

Yapay Sinir Ağı Tabanlı Dil Modelleri Genel Zekânın Temel Gerekliliklerini Sağlamakta Mıdır?

Do Artificial Neural Network-Based Language Models Fulfill the Foundational Requirements of General Intelligence?

HASAN ÇAĞATAY 

Social Sciences University of Ankara

Received: 30.11.2024 | Accepted: 30.06.2025

Abstract: Today's large-scale language models are based on artificial neural networks. Since 2020, large language models like the GPT series have captured the attention of academia, business, and the public, focusing their interest on neural network-based natural language processing technologies. While machine learning researchers hold diverse views on the GPT series' success in natural language processing and its significance regarding artificial general intelligence, there is consensus that its language processing capabilities have exceeded expectations. This study examines whether artificial neural networks can achieve general intelligence. It evaluates whether some fundamental cognitive skills necessary for general intelligence can be simulated by current artificial neural networks. To address this, the paper discusses how artificial neural networks operate through statistical processes and explores the relationship between these processes and key components of general intelligence and understanding.

Keywords: Artificial Neural Networks, Artificial General Intelligence, Natural Language Processing, Artificial Understanding, Functionalism.



Giriş: Yapay Genel Zekâ ve Yapay Güçlü Zekâ

Yapay sinir ağları (YSA), makine öğrenmesi algoritmaları arasında, Yapay Genel Zekâyâ (YGZ) ulaşma potansiyeli en çok tartışılan algoritmalarlardır. Bunda YSA'ların insan sinir sistemiyle olan yüzeysel de olsa benzerlikleri ve başarı düzeyleri, temel gerekçeler olmuştur. Bu teknolojilerin beklenenin ötesinde başarıları, bilim ve felsefe dünyasında YGZ'nin imkânı tartışmalarını alevlendirmiştir. Bu konuda iyimser bilim insanlarının (Kaku, 1999; Kurzweil, 2005) yanında, YSA'ların bu hedefe ulaşmada kritik eksikliklerden mustarip olduğunu ileri sürenler de vardır (Mitchell, 2019; Bender & Koller, 2020; Pepperell, 2022; Searle, 1980). Bugünün yapay zekâ (YZ) sistemlerinin, Yapay Genel Zekâyâ sahip olup olmadığı ya da bunu başarmak için doğru yolda olup olmadıkları problemlerini ele almadan önce, sıkça yapılan bir kavramsal hatayı düzeltmekte fayda vardır. Güçlü yapay zekâ ve genel yapay zekâ aynı anlama gelmez.

Güçlü yapay zekâ kavramı John Searle (1980) tarafından ortaya atılmıştır ve zekânın otantikliği ya da insanınki gibi olması ile ilgilidir. Yapay güçlü zekâ *tezi*, bir makinenin tıpkı insan gibi gerçek manasıyla düşünebileceğini iddia eder. Yapay güçlü zekâ *sistemleri* de insanlar gibi otantik biçimde düşünebilen teorik sistemlerdir. Böyle sistemlerin “gerçek anlamda düşünme” hakkındaki kavramsal analizimize bağlı olarak *bilinç*, *yönelimsellik*, belki *duygular* gibi bilişsel bileşenleri göstermesi beklenir. Bu daraltılmış ama köküne daha uygun bakış açısıyla bir primatın ya da köpeğin güçlü zekâyâ sahip olduğunu iddia etmek mümkündür. Zayıf yapay zekâ *tezi*, güçlü yapay zekâ *tezinin* karşıtı olarak makinelerin bizim gibi düşünsün ya da düşünmesin, düşünmeyi simüle edebileceğini öne sürer. Güçlü zekâ *sistemlerinin* karşıtı sayılabilecek zayıf yapay zekâ *sistemleri* ise gerçek anlamda düşünmenin karşıtı olarak, düşünmeyi sadece simüle etmektedir. Dikkat edilmelidir ki, güçlü yapay zekâ, zayıf yapay zekâ ikilisi zekanın niceliğiyle ilgili değildir.

