

Yapay Zekâ Etiği: Toplum Üzerine Etkisi

Tülay TURAN^{1*}, Gökhan TURAN¹, Ecir Uğur KÜÇÜKSİLLE²

¹Burdur Mehmet Akif Ersoy Üniversitesi, Teknik Bilimler Meslek Yüksekokulu, Burdur
²Süleyman Demirel Üniversitesi, Mühendislik Fakültesi, Isparta

Geliş Tarihi (Received): 16.01.2022, Kabul Tarihi (Accepted): 06.06.2022

✉ Sorumlu Yazar (Corresponding author*): tulayturan@mehmetakif.edu.tr

☎ +90 248 2134552 📠 +90 248 2134598

ÖZ

Yapay zekâ sistemlerini, günümüzde toplumun temel rollerine hizmet eden, birçok uygulama alanında bizlere fayda sağlayan ve gelecek yıllarda belki biz insanlara ihtiyaç duymadan, özerk kararlar alabilen sistemler olarak tanımlayabiliriz. Yapay zekâli sistemlerin giderek daha fazla özerklikle, yani daha az insan denetimi ile çalışabilmesi için öncelikle etik yapılarının oluşturulması gerekmektedir. Etik yapay zekâ, bireysel haklar, mahremiyet, eşitlik ve ön yargı yapmama gibi temel değerlerle ilgili iyi tanımlanmış etik yönergelerle bağlı kalan yapay zekâdir. Yapay zekâ etik uygulamaları, kuruluşların daha verimli çalışmasına, daha temiz ürünler üretmesine, zararlı çevresel etkileri azaltmasına, kamu güvenliğini artırmaya ve insan sağlığını iyileştirmesine yardımcı olacaktır. Etik olmayan yapay zekâ uygulamaları ise toplum için ciddi zararlı etkilere neden olabilecektir. Bu olumsuz etkileri sorumlu bir şekilde yönetmek ve yapay zekâ sistemlerini toplum yararına yönlendirmek için en önemli çözüm, etik yapay zekâ sistemlerinin geliştirilmesidir. Son yıllarda akademi, endüstri, hükümet ve sivil toplum tarafından yapay zekâ etiği ile ilgili yapılan çalışmalar bir temel sağlamaya başlamıştır. Bu çalışmada, yapay zekâ etiği ve toplum üzerine etkisi, iş gücü piyasası, eşitsizlik, gizlilik ve ön yargı konu başlıkları üzerinde tartışılıp, olası risklere ve tehditlere dikkat çekilmekte ve çözüm için önerilerde bulunmaktadır.

Anahtar Kelimeler: Eşitsizlik, gizlilik, iş piyasası, ön yargı, yapay zekâ etiği

Ethics of Artificial Intelligence: Impact on Society

ABSTRACT

We can define artificial intelligence systems as systems that serve the basic roles of society today, benefit us in many application areas, and can make autonomous decisions in the coming years, perhaps without the need for humans. In order for artificial intelligence systems to work with more and more autonomy, that is, with less human control, ethical structures must first be established. Ethical AI is AI that adheres to well-defined ethical guidelines regarding core values such as individual rights, privacy, equality and non-prejudice. Artificial intelligence ethical practices will help organizations operate more efficiently, produce cleaner products, reduce harmful environmental impacts, increase public safety and improve human health. Unethical artificial intelligence applications may cause serious harmful effects for society. The most important solution to responsibly manage these negative effects and direct artificial intelligence systems for the benefit of society is the development of ethical artificial intelligence systems. In recent years, studies on the ethics of artificial intelligence by academia, industry, government and civil society have begun to provide a basis. In this study, the ethics of artificial intelligence and its impact on society, labor market, inequality, privacy and prejudice are discussed, possible risks and threats are pointed out, and suggestions are made for solutions.

Keywords: Inequality, privacy, the labor market, bias, ethics of artificial intelligence

GİRİŞ

Alan Turing'in 1950 yılında yayınlanan makalesinde, makinelerin düşünüp düşünemeyeceği ve insanları iyi derecede taklit edebilecek makineler üretmenin mümkün olup olamayacağı sorularını yöneltmesi, yapay zekâ kavramının başlangıcı olarak görülmektedir (Turing, 2009). Yapay Zekâ; öğrenme, akıl yürütme ve eyleme dökme işlemlerini, insan sinir sisteminden ilham alarak gerçekleştiren ve bir dizi hesaplama teknolojisine dayanan bir bilimdir. Günümüzde; sürücüsüz arabalar, hastalık teşhis yazılımları, yaşlı bakımı için fiziksel yardımcı robotlar gibi birçok örnekleri yer almaktadır. Gelecekte ise, toplumu etkileyen daha fazla çalışma alanı ile yaşamımızda yer alacaktır. Gün geçtikçe daha özerk ve daha akıllı hale gelen bu sistemler hem olumlu hem de olumsuz sonuçları beraberinde getirmektedir. Bu bağlamda, yapay zekânın toplumda uygulanmasından kaynaklanan istenmeyen, olumsuz sonuç ve risklerinden kaçınmak için, bu sistemlerinin etik bir yapı çerçevesinde, insan haklarına saygı gösteren güvenli sistemler olarak geliştirilmesine dikkat edilmelidir (Stone, 2016).

