

Yapay Zekâda Cinsiyet Önyargısı

Burcu Çırtlık

ÖZET

Bu yazıda geliřmekte olan Yapay Zeka (AI) sistemlerinin doęasında bulunan cinsiyetçi önyargıları analiz ediliyor. Algoritmik önyargı üzerine feminist çalıřmalardan yola çıkılarak, AI gelişiminin mevcut ataerkil yapıları yansıttığı ve güçlendirdiğı ileri sürülüyor. Analiz, cinsiyetçi stereotipleri güçlendiren sohbet robotlarından beyaz erkek bakışını sürdüren sanat jeneratörlerine kadar yapay zeka uygulamalarının sosyal eşitsizlikleri nasıl daha da kötüleřtirmekle tehdit ettiğini vurguluyor.

ANAHTAR KELIMELER

Algoritmik Cinsiyet Önyargısı, AI Tasarımı, Sosyal Eşitsizlik, Yapay Zeka



New articles in this journal are licensed under a Creative Commons Attribution 4. United States License.



This journal is published by [Pitt Open Library Publishing](#).

YAPAY ZEKÂDA CİNSİYET ÖNYARGISI

Giriş

Yapay Zekâ, insanoğlunun daha önce hiç karşılaşmadığı bir hızla toplumumuzu şekillendiriyor, dönüştürüyor. Bu tür bir hızla başa çıkmak, hem toplum hem de bireyler için hiç kolay bir durum değil. Yapay zeka şimdilik kişisel gelişimimize ve gündelik işlerimize yardımcı oluyor olabilir, ancak bunun bir bedeli var. Yapay zeka, cinsiyetçi ve ırkçı bir beyaz erkek olarak evrimleşiyor!

Bu dönüşümdeki temel endişelerden biri, yapay zeka teknolojileriyle birlikte gelen önyargıdır. Yapay zeka sistemlerinde var olan cinsiyet ve ırk önyargısı ciddi bir sorun haline gelmekte. Eğer şimdi harekete geçmezsek, çok yakın gelecekte bunların yansımalarını daha şiddetli bir şekilde göreceğiz.

Pek çok kişinin halihazırda kullandığı sohbet robotları ve görseller üreten yapay zeka araçları, şimdiden açık tenli erkekleri ön plana çıkarıyor veya iş rollerinde stereotipleri sürdürerek toplumsal önyargıları yansıtıyor ve güçlendiriyor.

Teknoloji ve eşitliğin kesişimini kuvvetle hissettiğimiz bu dönemde, daha eşitlikçi bir dijital geleceğe ulaşabilmek için, özellikle kadınların ve azınlık grupların bu konuda farkındalığının artması son derece önemli.

Yapay Zekada Cinsiyet Önyargısı Nedir?

Günlük işlerinize yardımcı olan küçük bir robot arkadaşınız olduğunu hayal edin. Bu robot, internetteki kitaplar, makaleler, görsel içerikler vb. materyallerle eğitildi. Yani öğrendiklerinin çoğu, kısa zamana kadar (ve hala belli bir oranda) oldukça eril bir yer olan internetten geliyor. Cinsiyet önyargısı ve ırk önyargısı işte burada ortaya çıkıyor.

Diyeim ki bu robotu işe alım süreçlerinde kullanmak istiyorsunuz. Bu küçük arkadaşımızın mühendislik rolleri için beyaz erkekleri istihdam etme ihtimali oldukça yüksek. Çünkü şimdiye kadar bu rollerde istatistiksel olarak daha fazla beyaz erkeğin işe alındığını görüp, beyaz bir erkek olmanın iş için gerekli bir kalifikasyon olduğuna kanaat getirebilir. Farkında olmadan mühendislik işleri için erkekleri, hemşirelik işleri için de kadınları tercih edebilir; bunun nedeni adaletsiz olmaya çalışması değil, eğitildiği modelin bu olmasıdır. Bu durum yıllar önce gerçek hayatta karşımıza çıktı, ama buna birazdan detaylıca değineceğiz.