Genel zekâ dar zekâ ayrımı ise, sistemlerin çok sayıda farklı türden problemi çözüm çözemediğiyle ilgilidir. Zekânın otantikliğiyle değil, genelliğiyle ilgilidir. Bu durumda, bizden daha *genel* bir zekâyı simüle eden sistemler zayıf yapay zekâ sistemleri olabilir ve bizden daha *dar* problem çözme becerisi gösteren sistemler de gerçek anlamda düşünmenin şartlarını sağlıyorlarsa, *güçlü* anlamda zeki olabilir. Öyleyse, tıpkı yapay dar zekâ gibi, Yapay Genel Zekâ da ikiye ayrılmaktadır: Yapay Güçlü Genel Zekâ ve



Yapay Zayıf Genel Zekâ. Bu çalışmanın konusu güçlü değil zayıf zekâdır. Yani yapay zekâ sistemlerinin insan gibi anlayıp anlamamasından ve insan gibi düşünüp düşünmemesinden bağımsız olarak anlamayı, düşünmeyi ve genel zekâyı simüle edip edemedikleriyle ilgilenmektedir. Bu nedenle de bundan sonra, aksi açıkça belirtilmediği sürece “Yapay Genel Zekâ” ifadesi “Yapay Zayıf Genel Zekâ” yerine kullanılacaktır.

Yapay Genel Zekâ, basitçe insanların gerçekleştirebildiği tüm bilişsel görevleri gerçekleştirebilecek teorik (henüz geliştirememiş) yapay sistemleri ifade eder. Karşıtı olan *yapay dar zekâ* ise bugün sıkça karşılaştığımız tek bir ya da birkaç problemi çözmede uzmanlaşmış sistemleri ifade etmektedir. Ne zaman genel bir yapay zekâyı ulaşacağımız sorusu son derece tartışmaya açık bir soruyken (Chalmers, 2010; Moore 1965; Good, 1966; Vinge, 2003), son yirmi yılda yaşanan bazı gelişmeler, bazı düşünürlerin ilk YGZ'nin on yıllarla ölçülebilecek bir süreçte üretilebileceğinin ihtimalini daha fazla ciddiye almalarına sebep olmuştur (Kurzweil, 2005; Kaku, 1999).

Birçok durumda bir sistem ya YGZ'dir ya da değildir gibi kesikli bir YGZ analizi yapılsa da “Yapay Genel Zekâ”daki “genel” sıfatı kategorik değil, süreklidir ve dolayısıyla ölçütü de derece ifade eden bir deşışkendir. Bu nedenle, bir yapay zekâ sisteminin genel olup olmadığını sorarken, aslında sistemin *ne kadar* genel olduğunu öğrenmeyi amaçlarız. Bu derece farkını, niteliksel bir ayrıma taşıyan unsur neredeyse tüm araştırmacıların insanın zekâsının genelliğini bir sınır değeri olarak kabul edilmiş olmasıdır: Bir sistem bir insan kadar çeşitli sorunlarla başa çıkabiliyorsa, bu sistem YGZ'nin bir örneğidir.

GPT-3 (Generative Pre-trained Transformer 3 [Üretken Önceden Eğitilmiş Transformatör 3]), GPT-4 ve ardından geliştirilen yeni büyük dil modellerinin başarıları da bu sistemlerin altında yatan transformatör yapay sinir ağı mimarisinin YGZ için önemli olabileceğine işareti etmektedir. İnsanlarınkilere benzer metinler oluşturmak için derin öğrenmeyi kullanılarak geliştirilmiştir. 1 trilyon 700 milyar parametreye sahip olan GPT-4, şimdiye kadar oluşturulmuş en büyük dil modellerinden birisidir. Chat-GPT, bir yapay zekâ sistemiyle karşılıklı olarak hem doğal diller hem formal dillerde iletişim kurulabilmesine olanak verir. Bir anlamda çok gelişmiş bir sohbet robotu olduğu söylenebilir; ancak altında yatan modeller (GPT serisi) sayesinde oldukça başarılı metin üretme, geniş bilgi dağarcığı ve



sınırlı da olsa akıl yürütmedeki başarısı sergilediği gözden kaçırılmamalıdır. GPT serisi, makaleler, hikâyeler, şiirler ve hatta bilgisayar kodları gibi çeşitli formlarda son derece ikna edici ve tutarlı metinler üretebilmesi nedeniyle geniş çapta ilgi görmüştür.