Yapay zekâ etiği, kötü (etik olmayan) tasarım, uygun-suz uygulama veya yanlış kullanımdan kaynaklanan, toplumda yapay zekâ sistemleri nedeniyle oluşabilecek etik zararları en aza indirmek için, tasarımcı ve üretici firmaların nasıl davranması gerektiği belirleyen yapıdır. Yapay zekâ sistemleri için ilk etik kod, ünlü bilim kurgu yazarı Isaac Asimov tarafından 1942 yılında 'Üç Robotik Yasası' adında sunulmuştur (Aksoy, 2021). Günümüzde ise etik kuralları belirlemek için ortaya çıkan en son çalışmalardan biri, Montreal Üniversitesi ile Quebec Araştırma Fonu'nun birlikte hazırladığı ve on ilkedен oluşan beyanname'dir. Beyanname'nin temel amacı, insanların ve grupların temel çıkarlarını destekleyen, dijital teknoloji ve yapay zekâ alanları için geçerli olan etik ilkeleri ve değerleri belirlemektir. Beyanname'de yer alan ilkeler sırasıyla; refah, özerkliğe saygı, mahremiyet, dayanışma, demokratik katılım, eşitlik, çeşitliliğin dahil edilmesi, ihtiyaç, sorumluluk ve sürdürülebilir gelişme başlıklarıdır (URL-1, 2019).

Yapay zekâdan önce otomasyon sistemleri, yalnızca insan fiziksel gücüne yardımcı olan işlerde kullanılırken, yapay zekâ ile insan düşünce ve bilgi işlemini gerektiren işlemlerde de kullanılmaya başlanmıştır (Bostrom ve Yudkowsky, 2014). Dolayısıyla bu durum, işgücü piyasasında köklü değişikliklerin yaşanmasına sebep olacağı görüşünü ortaya çıkarmıştır. Yapay zekâ sistemlerinin gelişmesiyle şirketlerin işlerinin

daha verimli hale geleceği, daha çok kazanacakları öngörülmekteyken, bu şirketlerde çalışan insanlara duyulan ihtiyacın zamanla azalacağı düşünülmektedir. Bu durum kazanılan gelirin daha az kişiye bölüneceği ve beraberinde sosyal eşitsizliklerin artacağı anlamına gelmektedir.

Yaşanan teknolojik değişimlerin yeni zorlukları yaratacağı öngörülmektedir. Ancak toplum olarak bu teknolojilerden korkulması ve endişe duyulması, yapay zekâ gelişimini olumsuz etkileyebilir. Hatta yanlış adımlar atılmasına neden olabilir. Bu yüzden henüz işin başındayken yapay zekâ tabanlı teknolojilerin etik yapısının oluşturulması ve küresel boyutta bu çerçeveler içinde sistemlerin geliştirilmesi önemlidir.

Çalışmanın ikinci bölümünde yapay zekâ etiği, üçüncü bölümünde dört alt başlık ile yapay zekâ etiği ve toplum, dördüncü ve son bölümde ise sonuç ve öneriler bölümleri yer almaktadır.

YAPAY ZEKÂ ETİĞİ

Etik, bir bireyin veya bir grubun davranışlarını veya eylemlerini yöneten, neyin iyi neyin doğru olduğunu belirlemeye yardımcı olan ahlaki ilkeler, kurallardır. Bu tanıma göre yapay zekâ etiği de yapay zekâ sistemlerinin davranışlarının belirlenmesi için kullanılan ilke ve kurallar olarak tanımlanabilir.

Araştırma merkezleri, bilim akademileri, uluslararası kuruluşlar ve sivil toplum dernekleri, yapay zekâ gelişimi için etik bir çerçeve talep etmektedirler. Bu çerçevenin gerçekleştirilmesi için de küresel bir yaklaşım bütünlüğünün sağlanarak ve ortak bir yapının oluşturulmasını en temel unsur olarak görmektedirler (URL-2, 2020). Ancak, devletlerin etik ilkeleri arasındaki önemli farklılıklar göz önüne alındığında, ortak uluslararası yapay zekâ düzenlemelerinin üzerinde anlaşılıp anlaşılmayacağı belirsizliği ortaya çıkmaktadır.

Roosen (2020) çalışmasında, hayatta doğru dengeyi bulmanın ve etik sonuçlar elde etmenin zorlu bir süreç olduğunu vurgulamıştır. Ayrıca, gerçek dünya senaryolarını, yapay zekâ ile tasarlamak, geliştirme ve dağıtım süreçlerinin her aşamasında, etik çerçevenin dikkate alınmasının acil bir ihtiyaç olduğunu altını çizmiştir. Diğer bir görüşte ise yapay zekâ geliştiricilerinin, geliştirdikleri sistemlerin kullanımı ve erişimini belirlemede büyük bir yetkiye sahip olduklarını ve çalışmalarını toplum üzerine etkilerini dikkate alarak tasarlamaları gerekliliğini vurgulamıştır (URL-3, 2018).