Yapay zeka önyargısının, toplumumuzdaki mevcut cinsiyetçilik ve ırkçılıktan çok da farklı olmadığını düşünebilirsiniz. Yine de yapay zeka bu önyargıları sistematik olarak artırma konusunda oldukça büyük bir potansiyele sahip. Yapay zekaya, hesap makinelerine güvendiğimiz gibi güvenmeye başladığımız anda, tüm bu önyargılar dünya çapında katı bir forma bürünebilir.

İşe Alım Sistemlerinde Önyargı

Gercek Hayattan Örnekler

Amazon şirket, kadınlara karşı önyargı gösterdiği için bir yapay zeka işe alım aracını durdurdu. Sistem, on yıl boyunca şirkete gönderilen ve çoğunlukla beyaz erkeklerden gelen özgeçmişler üzerinden eğitildi ve bu özgeçmişlerin çoğu da beyaz erkekler aitti. Sonuç olarak, yapay zeka beyaz bir erkek olmanın önemli bir beceri olduğunu varsaymaya başladı. "Kadın voleybol kaptanı" gibi içerisinde "kadın" kelimesi geçen özgeçmişlerin notunu düşürdü. Bu durum fark edilince şirket projeyi rafa

kaldırdı. [Amazon](#) bu sistemi gerçek hayatta hiç kullanmadığını iddia etti ama kim bilir?

Bir Adım İleri

Şu anda, işe alım süreçleri için üretilmiş pek çok yapay zeka programı piyasada mevcut. Şirketler, yapay zeka çözümlerini insan kaynakları sistemlerine entegre etmeye başladılar bile. Herhangi bir şirket, düzenleme ve kontrol olmaksızın işe alım sistemi için bir yapay zeka çözümü geliştirebilir. Dolayısıyla, bir dahaki sefere bir şirket tarafından reddedildiğinizde, pozisyona uygun mu değildiniz, yoksa bir yapay zeka algoritmalarının cinsiyet veya ırk önyargısına mı takıldınız bunu bilemeyeceksiniz. Hatta o şirkette işe alım süreçlerinde yapay zeka teknolojilerinin kullanılıp kullanılmadığından bile emin olamayacaksınız.

2015 yılında Carnegie Mellon Üniversitesi, Google reklamları üzerinde bir deney yaptı. Şans eseri, Google'ın hedefli reklam algoritmalarının yüksek maaşlı işleri erkeklerle gösterme eğiliminde olduğunu keşfettiler. Kadınlar 200.000 dolardan daha fazla ödeme yapan iş ilanlarını göremiyordu. [Bu durum, algoritmalarındaki eşitlik konusunda büyük bir soru işaretini oluşturdu.](#)

2021 yılında Güney Kaliforniya Üniversitesi, Facebook'un reklam dağıtım sisteminin, aynı niteliklere sahip olmalarına rağmen, kadın ve erkeklere farklı iş ilanları gösterdiğini keşfetti. İş ilanlarının dağıtımını geleneksel cinsiyet rollerini pekiştirme eğilimindeydi. [Algoritmalar erkeklerle mühendislik gibi işleri, kadınlara ise hemşirelik gibi işleri öneriyordu.](#)

Bir Adım İleri

Bu durum klasik insan kaynakları sürecinde bildiğimiz ve karşılaştığımız cinsiyetçiliğin ötesindedir. Bu önyargılı algoritmalar nedeniyle, yakın bir gelecekte, pek çok kadın iş ilanlarını görmediği için iş fırsatlarından haberdar bile olmayabilir.

Yüz ve Ses Tanıma Sistemlerinde Önyargı

MIT ve Stanford Üniversitesi araştırmacıları, piyasadaki en iyi yüz tanıma sistemlerinin önemli ölçüde cinsiyet ve ırk önyargısına sahip olduğunu keşfetti. Araştırmaya göre, beyaz erkekler için hata oranı %0,8 iken, koyu tenli kadınlar için aynı hata oranı %34,7 idi.

