Önyargının Tanımı ve Ortaya Çıkışı

**Algoritmik önyargı**, yapay zekâ sistemlerinin kullandığı öğrenme algoritmalarına ve girdilerine bağlı olarak elde edilen çıktılar üzerinde mevcut olan, olumlu veya olumsuz, önceden edinilmiş deneyimlere dayalı bir düşüncedir. Hukuki açıdan herkes kanun önünde eşit olsa da YZ alanında öğrenmenin geçmiş deneyimlere bağımlılığı, önyargıyı istatistiki değerlerin bir tezahürü olarak karşımıza çıkarabilmektedir.

Bu önyargılar genellikle YZ sistemlerinin tasarımı ve eğitimi sırasında ortaya çıkar. Şeffaf olmayan algoritmik karar alma süreçleri nedeniyle ("kara kutu" problemi olarak adlandırılır), hukuka aykırı sonuçların doğması muhtemeldir.

## 1.2 Algoritmik önyargının temel nedenleri şunlardır:

- **Veri Setlerindeki Yanlılıklar:** YZ modellerini eğitmek için kullanılan veri kümelerinin niteliği ve içerdiği önyargılar, doğrudan modelin sonuçlarını etkiler. Örneğin, erkeklerin kadınlardan daha fazla suç işlediğine dair verilerle eğitilmiş bir YZ sistemi, erkeklerin aleyhine bir yanlılık gösterebilir. Geçmiş işe alım verileri belirli bir cinsiyet veya etnik gruba yönelik önyargılar içeriyorsa, YZ sistemi bu önyargıları öğrenebilir ve gelecekteki kararlarında yansıtabilir.
- **İnsan Kaynaklı Etik Dışı Veriler:** Bir insan ürünü olan yapay zekânın tahmine dayalı karar alma süreçlerini hukuka aykırı bir yola sokabilecek etik dışı, toplumsal ahlak ve değer yargıları ile bezeli veriler insanlar tarafından YZ kodlarına ve veri setlerine yüklenebilir.
- **Hatalı Kategorizasyon ve Asılsız Bilgi İşleme:** Yapay zekâ bazen topladığı verileri yanlış kategorize edebilir veya yeniden ve sürekli öğrenme metotlarını kullanarak her türlü dijital mecradan kendisine veri akışı sağlarken asılsız ve yanlış bilgileri de doğruymuşçasına işleyebilir.
- **Dinamik Öğrenme Algoritmaları:** Derin öğrenme algoritmaları, girdileri kendi kendine ağırlıklandırdığı için önyargı oluşma riski taşır.

## 1.3 Algoritmik Önyargının Yol Açtığı Problemler

Algoritmik önyargı ve YZ'nin kontrolsüz kullanımı, toplumda çeşitli ciddi problemlere yol açmaktadır:

- **Hak İhlalleri ve Adil Olmayan Sonuçlar:**
  - Algoritmik karar alma süreçlerinin şeffaf olmayışı, kredi, istihdam, eğitim, sağlık, adalet gibi önemli konularda **ağır hak ihlalleri** yaratabilmektedir. Özellikle işe alım süreçlerinde YZ tabanlı sistemlerin cinsiyet, ırk veya yaş gibi faktörlere dayalı önyargılar sergilediği rapor edilmiştir.
  - Vergilendirme sürecinde kullanılan YZ teknolojisinin algoritmik karar alma yetisine bağlı olarak mükellefler üzerinde **ayırıcı ve önyargılı kararlar** alma ihtimali bulunmaktadır. Hollanda'daki çocuk yardımları vakası veya Avustralya'daki Robodebt vakası gibi örnekler, YZ kaynaklı mağduriyetleri göstermektedir.
  - Yapay zekânın insan gözetimsiz modelinin kullanılması, kişisel ve özel verileri izinsiz toplama, işleme, önyargılı çıkarımlar yapma ve yanlış profillemeye gibi hukuka aykırı sonuçlar doğurma ihtimali taşımaktadır.
- **Bilgi Ortamında Dezenformasyon ve Manipülasyon:**
  - Üretken YZ'nin yaygınlaşması, yanlış bilginin miktarını, kalitesini ve kişiselleştirilmesini artırma potansiyeline sahiptir. YZ araçları, **sahte haberler ve deepfake'ler** oluşturmak için kullanılabilir; bu tür içerikler, insan yapımı sahte içeriklerden ayırt edilmesi zor olabilir.
  - Dijital dünyada **sahte içerik üretimi ve dijital manipülasyon** önemli bir olumsuzluktur. Deepfake içerikler, gerçeğin bir kopyasını üreterek sahte bilgi sorununu daha karmaşık hale getirmektedir.
- **Gözetim Toplumu ve Veri Mahremiyeti İhlali:**
  - Dijital iletişim teknolojilerinin yaygınlaşmasıyla birlikte bireyler dijital ortamda iz bırakmakta ve **dijital gözetim araçlarına** maruz kalarak büyük verinin ham maddesine dönüşmektedir. Bu durum, gözetimin tarih boyunca toplumsal denetim aracı olmasından öte, dijitalleşme ile birlikte **daha sürekli ve kapsamlı** hale gelmesini ifade eden "gözetim toplumu" kavramıyla açıklanmaktadır.
  - Dijital gözetim amaçlı veri toplama faaliyetleri, **kışisel veri güvenliğini tehlikeye** atarak dijital ortamların güvenilirliğini sorgulanır hale getirmektedir. Algoritmalar tarafından gerçekleştirilen derin veri analizleri, şirketlere tüketicileri izleme ve analiz etme konusunda geniş fırsatlar sunarak **kışisel veri mahremiyetini tehdit eden** bir unsura dönüşmektedir.
- **Etik ve Güven Sorunları:**
  - Algoritmaların verilerle eğitilmesi ve bu verilerin yargılarını yansıtabilmesi, özellikle önyargılı verilerle eğitilen algoritmaların ırkçı ve cinsiyetçi önyargıları açığa çıkarabilmesine neden olabilmektedir. Bu durum, halkla ilişkilerde **algoritmik önyargı sorununu** ortaya çıkarmaktadır.
  - Yapay zekânın otomasyon yanlılığı veya taraftarlığı da bir sorun olarak karşımıza çıkabilir; insanlar YZ kararlarını sorgusuz uygulamaya eğilimli olabilir.

- Algoritmaların bir “kara kutu” gibi görünmez olması, halkla ilişkiler uygulayıcılarının süreci yönetirken ortaya çıkabilecek belirli bir eğilimi ya da etkiyi denetleme fırsatını azaltmaktadır. Algoritmik halkla ilişkiler, ahlaki körlüğe neden olabilir; karar süreçlerinde algoritmaların ön plana çıkması, halkla ilişkiler uygulayıcıları için “**ahlaki körlük**” riski taşımaktadır.
- Şeffaflık ve güven yapı taşları üzerine kurulan halkla ilişkiler süreci, algoritmaların yanıltıcı veya yanlış bilgi tabanlı analizlere açık hale gelmesiyle risk taşımaktadır.
- **İşgücü Piyasasındaki Dönüşümler:**
  - Algoritmaların kullanımının artmasıyla birlikte daha önce zaman alan işler kısa sürede yapılabilmekte, bu da **otomasyonlaşmanın artmasına** neden olabilmektedir.
  - Yeni teknolojilerle birlikte **yeni meslekler** de ortaya çıkabilir; örneğin, halkla ilişkiler ajanslarında artık yazılımcıların da görev aldığı belirtilmiştir.