Yapay Zekâ Etiği: Toplum Üzerine Etkisi

Tablo 1. Yapay zekâ etiği ile ilgili güncel çalışmalar

	Çalışma Adı	Elde Edilen Sonuç
1	Management perspective of ethics in artificial intelligence (Baker, 2021)	Çalışmalarında yapay zekâ etiği konusunun karmaşık ve cevaplanmasının zor bir konu olduğunu belirtmişlerdir.
2	Moral consideration of nonhumans in the ethics of artificial intelligence (Owe ve Baum, 2021)	Yapay zekâ etiğinin, yapay zekânın kendisine nasıl değer verileceği ve tüm bunların yapay zekâ sistem tasarımına nasıl dahil edileceği konusunda özgün çalışmalar yapılması gerekliliğini belirtmişlerdir.
3	From posthumanism to ethics of artificial intelligence (Nath ve Manna, 2021)	Yapay zekâ veya herhangi bir akıllı teknolojinin en iyi karara varmak için kalıpları analiz etmemize yardımcı olabileceği, ancak kesinlikle bu makinelerin etik problemlere yol açabileceği belirtilmiştir.
4	Artificial intelligence for good health: a scoping review of the ethics literatüre (Murphy ve ark., 2021)	Sağlık alanında yapay zekâ yı çevreleyen etik konuların hem geniş hem de karmaşık olduğunu belirtmişlerdir. Ayrıca yapay zekâ nın sağlık sistemlerini iyileştirme potansiyeline sahip olsa da etik yapısı dikkate alındığında temkinli bir iyimserlikle yaklaşılması gerektiğini öne sürülmektedir.
5	Ethical Management of Artificial Intelligence (Brendel ve ark., 2021)	Sonuç olarak, farklı disiplinlerden araştırmacıların, yapay zekâ hakkında etik bir söylemi geniş çapta kabul ettiklerinin, bununla birlikte, etik düşünceleri tetiklemeye istekli olan yöneticilerin, yapay zekâ etiğini nasıl oluşturabilecekleri ve yönetebilecekleri konusunda sınırlı destek aldıkları belirtilmiştir.
6	Ethics and privacy of artificial intelligence: Understandings from bibliometrics (Zhang ve ark., 2021)	Yapay zekâyı çevreleyen etik konular üzerine araştırmalara katkıda bulunan temel araştırmaları ve araştırma topluluklarını çalışmalarında listelemişlerdir.
7	Ethics and Governance of Artificial Intelligence: Evidence from a Survey of Machine Learning Researchers (Zhang ve ark., 2021)	Yapay zekâ etiği araştırmalarına öncelik verilmesi gerektiğini ve makine öğrenimi kurumlarının potansiyel zararlarının değerlendirilmesi için yayın öncesi incelemesi gerekliliği belirtilmiştir.
8	Ethical framework for Artificial Intelligence and Digital Technologies (Ashok ve ark., 2022)	Yapay zekânın ön yargıları, klişeleri ve açıklanamayan otonom kararları yaymasında karşılaşılan aksiliklerin giderilmesi için, mevcut ve gelecekteki teknolojiler için yapay zekâ etiğinin önemi vurgulanmıştır.
9	Organisational responses to the ethical issues of artificial intelligence (Stahl ve ark., 2022)	Yapay zekâyı kullanan kuruluşların bu teknolojilerin neden olabileceği etik sorunların farkında olduklarına ve aynı zamanda bu sorunlar ile ilgili bir şeyler yapmak için sorumluluk almaya istekli olduklarını belirtmişlerdir.
10	Artificial intelligence ethics has a black box problem (Bélisle Pipon ve ark., 2022)	Yapay zekâ etik ilkeleri oluşturulurken akademi, endüstri, hükümet ve sivil toplum paydaş katılımının sağlanması gerekliliği vurgulanmıştır.

Yapay zekâ etiği ile ilgili güncel yayınları incelediğimizde elde edilen sonuçlar Tablo 1'de gösterilmektedir.

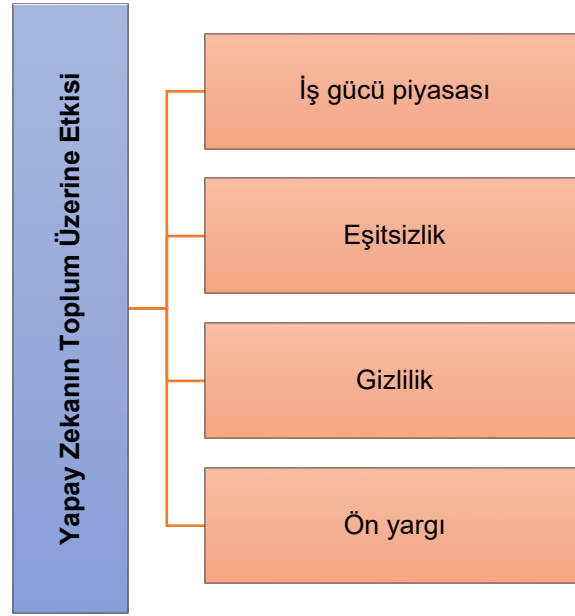
Bugün yapay zekânın etiği, doğru cevaplardan çok doğru sorularla ilgilidir. Yapay zekânın insan zekâsına eşit olup olmayacağını bilinmemektedir. Ancak hızla ve öngörülemez bir şekilde geliştiği için, bu geçişi kolaylaştıracak ve olumsuz sonuç riskini azaltabilecek önlemler hakkında düşünmek son derece önemlidir.

Yapay zekâ sistemleri neyin iyi veya doğru olduğunu nasıl belirler? Algoritmaların mahremiyet, veri gizliliği gibi temel insan haklarını ihlal etmediğinden nasıl emin olabiliriz? Kararlar ve eylemler tamamen otomatik olduğunda hesap verebilirliği nasıl sağlayabiliriz? Verilen kararlar evrensel olarak mı iyi yoksa sadece bazıları için mi iyidir? Bazı ölçütlere karşı iyi ama diğerleri için o kadar iyi değil midir? Bu ve benzeri tartışmalar,

sorular yapay zekâ etiği üzerine birçok çalışmanın yapılmasına zemin hazırlayarak, gelişmesine olumlu katkı sağlayacaktır.