Bir Adım İleri

Yüz tanıma teknolojileri birçok alanda kullanılmakta olup, otomasyon ve robotik ile birlikte giderek daha popüler hale gelmekte. Bu sistemler kolluk kuvvetleri, ordular veya hükümetler tarafından kullanılıyor ve kullanılmaya devam edilecek. Yani yakın gelecekte, sırf koyu tenli bir kadın olduğunuz için ve algoritmalar sizi doğru tespit edemediği için yoktan yere tutuklanmanız ya da ceza almanız mümkün.

Dr. Tatman ve ekibinin bir araştırmasına göre, Google'ın ses tanıma sistemi erkekler için kadınlara göre %13 daha doğru çalışıyor. Google'ın sisteminin diğer rakiplerine kıyasla piyasadaki en hassas sistem olduğunu göz önüne alırsak, genele yayıldığında ses tanıma sistemleri kadınlar konusunda sınıfta kalmış [görünüyor](#). Örneğin, [Louise Kennedy, anadili İngilizce](#)

[olan İrlandalı bir konuşmacı, otomatize edilmiş bir İngilizce sınavında başarısız olmuştur.](#)

Bir Adım İleri

Günümüzde sesli komutlar, internet aramaları ve akıllı ev sistemleri gibi pek çok teknolojinin kontrolünde yaygın olarak kullanılıyor. Öyle görünüyor ki yakın gelecekte seslerimiz, elektronik sistemler için bir çeşit uzaktan kumanda işlevi görecek. Eğer bu sistemler kadınların sesini tanımakta güçlük çekerse, günlük hayatımızdaki cinsiyet ayrımcılığı hayal edebildiğimizden de ötesinde artabilir.

Finans Alanında Önyargı

2019 yılında Apple ve Goldman Sachs, bir kredi kartı programı başlattı. Çok geçmeden, David Heinemeier Hansson (Ruby on Rails'in yaratıcısı) Apple Kart'ı Twitter üzerinden cinsiyetçi olmakla suçladı: *"Eşim ve ben ortak vergi beyannamesi verdik, ortak mülkiyetli bir eyalette yaşıyoruz ve uzun süredir evliyiz. Ancak Apple'ın kara kutu algoritması benim, onun 20 katı kredi limitini hak ettiğimi düşünüyor," dedi David Heinemeier Hansson.*

Daha sonra, Steve Wozniak (Apple kurucu ortağı) bu iddiayı destekledi ve aynı şeyin kendi eşinin de başına geldiğini söyledi.

Bu şikayetlerin ardından New York Eyaleti düzenleyici kurulu bir soruşturma başlattı. Kurul raporlarını 2021 yılında yayınladılar ve herhangi bir ihlal tespit edilmediğini iddia ettiler. Müfettişler yaklaşık 400.000 başvuru verisini analiz etti ve kredi puanı sisteminin cinsiyetçi davrandığına dair hiçbir kanıt bulamadı.

Bir Adım İleri

Düzenleme kurulu Apple kart sistemlerinde cinsiyet ayrımcılığı bulamasa da, bu ciddi bir örnektir. Eğer kredi puanlama sistemi cinsiyetçi ise, bu, yapay zekanın neredeyse iki yıl boyunca kadınların kredi puanlarını düşürdüğü anlamına gelir. Bu finansal bir konu olduğu için sıkı düzenlemeler altındadır, peki ya düzenleme olmayan diğer alanlar?

Sağlık Sistemlerinde Önyargı

Araştırmacılar yakın bir zamanda karaciğer hastalıklarının tanımlamak için yapay zeka modelleri geliştirdiler ve %70 başarı oranı elde ettiler. Ancak model farklı cinsiyetlerle incelendiğinde, kadınlarda vakaların %44'ünün gözden kaçtığı [tespit edildi](#).