## 1.4 Çözüm Yolları ve Öneriler

YZ'nin toplumsal etkilerinin olumlu yönde ilerlemesi ve algoritmik önyargı gibi sorunların ele alınması için çeşitli öneriler sunulmuştur:

- **YZ Okuryazarlığı ve Bilinçlendirme:**
  - YZ sistemlerini ve toplumsal etkilerini anlama, değerlendirme ve analiz etme yeteneği olan **YZ okuryazarlığının geliştirilmesi** hayati öneme sahiptir. Kamuoyunda YZ'nin temel prensipleri, veri ve doğal dil işleme konularında bilinçlendirme yapılmalıdır.
  - Yapay zekâ teknolojisinin net bir şekilde anlatılması, sınırlarının ne olduğu, gerçekte nasıl çalıştığı ve önüne konan bilgilere nasıl yaklaşılması gerektiği konularında insanların bilinçlendirilmesi önemlidir.
- **Şeffaflık ve Açıklanabilir YZ (Explainable AI - XAI):**
  - YZ'nin temeldeki matematiksel modellerinin nasıl çalıştığının ve hangi verilerle eğitildiğinin anlaşılması gerektiğini belirten “**Açıklanabilir YZ**” kavramı önemlidir.
  - Algoritmik karar alma süreçlerinin şeffaf olmayışı nedeniyle yaşanabilecek hak ihlallerini önlemek için YZ'nin tasarım ve uygulama aşamalarında **denetlenmesi gerekmektedir**.
- **Veri Kalitesi ve Çeşitliliği:**
  - YZ sistemlerinde veri meselesi ve verinin niteliği modellerin sonuçlarını doğrudan etkilediği için önemlidir. Algoritmik önyargıyı engellemek için, YZ sistemlerini eğitirken kullanılan veri setlerinin **çeşitliliği ve temsil gücü** artırılmalı ve algoritmaların adil kararlar vermesini sağlamak için sürekli denetim mekanizmaları geliştirilmelidir.
- **İnsan Faktörünün Merkeziliği ve Denetim:**
  - YZ temelli dezenformasyon tespiti ve doğrulama çalışmalarında **insan faktörünün merkeziliği** kamuoyuna doğru şekilde aktarılmalıdır. YZ içeren bir

iş akışı geliştirildiğinde, başta, ortada ve sonda mutlaka bir gazeteci ve editör bulunmalıdır.

- YZ denetiminde genel bir denetim sistemi henüz kurulmamış olsa da **teknik, etik ve hukuki temelli denetim** yaklaşımları önerilmektedir.
- **Üçüncü taraf denetimi**, YZ denetimine yeni bir perspektif getirmektedir. Bağımsız ve tarafsız olması nedeniyle etik temelli denetimlerde daha etkili olabileceği düşünülmektedir. Sivil toplum kuruluşları, dernekler, avukatlar, akademisyenler gibi temel hak ve özgürlüklere dair uzmanlığa sahip gruplar, üçüncü taraf denetçisi olarak önemli bir rol oynayabilir.
- **Regülasyon ve Hukuki Çerçeve:**
  - Regülasyon konusunun ciddiyetle ele alınması, ancak esas olarak okuryazarlık kavramının merkezde tutulması önerilmiştir. YZ teknoloji şirketlerini de içine katacak **watermark/etiket koyma zorunluluğu** gibi somut düzenlemeler yapılması gerekmektedir.
  - Mevcut yasal düzenlemelerin algoritmaların gelişim hızına ayak uyduramadığı ve bu durumun etik ihlal riskini artırdığı belirtilmiştir. Bu nedenle, **yasal düzenlemelerin güncellenmesi** ve veri güvenliği kriterlerinin artırılması kaçınılmazdır.
- **İş Birlikleri ve Ortaklıklar:**
  - Doğrulama aktörleri, teknoloji şirketleri ve akademisyenler arasında **açık iletişim kanalları** oluşturulmalı ve iş birlikleri geliştirilmelidir. Küresel teknoloji şirketlerinin yerel bağlamları gözeterak uygulama geliştirmeleri ve ortaklıklar kurmaları gerektiği vurgulanmıştır.

Sonuç olarak, yapay zekânın toplumsal etkileri geniş ve karmaşıktır. Algoritmik önyargı, dezenformasyon ve mahremiyet ihlali gibi ciddi riskler taşısa da doğru yaklaşımlarla (YZ okuryazarlığı, şeffaf algoritmalar, güçlü regülasyonlar, insan denetimi ve iş birlikleri) bu risklerin azaltılması ve YZ'nin topluma faydalı bir şekilde entegre edilmesi mümkündür.