YAPAY ZEKÂ ETİĞİ VE TOPLUM

Bu bölümde yapay zekâ etiği ve toplum üzerine etkileri Şekil 1'de gösterilen alt başlıklar ile incelenmiş olup, yapay zekâ sistemlerinin gelişmesi ile karşı karşıya gelebilecek zorluklar ve bu zorlukların çözümleri literatür taraması yapılarak açıklanmıştır.



Şekil 2. Yapay zekâ etiği ve toplum üzerine etkileri

İş Gücü Piyasası

Teknolojik gelişmeler ve ilerlemeler, insanlar için iş potansiyellerinin ellerinden alınacağı kaygısını da beraberinde getirmektedir. Günümüzde yapay zekâ ve makine öğrenimindeki gelişmeler için de aynı kaygıların toplum tarafından duyulduğu görülmektedir. Leontief, tıpkı atların içten yanmalı motorların icadıyla modasının geçmesi gibi, insanların da yerini makinelerin alacağından endişeliydi (Perkins, 1983). Ancak geçmişte, otomasyon kısa vadede insan emeğinin yerini almış, ancak uzun vadede yeni işlerin yaratılmasına yol açmıştır (David, 2015).

Bununla birlikte, yapay zekânın önümüzdeki yirmi yıl içinde kitlesel işsizlik yaratabileceğine dair yaygın endişeler bulunmaktadır. Frey ve Osborne (2017) çalışmalarında yeni bilgi teknolojilerinin, yakın gelecekte çok çeşitli mesleklerde, önemli bir istihdam payını riske atacağı sonucuna varmışlardır. Mevcut iş sınıflarının riske atılması sonucunda ise, iş sektöründe kitlesel iş-

sizliğe neden olarak mevcut sosyal ve ekonomik bölgelerin daha da kötüleşmesi mümkündür. Bu durum ile karşı karşıya gelecek öncelikli grupların ise düşük eğitimli kişiler ve genç toplum olacağı öngörülmektedir.

İş gücü piyasası üzerinde yaşanan bu değişimler, toplum üyeleri üzerinde farklı etkiler ile hissedilecektir. İngiltere İstihdam ve Beceriler Komisyonu'na göre, daha az teknik beceriye sahip olan kişiler, gelecekte daha çok zorluklar ile karşılaşacaklardır (United Kingdom Commission, 2014). İş gücü piyasasına yeni giren gençlerin ise, yapay zekâ ile çalışan ilk nesil olacakları için, önceden yaşanmamış problemlere çözümler bulmak zorunda kalacakları belirtilmiştir (Biavaschi ve ark., 2012).

Tesla grubun yöneticisi Elon Musk geliştirdikleri sürücüsüz kamyonların önümüzdeki on yıl içerisinde yaygın olarak kullanılacağı vaadini vermiştir (URL-4, 2018). Peki bu vaat, dünya üzerinde kamyon taşımacılığı sektöründe çalışan milyonlarca kişi üzerine nasıl etki edecektir? Öte yandan, sürücüsüz kamyonların

Yapay Zekâ Etiği: Toplum Üzerine Etkisi

kaza riskinin daha düşük olduğu biliniyorsa, etik olarak sürücüsüz kamyonların seçilmesinin daha iyi bir karar olup olmayacağı ikilem beraberinde gelmektedir.

Yaşlı ve hasta bakım sektöründe, artan bir şekilde yapay zekâli yardımcı sistemlerin kullanılması, bu sektörde ağırlıklı olarak çalışan kadın çalışanları orantısız etkileyecektir. Ayrımcılık, ön yargı ve eğitim eksikliği ile karşı karşıya kalmalarına ek olarak bu durumun birçoğunun işsiz kalmasına neden olabileceği görülmektedir (Robinson ve ark., 2015).

Furman ve Seamans (2019)'e göre yapay zekânın teknolojilerindeki gelişmeler ekonomik büyümeyi yanında getirerek, farklı iş sektörlerinin durumunu olumlu yönde değiştirebilir. Örneğin, ekonomik büyüme sonucuna bağlı olarak, ailelerin gelirinin artabileceği ve ailelerin dinlemek için daha çok tatil, seyahat planlamasında bulunabileceği ve restoranlarda daha çok yemek yiyebileceği öngörülmektedir.

Son olarak Mokyry ve ark. (2015) çalışmalarında, uzun vadede teknolojik işsizliğin gerçekleşme olasılığının çok düşük görüldüğünü, bununla beraber geçiş yolunda bazı işçiler ve endüstriler için istenmeyen sonuçları beraberinde getirebileceğini ortaya koymuşlardır. Diğer bir deyişle, yapay zekâ uzun vadeli işsizlik sorunları yaratmasa da, kısa vadede otomatikleştirilebilir alanlarda işçilerin gelir elde etmelerini engelleyecek şekilde etkiler ortaya çıkarabilecektir.