Akl yürütme kısıtlılıklarına rağmen, günümüz büyük dil modelleri “yapay orta genellikte zekâ” etiketini hak etmeye en yakın sistemler sayılabilirler; çünkü dillerin fikirleri ifade etmekteki tek aracımız olduğu düşünüldüğünde, büyük dil modellerine, biyolojiden fiziğe, satranç hamlelerinden edebiyata, duygulardan etiğe birçok alanda soru sormak ve insanlarınla karşılaştırılabilir cevaplar almak mümkündür. Büyük dil modelleri, çoklu uzmanlık alanında düşünmeyi simüle etme becerisine sahiptir ve bunlarda insanın her bilişsel becerisinin kıvılcımlarını sergilemektedir.

“Her bilişsel becerinin *kıvılcımlarını* sergilemek” iddiası hem desteklenmeye muhtaçtır hem de son derece kritik bir savdır. Çünkü bu iddia doğruysa, büyük dil modelleri, birçok bilişsel beceride insandan niceliksel olarak daha geride de olsa, insanın yaptığı tüm kritik bilişsel becerileri sergileyebiliyorsa, mevcut sistemler geliştirildikçe düşünmeyi ya da anlamayı simüle edebileceğimiz anlamına gelir. John Searle (1980) ve Robert Pepperell (2022) gibi düşünürler bu iddiaya katılmazlar. John Searle, makinelerin yönelimselliğin *kıvılcımını* sergilemediğini ve bu nedenle düşünemeyeceğini iddia eder. Ona göre, organik olmayan ya da insan beyninin özel nedensel özelliklerine sahip olmayan bugünün teknolojileriyle gerçek anlamda düşünen bir sistem oluşturulamaz. Robert Pepperell ise bugünün makinelerin bilinç, kavrama ve ödül gibi biliş fonksiyonlarını gerçekleştiriyor olmasının *yapay anlamayı* olanaksız kıldığını savunur. Burada özellikle John Searle’ün yapay *güçlü* ve *genel* zekâyı ilgilendiriyor gördüğünün altını çizmekte fayda var. O, makinelerin düşünmeyi simüle etmesinin ötesinde, makinelerin insanlar gibi düşünmesi problemini ele almaktadır. Oysa bu yazı, daha önce de belirttiği gibi, yapay *zayıf* genel zekâ ile ilgilidir. Dolayısıyla, insan gibi düşünmeyen ama insan gibi düşünüyor gibi davranan sistemlerin imkânlılığı bu yazının temel problemdir. Bu durum her bilişsel becerilerin kıvılcımlarını sergilemek tezimizin John Searle’ü haksız çıkarmadan doğrulanabilmesinin önünü açmaktadır.

Yapay Sinir Ağları Nasıl Çalışmaktadır?