## 1.5 Anahtar Terimler Sözlüğü

- **Yapay Zekâ:** İnsan zekâsını taklit eden sistemler olarak tanımlanan, işlenen veri ve kodlara bağlı olarak öğrenebilen, analiz edebilen ve karar alabilen dijital araçlar.
- **Üretken Yapay Zekâ:** Kitlelerin kullanımına sunulan ve yeni içerikler üretebilen YZ araçları.
- **Dezenformasyon:** Yanlış veya yanıltıcı bilginin yayılması. YZ'nin bu alandaki etkisi tartışma konusudur.
- **Derin Sahte:** Yapay zekâ desteğiyle üretilen, genellikle bir kişinin yüzünün veya sesinin başka birinin üzerine yerleştirilmesiyle oluşturulan gerçekçi ancak sahte içerikler.
- **Algoritma:** Başlangıçta matematiksel problemlerin çözümü için kullanılan, daha sonra bilgisayar teknolojilerinde yaygınlaşan, belirli bir sürecin ilerleyişini açık ve kesin bir

şekilde belirleyen, karmaşık girdileri anlamlı çıktılara dönüştüren sonlu talimatlar dizisi.

- **Büyük Veri:** Geleneksel araçlarla işlenemeyen, yüksek hacimli, çeşitli ve hızlı veri kümeleri. Yönetim ve halkla ilişkiler gibi birçok alanda karar alma süreçlerinde kritik içgörüler sağlar.
- **Algoritmik Halkla İlişkiler:** Algoritmaların halkla ilişkiler süreçlerine dahil olmasıyla veri analizi, dijital medya takibi, kriz iletişimi gibi uygulamaların algoritmik teknolojiler aracılığıyla yürütüldüğü, hedef kitlelerin algoritmik gözetime açık hale geldiği yeni bir halkla ilişkiler yaklaşımı.
- **Dijital Gözetim:** Bireylerin dijital ortamda bıraktıkları ayak izleri aracılığıyla sağlık bilgileri, telefon ve internet kullanımı, kredi kartı işlemleri gibi kişisel verilerin izlenmesi.
- **Algoritmik Önyargı:** Yapay zekâ sistemlerinin tasarım ve eğitimi sırasında ortaya çıkan, kullanılan öğrenme algoritmalarına ve girdilerine bağlı olarak elde edilen çıktılar üzerinde mevcut olabilen, adil olmayan veya ayrımcı sonuçlara yol açan eğilimler.
- **Kara Kutu:** Yapay zekâ modellerinin kararlarını hangi veri setleri ve kodlara dayanarak, hangi gerekçelerle aldığına dair şeffaflığın olmaması durumu.
- **Otomasyon Yanlılığı/Taraftarlığı:** İnsanların yapay zekânın hatasız ve insan aklının üzerinde kararlar alabileceğine inanarak, YZ kararlarını sorgusuzca uygulaması ve sürekli takdir etmesi.
- **Üçüncü Taraf Denetimi:** Bir şirketin veya kurumun kendi bünyesi dışındaki bağımsız kişi veya gruplar tarafından yapılan denetim. Özellikle YZ'deki algoritmik önyargı ve etik sorunların denetiminde önemli bir rol oynar.
- **Veri Okuryazarlığı:** Veriyi anlama, analiz etme ve kullanma becerisi; liderler ve çalışanlar için dijital dönüşümde kritik bir yetkinlik.
- **Regülasyon:** Yeni teknolojilerin, özellikle yapay zekânın kullanımıyla ilgili yasal düzenlemeler ve standartlar belirleme süreci.
- **Aşırı Öğrenme:** Makine öğrenmesi algoritmalarında, modelin eğitim verisine çok fazla uyum sağlayarak yeni verilere genelleme yeteneğini kaybetmesi durumu, istisnai durumlarda hatalı sonuçlar üretebilir.