Eşitsizlik

Eşitsizlik, genellikle ekonomik gelişim ve büyümenin yaşandığı durumlarda daha da kötüleşen yapısı ile ortaya çıkmaktadır. Angus Deaton (2013)'e göre, "*İnsanlar üzerinde iyi olarak tanımlanan bir ilerleme, geride eşitsizlik mirası bırakmaktadır*". Sermaye kazançlarında yaşanan büyük artışlar, artan eşitsizlikle ilişkilendirilebilir. Yapay zekâ teknolojisinin, sermaye geliri için olumlu bir etki rolüne sahip olduğu düşünüldüğünde, eşitsizliği de yanında getireceği öngörülmektedir. Örneğin yapay zekâ, uzun vadeli işsizlik sorunları yaratmasa da kısa vadede otomatikleştirilebilir alanlarda işçilerin gelir elde etmelerini engelleyebilir. Bu dönem, nüfusun bir kısmının gelir elde etmesini engelleyerek, eşitsizliğe katkıda bulunacaktır.

Bryson (2019)'a göre, yapay zekânın yükselişi servet eşitsizliğine ve siyasi kargaşaya yol açabilecektir. Eşitsizlik, siyasi kutuplaşma ile oldukça ilişkilidir (McCarty ve ark., 2016). Kutuplaşmanın olası bir sonucu olarak, inançların grup içi statüye veya bağlılığa işaret etmek

için kullanıldığı kimlik politikasında artış meydana gelecektir (Iyengar ve ark., 2012).

Beyaz Saray'ın yapay zekâ, otomasyon ve ekonomi üzerine yayınladığı rapor incelendiğinde, yapay zekânın potansiyel faydalarının eşit olmayan bir şekilde birikmemesinin ve mümkün olduğunca çok insan için erişilebilir hale getirilmesinin önemi vurgulanmaktadır (URL-5, 2016).

Sonuç olarak Yapay zekâ teknolojisi, sahiplerini toplumun geri kalanından çok daha zengin hale getirmek için üretkenliği yoluyla yeterli gelir elde etme potansiyeline sahiptir. Dahası, bu üretkenlik belirli emek biçimlerinin yanı sıra çalışma koşullarının da yerini alarak, birçok çalışana tehdit etmektedir. Bu gibi gelişmeler yapay zekâ teknolojisinin eşitsizlikte bir artışa neden olacağı görüşünü desteklemektedir. Geleceğin politikacılarının, ülke genelinde yaşam standardında eşitliği artırmak için, yapay zekânın üretkenliğini kullanan bir gelecek yaratmaya odaklanmaları gerekmektedir.

Gizlilik

Yapay zekâ sistemlerinin gelişimi ile kişisel bilgilerin kullanılma oranı da gün geçtikçe artmaktadır. Bu durum bilgilerin gizlilik çıkarlarına müdahale edebilecek şekillerde kullanılmasına neden olabilmektedir. Teknoloji şu anda, saklanabilir verilerin üretilebildiği ve uzun vadeli kayıtların tutulabileceği aşamadır. Dijital kayıtlar, örüntü tanıma algoritmaları kullanılarak aranabilir. Bu durum, belirsizlik nedeniyle varsayılan anonimlik varsayımının kaybedildiği anlamına gelmektedir (Selinger ve Hartzog, 2017).

Makine öğrenimi, verilerden bilgi çıkarmaya ve yeni kalıplar keşfetmeye olanak tanırken, görünüşte zararsız verileri, kişisel özel verilere dönüştürebilir. Bunun sonucunda, bu bilgiler herhangi bir kişinin yüz tanıma sistemi ile tanınmasından, alışveriş alışkanlıklarının belirlenmesine kadar geniş bir yelpazede kullanılabilir (Pasquale, 2015). Hatta kişinin politik veya ekonomik eğilimlerini ortaya çıkararak, bunları değiştirmek için hangi stratejilerin etkili olabileceğini bile gösterebilir (URL-6, 2017).

Azınlıkların orantısız bir şekilde etkileyecek kararların alınması (Spielkamp, 2017) veya Amazon şirketinin işe alım algoritmasında yaşanan erkeklere öncelik verilmesi (URL-7, 2018) örneklerinde görüldüğü gibi, gizlilik konu başlığı, çoğu zaman yapay zekâ sistemlerinin sınırlamalarını ve başarısızlıklarını ortaya çıkarır. Ayrımcılık ten rengi, cinsel kimlik ve ulusal köken gibi kişisel özelliklere dayanmaktadır. Bu niteliklerin açık

Yapay Zekâ Etiği: Toplum Üzerine Etkisi

bir şekilde ilgili kişinin çıkarlarına aykırı olan otomatik karar verme sistemlerinde kullanılması, beraberinde kişisel mahremiyetin ve gizliliğin ihlal edilmesine yol açacaktır.

Yaşlı ve hasta bakım hizmetlerinde kullanılacak yapay zekâ sistemleri tasarlanırken, kullanıcıların mahremiyetine ve kişisel bilgi gizliliğine ayrıca özen gösterilmelidir. Çünkü insanların evlerinde olacak bu sistemler, kişilerin giyinmesine, banyo yapmasına yardım edecek, özel hayatlarına müdahil olacaklardır.

Sonuç olarak, yapay zekâ sistemlerinde kişisel bilgilerin kullanılıp kullanılmayacağı veya nasıl kullanılacağına belirlenmesi için etik bir çerçevenin oluşturulması gerekliliği ortaya çıkmaktadır. Etik çerçevenin, hem kişisel bilgileri kullanımından kaynaklı olumsuz bir etkiye bireyleri koruması hem de yapay zekâ gelişimini kısıtlamayan bir yapıda olması önemlidir.