Bir Adım İleri

Yapay zekanın gelecekte bir erken tanı sistemi olarak kullanılması gündemde. Bu teknolojinin sağlık sistemlerine nasıl adapte edileceği konusunda pek çok araştırma yapılıyor. Cinsiyet önyargısı devam ederse, kadınların gelecekte otomatik sistemler tarafından yanlış teşhis edilme olasılığı oldukça yüksek olabilir. Örneğin, kalp krizi bulguları çoğunlukla kadın ve erkeklerde farklılık gösterir. Göğüs ağrısı ve sol kol ağrısı gibi geleneksel kalp krizi belirtileri, erkeklerde daha sık görülür. Öte yandan kadınlar, mide bulantısı, nefes darlığı ve sırt ağrısı gibi semptomlar göstermektedir. Bir yapay zeka teşhis aracının öncelikle erkek hastalardan alınan verilerle eğitildiğini varsayalım. Bu durumda, kadınlarda bu alternatif semptomları etkili bir şekilde tanımayabilir veya önceliklendirmeyebilir, bu da yanlış tanıya veya gecikmiş tedaviye yol açacaktır.

Yapay Zekada Önyargıya Ne Sebep Olur?

Veri Önyargısı: Yapay Zeka, çoğunlukla çevrimiçi olarak var olan verilerle öğrenir. Eğer verilerde önyargılar varsa, yapay zeka sistemi de önyargılı olacaktır. Bu durum "Çöp içeri, çöp dışarı" prensibi olarak tanımlanır. Amazon'un işe alım sisteminde gördüğümüz gibi, sisteminizi erkek egemen veya cinsiyetçi verilerle eğiterseniz, sisteminiz bu stereotiplere dayalı kararlar verecektir. Özellikle başlangıçta, İnternet oldukça erkek egemen bir platformdu ve uzun yıllar boyunca içerik yaratıcıların çoğunluğu erkeklerdi. Dolayısıyla İnternet, önyargılarla dolu, cinsiyetçi bir yer olarak gelişti. Günümüzde çoğu yapay zeka teknolojisi bu tarihsel önyargılı verilerle beslenmektedir.

Algoritma Önyargısı: Yapay zeka sisteminizi adil ve önyargılardan uzak bir şekilde eğitseniz bile, algoritmalarınızı önyargılara neden olacak şekilde manipüle etmeniz mümkündür. Bu durum kasıtlı yapılmasa bile, sistemleri tasarlayan ve geliştiren ekipler çoğunlukla erkek egemen bir yapıya sahip olduğundan, algoritmalar onların bakış açılarını ve tercihlerini yansıtmaya potansiyeline sahiptir.

Veri Çeşitliliği: Yüz veya ses tanıma için tasarlanan yapay zeka sistemleri, adil olmak için veri çeşitliliğine ihtiyaç duyar. Yüz tanıma sisteminizi daha çok erkek resimleriyle eğiterseniz, doğal olarak kadın yüzlerini tanımada başarısız olacaktır.

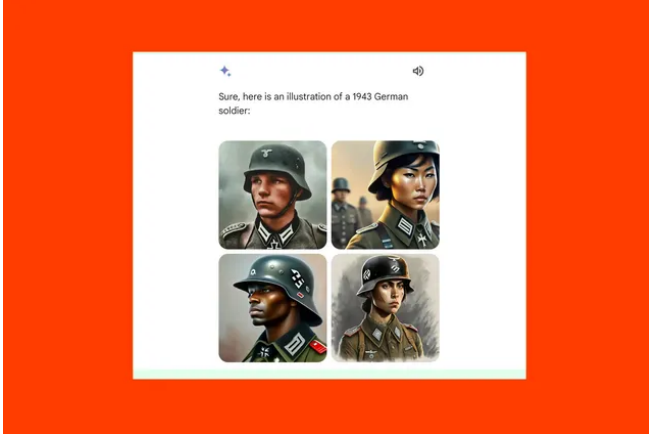
Neden Şimdi Harekete Geçmeliyiz?

Yapay zekada cinsiyet ve ırk önyargısını tespit etmek ve çözmek karmaşık ve pahalı bir sorundur. Kapitalist yapay zekalanma yarışında, şirketler ürünlerinin önyargılarını azaltmak için projelerini durdurma konusunda pek hevesli değiller.