Bugünün yapay sinir ağlarının ve özelde büyük dil modellerinin, her

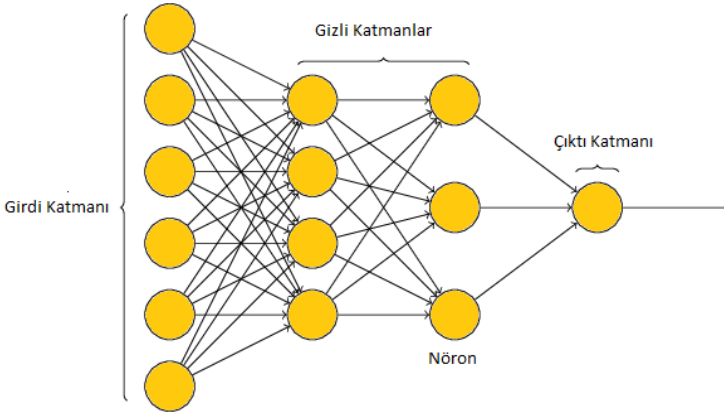


bilişsel becerinin kıvılcımlarını sergileyip sergilemediği sorunun cevabını daha sonra verebilmek için, burada modern büyük dil modellerinin (BDM) nasıl çalıştığını basit ve sezgisel bir düzeyde anlatmakta fayda var. Konumuz olan BDM'ler transformatör mimarisinde yapay sinir ağlarıdır. Yapay sinir ağları (YSA), insan beyninin çalışma şeklini taklit etmek için tasarlanmış bir makine öğrenimi modeli türüdür. Ancak burada, YSA'ların sinir sisteminin bazı özelliklerini taklit ettiğini belirtirken, bu taklit etmenin bir ana fikri alıp birçok detayı göz ardı ettiğinin altını çizmekte fayda var. Her şeyden önce, bugün sinir sisteminin nasıl çalıştığına dair bilmediğimiz birçok bileşen var. Belirsizlik alanlarından biri, glial hücrelerin bilgi işlemdeki rolüdür. Glial hücreler, geleneksel olarak nöronlar için destek hücreleri olarak görülürken, bunların nöral iletişimde ve belki de bilişte daha doğrudan bir rol oynayabileceğine dair artan kanıtlar ortaya çıkmaktadır (Allen & Eroglu 2017; Fields, Woo & Basser, 2015). Dahası, nöroplastisitenin öğrenme ve için hayati önem taşıdığı bilinirken, bu değişikliklerin moleküler düzeyde nasıl ve neden meydana geldiği hakkındaki ayrıntılar hâlâ araştırma konusudur (Price & Duman, 2020). Son olarak bu bilgiye sahip olsaydık ve beyni olduğu gibi simüle etmek isteseydik bile, bunu yapmanın gerektirdiği hesaplama kapasitesine sahip değiliz. Öyleyse YSA'lar sinir sisteme dair her şeyi değil, sadece bazı ana mekanizmayı simüle etmeye çalışmaktadır. Bu ana mekanizmaların arkasındaki temel öncüller, 1) nöronlar arasında bağların olduğu, 2) bu bağların niceliğinin zaman içerisinde ateşleme örüntülerine göre değiştiği ve 3) bu değişim sonucunda da nöronlar arasındaki bağların verinin karşılığı olan gerçeklikle uyumlu hale gelerek 4) girdiler için doğruya yakın bir çıktılar vermeye başladığıdır. Bu öncüller, sinir sisteminin karmaşık etkileşiminin çok küçük bir parçasını ifade etmektedir. Öte yandan, makine öğrenmesinin böyle yüzeysel bir esinlenmeyle harekete geçmesinin bir avantajı da vardır: Makine öğrenmesi, sinir sisteminden aldığı esini temel ve basit bir düzeyde tutarak, sinir sisteminden ödünç aldığı bileşenleri, matematiksel olarak savunulabilir ve ideal bir formda tutar ve böylece bu bileşenlerin analitik olarak geliştirilmesinin önünü açar.

YSA'ların insan sinir sisteminden bu önemli farkının altını çizdikten sonra, burada YSA'ların nasıl çalıştığından söz edilecektir: Bir YSA, birbirine bağlı çoklu yapay nöron katmanlarından oluşur. Her bir katman bir



önceki katmandaki nöronlardan girdi sinyalleri alır, bu girdiyi bir aktivasyon fonksiyonunu kullanarak işler ve ardından çıkış sinyalini ağına bir sonraki katmanındaki diğer nöronlara iletir. Ağdaki nöronlar arasındaki bağlantılar, organik nöronlar arasındaki iletişimi sağlayan sinaplara karşılık gelir. Zaman içerisinde bu bağlantıların gücü, yani ağırlıkları değişir. Bu da hesaplamanın genelinde bazı girdilerin diğerlerinden daha önemli olmasına olanak sağlar.



Şekil 1: Basit bir yapay sinir ağı. Bu yapay sinir ağının girdi katmanında 6, çıktı katmanında 1, aradaki gizli katmanlarında 4 ve 3 nöron bulunmaktadır.

Bir YSA'nın işlevi, her bir nöron katmanının girdi sinyalini daha üst-düzye bir temsile dönüştürerek, adım adım çıktı katmanına taşınması ve her şey yolunda giderse çıktı katmanının tahmin ettiği değişkenin gerçek değere yakınsamasıdır. Söz gelimi, Şekil-1'deki 6 girdi, bir şehrin yüzölçümü, nüfusu, kişi başına düşen geliri, ortalama eğitim düzeyi, Gini endeksi (bir gelir dağılımı eşitlik endeksi) ve şehirdeki üniversite sayısı, sinir ağının çıktı katmanında tahmin etmeye çalıştığı değişken de o şehirdeki suç oranı olabilir. Bu durumda çıktı katmanındaki nöron suç oranını temsil edecektir. Farklı girdiler için girdiler ve ağırlık değerlerine bağlı olarak farklı çıktılar hesaplanacaktır.