Ön yargı

İnsan ön yargısı, psikoloji biliminde geçmişten günümüze üzerine birçok çalışma yapılan bir konu başlığıdır. Son yıllarda hızla gelişen yapay zekâ sistemlerinde de insan ön yargısının ne kadar etkili olduğu ve bu sistemlerin ön yargı olmadan nasıl tasarlanabilecekleri konusunda birçok çalışma yer almaktadır.

Yapay zekâ sistemlerinde ön yargılı sonuçların oluşmasındaki temel neden, eğitim verileridir (Hurtado ve ark., 2021). Verilerde cinsiyet, ırk, din ve yaş bilgileri yer almasa bile, insanlara ait çarpık görüşlerin veya tarihsel, sosyal eşitsizlikleri temsil edebilen verilerin yer alması, ön yargıya dayalı kararların alınmasına ve öğrenilmesine neden olmaktadır. Microsoft 2016 yılında, insanlar ile etkileşime girerek öğrenmesini planladığı Tay isimli sohbet robotu adına Twitter'da bir hesap açmıştır. Ancak bazı grupların sohbet robotuna yoğun bir şekilde ırkçı, cinsiyetçi ve dini mesajlar atarak, ön yargılı bilgiler öğrenmesine neden olmaları yüzünden, 16 saat gibi kısa sürede hesap kapatılmıştır (Hovy ve ark., 2021).

ProPublica kuruluşunun yaptığı habere göre, Amerika Birleşik Devletinde bir sanığın yeniden suç işleme olasılığını değerlendirmek için COMPAS adında bir yapay zekâ yazılımı kullanılmıştır. Yazılımın, aynı suçu işleyen iki farklı ten rengindeki bireyler için; beyaz tenli kişilerin yeniden suç işlemeceğine, siyah tenli kişilerin ise yeniden suç işleyeceğine dair öngöründe bulunduğu görülmüştür. Bu yazılımın siyah tenli Amerikalılara karşı güçlü bir ön yargı gösterdiği ortadadır (URL-8, 2016).

Datta ve ark. (2015) çalışmalarında, otomatik reklam dağıtım araçlarının erkeklere yüksek ücretli işler için reklam dağıtım olasılığının, kadınlara göre daha yüksek olduğunu bulmuşlardır. İş başvurularını değerlendirmek için kullanılan yapay zekâ yazılımlarının da ön yargılı kararlar verebildiği görülmüştür. Buna örnek olarak, Amazon yazılımı gösterilmektedir. Kendi kendine öğrenebilen Amazon yazılımının iş alım değerlendirmelerinde erkek bireylerin başvurularını en üst sıralara alarak, önemli ölçüde erkek çalışanları tercih ettiği görülmüştür.

Yüz tanıma sistemlerinin de ön yargı içermeyen kararlar vermesi için, veri setleri içerisinde ırkların, cinsiyetlerin, ten renklerinin eşit bir şekilde, bol örnek ile temsil edilmesi gerekmektedir. Yüz tanıma sistemlerinde kullanılan "Faces in the Wild" veri kümesinde %70 erkek ve %80 beyaz renkli insanlara ait verilerin bulunduğu görülmektedir. Bu veri setinin kullanılmasının kadın ve siyah tenli insanlar için ön yargılı sonuçları beraberinde getireceği ortadadır. New York City'deki Bilgisayar Makineleri Derneği (ACM), "Etnik, ırk, cinsiyet ve diğer insan özelliklerine dayalı açık ön yargı" nedeniyle yüz tanıma teknolojilerinin özel ve devlet kurumları tarafından kullanımının durdurulması çağrısında bulunmuştur (URL-9, 2020).

Günümüz yapay zekâ sistemlerinde tüm ön yargıları ortadan kaldırmanın hızlı bir çözümü olmamakla birlikte, en aza indirmek için kullanılan bazı araçlar bulunmaktadır. IBM Github üzerinde AI Fairness 360 adını verdiği bir kitaplık yayınlamıştır. Bu kitaplığın şu an için ikili sınıflandırma problemlerindeki ve denetimsiz öğrenme algoritmalarındaki ön yargıları tespit ettiği belirtilmektedir. IBM tarafından paylaşılan Watson OpenScale yapay zekâ sistemlerinde gerçek zamanlı ön yargı denetimini gerçekleştirerek, ön yargının azaltıldığı vurgulanmıştır (URL-10, 2020).

SONUÇ

İçinde yaşadığımız toplum, yapay zekâ sayesinde daha önce insanlık tarihinde gözlemlenen gelişmelerden hem daha büyük hem de daha hızlı bir ölçekte teknolojik bir devrim yaşamaktadır. Bu gelişimin toplum üzerindeki etkilerine yön verebilmek için, küresel bir boyutta yapay zekâ etik yapısının oluşturulması ve genel çerçevelerin belirlenmesi gerekmektedir.

Bu çalışmada, yapay zekânın toplumu etkileyen dört ana başlığı incelenmiş olup; iş gücü değişimi, eşitsizlik, gizlilik ve ön yargı konularında aşağıda yer alan öneri ve çözümler sunulmuştur. Yapay zekâ ile ortaya çıkabilecek işgücü değişimi, teknolojik işsizlik gibi durumlar

Yapay Zekâ Etiği: Toplum Üzerine Etkisi

ile başa çıkabilmek için devletlerin ekonomik politikalar izlemesi gerekebilir. Yerinden edilmiş işçiler için yeni eğitim programları sunulabilir. Sosyal ve mali destek almaları sağlanabilir. Ayrıca hükümetler yapay zekâ teknolojilerine yatırım yaparak hem inovasyona katkı sağlayabilir hem de elde edilen gelirleri tüm çalışanların yaşam koşullarını desteklemek için kullanabilir.