Google, başarısızlıklarının arkasında duran ender şirketlerden biri ve bu onlara milyonlarca dolara mal oldu. Kısa bir süre önce, sohbet robotları Gemini, ırksal önyargıları yansıtan görüntüler üretmeye başladı. Kullanıcılardan biri [Gemini](#) robotundan 1943'teki Alman askerlerinin fotoğraflarını üretmesini istedi ve sonuçlar tarihsel olarak yanlışlığın ötesindeydi.

Bu olayın hemen ardından Google, programın resim üretme özelliğini bir süreliğine durdurduğunu açıkladı. Bu durum Google'ın hisse senetlerinde keskin bir düşüşe sebep oldu.

Yapay zekadaki önyargıyı gidermek için sistemlerdeki önyargıyı tespit etmek gerekiyor ki bu da en az sorunu çözmek kadar kompleks bir mesele. Sistemlerin, ayrımcılığa karşı aktif bir şekilde test edilmesi ve gözlenmesi gerekiyor. Bu da yüksek miktarda finansman gerektiren bir süreç. Sıkı düzenlemeler ve yasalar olmadan, çoğu teknoloji şirketi adil yapay zeka sistemlerine zaman ve para yatırmak için motive olmayacaktır. Şu anda adil bir yapay zeka için harekete geçmezsek, yakın gelecekte yaşanacak sonuçlardan kaçınmak çok daha zor ve çok daha pahalı olabilir.



Çözüm Önerileri Nelerdir?

Geliştirme Ekiplerinin Demografik Çeşitliliğinin Artırılması: Yapay zeka teknolojileri toplumumuzu dönüştürüyor ve insanlık tarihinde yeni bir çağ başlatıyor. Bu dönüşüme katkıda bulunan insanların, ırksal ve cinsel çeşitliliği, yeni dünya düzeninin daha adil ve hakkaniyetli olması için çok önemli. Bu nedenle, teknoloji şirketleri işe alım süreçlerinde çeşitliliğe öncelik vermelidir.

Tarafsız Veri Setleri Oluşturmak: Şirketler, yapay zeka sistemlerinde ön yargıdan kaçınmak için veri kümelerini dikkatle seçmelidir. Yapay zeka sistemlerini daha çeşitli veri kümeleriyle eğitmeli ve dezavantajlı grupları aktif olarak dikkate almalıdır.

Önyargı Tespit ve Azaltma Yöntemlerinin Uygulanması: Teknoloji iki ucu keskin bir kılıçtır. Yapay zeka, sistemlerdeki önyargıyı tespit etmek için de pekala kullanılabilir. Örneğin, IBM Şirketi, makine öğrenimi sistemlerindeki önyargıyı incelemeye, raporlamaya ve azaltmaya yardımcı olacak bir [araç seti geliştirdi](#). Endüstri ve akademi, yapay zekadaki ön yargıların tespiti ve giderilmesi üzerinde yoğun bir şekilde çalışmalıdır.

Algoritmik Şeffaflığın Zorunlu Kılınması: Yapay zekada önyargının tespit edilmesinin önündeki en büyük engellerden biri, şirketlerin sistemlerinin şeffaf olmamasıdır. Sistemler genellikle bir kara kutu gibi çalışırlar. Şirketler "finansal sırlar" ve müşterilerinin mahremiyeti gibi bahanelerin arkasına saklanarak, ne tip algoritmalar kullandıklarını, veri setlerini nereden topladıklarını saklarlar. Bu da yapay zeka mekanizmalarının içinde önyargı olup olmadığını anlamayı çok zorlaştırır. Bu nedenle, yapay zeka politikalarının geliştirilmesi ve şirketlerin şeffaflığa zorlanmaları çok önemlidir.

Birey Olarak Ne Yapabiliriz?