Ön yargıyı yapay zekâ sistemlerinde azaltmanın ya da kaldırmanın kolay olmayacağı görülmektedir. Çünkü birçok makine öğrenimi modeli insan tarafından oluşturulan verilerden oluşturulduğundan, insan ön yargıları eğitim verilerinde kolayca çarpık bir dağılımla sonuçlanabilir. Geliştiricilere bu boyutta çok önemli bir iş düşmektedir. Ön yargıları tanımaya ve bunlara karşı koymaya çalışırlarsa, daha iyi ve temiz veri setleri kullanırlarsa yapay zekâ uygulamalarında karşılaşılan adaletsizlik ve ayrımcılığı en aza indirebilirler.

Yapay zekânın maddi yönden, düşük gelirli insanlara fayda sağlamaktansa, yüksek gelirli insanlara daha fazla fayda sağlamak için kullanıldığını düşünülebilir. Bu durumu aksine çevirip toplum içinde eşitsizliği azaltmak için düşük gelirli insanların ihtiyaçlarına yönelik tasarım ve gelişimlerin hükümetlerce desteklenmesi sağlanabilir.

Kişisel bilgiler içeren veri setlerinin, yapay zekâ sistemlerinde kullanılması, herhangi bir siber saldırıda bu bilgilerin kötüye kullanılması durumunu ortaya çıkarabilir. Bu bağlamda, kişisel bilgilerin gizliliğinin korunması için veri güvenliğinin sağlanması ön şart olmalıdır. Bununla beraber bilgileri kullanılan kişilere, bilgilerinin hangi projelerde kullanıldığı bilgisi verilmeli ve kişinin onaylamadığı işlemlerde bilgilerinin veri setinden silinmesi sağlanmalıdır.

Sonuç olarak, tüm yeni teknolojiler her zaman iyi sonuçlar üretmesi umudu ile geliştirilir. Yapay zekâ da biz insanlara yardım etmek ve dünyayı daha iyi bir yer haline getirmek için geliştirilmiştir. Beklenen iyi sonuçların elde edilebilmesi için bu sistemlerin etik kurallara uygun olarak geliştirilmesi gerekmektedir. Henüz insanlar üzerinde bile genel bir etik yapının oluşturulmadığı günümüzde yapay zekâli sistemler için bunu gerçekleştirmenin oldukça zor olacağı görülmektedir. Ayrıca, yapay zekânın gelişiminden kaynaklanan etik kaygıların çeşitliliği ve karmaşıklığı da gelişimini etkileyecektir. Ancak, yapay zekâ etiği konusunda hükümetler, geliştiriciler, akademisyenler tarafından günümüzde birçok çalışma yapılmakta olup, bu çalışmalar yapay zekâ etik gelişimi için umut vadecidir.

KAYNAKLAR

- Aksoy, H. (2021). Yapay zekâli varlıklar ve ceza hukuku. *Uluslararası Ekonomi Siyaset İnsan ve Toplum Bilimleri Dergisi*, 4(1): 10-27.
- Ashok, M., Madan, R., Joha, A., Sivarajah, U. (2022). Ethical framework for artificial intelligence and digital technologies. *International Journal of Information Management*, 62, 102433; DOI: 10.1016/j.ijinfomgt.2021.102433
- Baker Brunnbauer, J. (2021). Management perspective of ethics in artificial intelligence. *AI and Ethics*, 1(2); 173-181.
- Bélisle Papon, J. C., Monteferrante, E., Roy, M.C., Couture, V. (2022). Artificial intelligence ethics has a black box problem. *AI & SOCIETY*, 1-16.
- Biavaschi, C., Eichhorst, W., Giulietti, C., Kendzia, M. J., Muravyev, A., Schmidl, R., Zimmermann, K. F. (2013). *Youth unemployment and vocational training*. Now Publishers Incorporated, Germany.
- Bostrom, N., Yudkowsky, E. (2014). The Cambridge handbook of artificial intelligence. In: *The ethics of artificial intelligence*. Frankish, K., Ramsey, W. M. (eds.), Cambridge University Press, Cambridge, UK, 316-334.
- Brendel, A.B., Mirbabaie, M., Lembcke, T.B., Hofeditz, L. (2021). Ethical management of artificial intelligence. *Sustainability*, 13(4): 1974.
- Bryson, J.J. (2019). The past decade and future of AI's impact on society. *Towards a new enlightenment*, Turner, Madrid, Spain, 150-185.
- Datta, A., Tschantz, M. C., Datta, A. (2015). Automated experiments on ad privacy settings: A tale of opacity, choice, and discrimination. *Proceedings on Privacy Enhancing Technologies*, 1: 92-112.
- David, H.J.J.O.E.P. (2015). Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives*, 29(3): 3-30.
- Deaton, A. (2013). *The great escape: health, wealth, and the origins of inequality*. Princeton University Press, ABD.
- Frey, C.B., Osborne, M.A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, 114: 254-280.
- Furman, J., Seamans, R. (2019). AI and the economy. *Innovation policy and the economy*, 19(1): 161-191.
- Hovy, D., Yang, D. (2021). The importance of modeling social factors of language: Theory and practice. *The 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, June 6-11, 2021, Book of Proceedings, 588-602.
- Hurtado, J.V., Londoño, L., Valada, A. (2021). From Learning to relearning: A framework for diminishing bias in social robot navigation. *Frontiers in Robotics and AI*, 8: 69-88.
- Iyengar, S., Sood, G., Lelkes, Y. (2012). Affect, not ideology: social identity perspective on polarization. *Public Opinion Quarterly*, 76(3): 405-431.
- McCarty, N.M., Poole, K.T., Rosenthal, H. (2016). *Polarized America: The dance of ideology and unequal riches*. MIT Press, Cambridge, UK.
- Mokyr, J., Vickers, C., Ziebarth, N. L. (2015). The history of technological anxiety and the future of economic growth: Is this time different? *Journal of Economic Perspectives*, 29(3): 31-50.

Yapay Zekâ Etiği: Toplum Üzerine Etkisi

- Murphy, K., Di Ruggiero, E., Upshur, R., Willison, D. J., Malhotra, N., Cai, J. C., Gibson, J. (2021). Artificial intelligence for good health: a scoping review of the ethics literature. *BMC Medical Ethics*, 22(1); 1-17.
- Nath, R., Manna, R. (2021). From posthumanism to ethics of artificial intelligence. *AI & SOCIETY*, 1-12.
- Owe, A., Baum, S. D. (2021). Moral consideration of nonhumans in the ethics of artificial intelligence. *AI and Ethics*, 1(4): 517-528.
- Pasquale, F. (2015). *The Black Box Society: The secret algorithms that control money and information*. Harvard University Press, UK.
- Perkins, C. D. (1983). *The long-term impact of technology on employment and unemployment*. National Academy Press, USA.
- Robinson, L., Cotten, S. R., Ono, H., Quan-Haase, A., Mesch, G., Chen, W., Stern, M. J. (2015). Digital inequalities and why they matter. *Information, Communication & Society*, 18(5): 569-582.
- Roosen, J. (2020). AI Ethics: Why does it matter. *Institute for Ethics in Artificial Intelligence*, 1-7.
- Selinger, E., Hartzog, W. (2017). Obscurity and privacy. In: *Spaces for the future: A companion to philosophy of technology*, Pitt, J., Shew, A. (eds.), New York, USA.
- Spielkamp, M. (2017). Inspecting algorithms for bias. *Technology Review*, 120(4): 96-98.
- Stahl, B. C., Antoniou, J., Ryan, M., Macnish, K., Jiya, T. (2022). Organisational responses to the ethical issues of artificial intelligence. *AI & SOCIETY*, 37(1); 23-37.
- Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., Teller, A. (2016). Artificial intelligence and life in 2030: the one hundred year study on artificial intelligence. Report of the 2015 Study Panel, September.
- Turing, A. M. (2009). Computing machinery and intelligence. In: *Parsing the turing test*, 1: 23-65.
- URL-1 (2019). <https://www.canasean.com/the-montreal-declaration-for-the-responsible-development-of-artificial-intelligence-launched/> (Erişim Tarihi: 17.12.2021)
- URL-10 (2020). <https://www.ibm.com/tr-tr/cloud/watson-openscale> (Erişim Tarihi: 19.12.2021)
- URL-2 (2020). <https://www.un.org/en/chronicle/article/towards-ethics-artificial-intelligence> (Erişim Tarihi: 17.12.2021)
- URL-3 (2018). <https://futureoflife.org/2018/01/10/shared-benefit-principle/> (Erişim Tarihi: 17.12.2021)
- URL-4 (2018). https://www.ntv.com.tr/galeri/teknoloji/tesla-semi-ilk-defa-yol-larda,0kjbif0chEW6BtYpKzPC2w/wsbl1-DF1UqWVK_eA2xdkA (Erişim Tarihi: 18.12.2021)
- URL-5 (2016). <https://obamawhitehouse.archives.gov/sites/whitehouse.gov/files/documents/Artificial-Intelligence-Automation-Economy.pdf> (Erişim Tarihi: 18.12.2021)
- URL-6 (2017). <https://www.theguardian.com/politics/2017/feb/26/robert-mercer-breitbart-war-on-media-steve-bannon-donald-trump-nigel-farage> (Erişim Tarihi: 18.12.2021)
- URL-7 (2018). <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scrapsecret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G> (Erişim Tarihi: 19.12.2021)
- URL-8 (2016). <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> (Erişim Tarihi: 20.12.2021)
- URL-9 (2020). <https://www.acm.org/binaries/content/assets/public-policy/ustpc-facial-recognition-tech-statement.pdf> (Erişim Tarihi: 20.12.2021)
- Zhang, B., Anderljung, M., Kahn, L., Dreksler, N., Horowitz, M. C., Dafoe, A. (2021). Ethics and governance of artificial intelligence: Evidence from a survey of machine learning researchers. *Journal of Artificial Intelligence Research*, 71; 591-666.
- Zhang, Y., Wu, M., Tian, G. Y., Zhang, G., Lu, J. (2021). Ethics and privacy of artificial intelligence: Understandings from bibliometrics. *Knowledge-Based Systems*, 222; 106994; DOI: 10.1016/j.knosys.2021.106994.