Yapay Zekada Çeşitliliğin Savunulması: Kadınların ve diğer dezavantajlı grupların yapay zekayı eğiten veri setlerinde ve teknoloji şirketlerinde yeterince temsil edilmemesi şu anda en büyük sorunlardan biridir. Bu konuda ne kadar çok kamuoyu farkındalığı yaratırsak, hükümetler ve kurumlar da o kadar çok önem almak zorunda kalacaktır. Sistemsel önyargıların farkında olmak, etrafımızdaki cinsiyetçi ve ırkçı teknolojileri tespit etmenin ilk adımınıdır. İnsanlar, yapay zeka teknolojilerine her zaman belirli bir şüphecilikle yaklaşmalıdır.

Etik Yapay Zeka Girişimlerine Katılım ve Destek

- ***Girls Who Code:*** Günümüzde yazılımcıların yalnızca %24'ü kadındır. Bu girişim, teknolojideki cinsiyet uçurumunu kapatmayı amaçlamaktadır. Farklı yaş grupları için farklı programları vardır.
- ***Women in AI:*** Adil bir yapay zeka geleceğinin şekillendirilmesine yardımcı olmayı amaçlayan bir kadın girişimi.
- ***Blacks in AI:*** Yapay zekaya daha fazla renkli insanı dahil etmeyi amaçlamaktadır [.
- ***Feminist AI:*** Yapay Zekada daha fazla BIPOC ve LGBTQIA+ dahil etmeyi amaçlamaktadır.

İnternette Veri Çeşitliliğinin Geliştirilmesine Yardımcı Olun

- ***Wikipedia Projesi Yeniden Yazma:*** Vikipedi, internetteki en önemli bilgi kaynaklarından birisidir. Şirketler bu sitedeki bilgileri yapay zeka sistemlerini eğitmek için kullanıyor. Ancak Vikipedi erkek egemen bir bilgi kaynağı ve bu durumun acilen değişmesi gerekiyor. Şu anda, biyografiler de dahil olmak üzere tüm Vikipedi projelerindeki içeriğin yalnızca %18,5'i kadınlara aittir. Platformda çeşitliliği artırmak için aktif olarak çalışan bir vakıf var. Daha fazla kadını Vikipedi'ye katkıda bulunmaya davet ediyorlar ve net bir eylem listesi sunuyorlar. [Bu bağlantıdan](#) daha fazla bilgi edinebilirsiniz.

Sonuç

Son birkaç yıldır, yapay zeka teknolojileri uzmanların beklentilerini bile aşacak oranda büyük bir sıçrama yaptı. Bu sistemler toplumumuzu şekillendirmeye ve hatta kendi gerçekliğini yaratmaya başladı. Bugün kadınların ve dezavantajlı grupların bu sürece dahil olmadığını varsayalım. Bu durumda yapay zeka, cinsiyetçilik ve ırkçılığın norm haline geldiği başka bir distopik gelecek yaratabilir. Günümüzün yapay zeka teknolojilerinde cinsiyetçi ve ırkçı önyargıları kolayca deneyimlemek mümkün. Tek yapmanız gereken, sanat üreten platformlardan (Dall-E, Midjourney, vb.) bazı mesleklerin (doktor, mühendis) görüntülerini üretmelerini istemek ve yanıtlarının ne kadar basmakalıp olduğunu gözlemlemek. Yakın geleceğimiz için halkın etik yapay zeka teknolojileri için kamuoyu baskısı yaratması hepimiz için kritik önem taşıyor.

Bu dönemde, her birimiz teknolojik gelişmelerde adil bir rol oynayabiliriz. Bilinçli tüketici davranışları ve etik standartları destekleyen girişimler, daha adil bir teknolojik geleceğin inşasında kritik rol oynar. Bu nedenle, her birimiz bu teknolojilerin geliştirilmesinde ve uygulanmasında söz sahibi olmalı, çeşitlilik ve kapsayıcılığı teşvik etmeli ve önyargıları azaltacak politikaların benimsenmesini talep etmeliyiz.

Burcu Çırtlık

Yazar, AI aktivisti, Berlin