Bir YSA'nın eğitimi, yani öğrenme süreci, genellikle, ağ tarafından üretilen gerçek çıktı ile istenen çıktı arasındaki farkı azaltmak için nöronlar arasındaki bağlantıların ağırlıklarının ayarlanmasını içerir. Bu genellikle geriye doğru yayılım (backpropagation) olarak bilinen bir teknik kullanılarak



yapılır. Yani yapay sinir ağına tahmininin ne kadar iyi olduğuna bakılır ve buna göre sinir ağına parametreleri değiştirilerek daha iyi bir tahmin yapılması sağlanır. Böylece yapay sinir ağı her yeni veriyle birlikte dönüşmekte ve bir süre sonra doğruya daha yakın bir çıktı hesaplamaya başlamaktadır.

YSA'ların nasıl çalıştığına dair bu genel resmi sunduktan sonra, burada transformatör mimarili doğal dil işleme sistemlerinin nasıl çalıştığına dair yine ana hatlarıyla bir iki temel bilgi verilecektir: Doğal dil işleme sürecinde kelimelerin nasıl vektörlere çevrildiği ve nasıl olup da birbirinden farklı uzunluktaki girdilerin işlendiği gibi teknik konulara girilmeyecek ancak basitçe GPT serisi gibi otoregresif dil modellerinin nasıl eğitildiği ve nasıl tahminde bulunduğu sezgisi verilecektir. Otoregresif modeller, bir söz dizisinin ardından hangi söz parçasının geleceğini tahmin etmek üzere eğitilirler. Tam olarak gerçeği yansıtmamasına rağmen, otoregresif modellerin, basitleştirilmiş hâliyle, bir dizinin ardından hangi kelimenin geleceğini tahmin ettiği varsayımı, bizim amacımız açısından bir sorun oluşturmamaktadır. Öyleyse, otoregresif dil modelleri “Her sabah uyandığımda mutlaka kahve içerim; çünkü kahve kendime” gibi bir söz dizisini girdi olarak alarak, bu diziyi “gelmeme” gibi duruma uygun sözlerle sürdürmeye çalışmaktadır. Büyük ölçekte bir otoregresif dil modelleri, eğitilirken internette yer alan neredeyse tüm söz dizileri için böyle sayısız tahminde bulunmaya çalışmaktadır. Yanlış tahminler yaptığında yapay nöronları arasındaki ağırlıklar, daha doğru tahminde bulunması için, geriye doğru yayılım yöntemiyle tekrar ayarlanır. Eğitildikten sonra, otoregresif dil modeli, dizideki önceki kelimelere dayanarak bir sonraki kelimeyi tahmin eder ve metin oluşturur. Model, her kelimeye, dizide bir sonraki olarak görünme olasılığına dayalı olarak bir olasılık atar ve bu olasılıklara göre sonraki kelimeyi belirler. Model dizideki bir sonraki kelimeyi atadıktan sonra, nu yeni kelimeyi de girdileri arasına alır ve daha sonraki kelimeyi tekrar tekrar tahmin eder. Bu süreç, otoregresif dil modelinin tutarlı, dilbilgisi kurallarına uygun, mantıklı ve bağlama uygun metin oluşturmasını sağlar.

Anlamanın İki Temel Koşulu

Anlamanın temel şartlarını tartışmaya başlamadan önce, büyük dil işleme modellerinin anlamaya göre zayıf epistemik koşulları olan bilmeyi



simüle edip edemeyeceğini tartışmak yerinde olacaktır. Platon'un bilgi anlayışı büyük dil modellerinin bilgiye sahip olup olamayacağına dair tartışmalar için yararlı bir başlangıç noktası teşkil edebilir. Bu anlayışa göre, bilgi gerekçelendirilmiş doğru inançtır ve şu şekilde analizi verilebilir: “*S, P'nin doğru olduğunu bilir ANCAK VE ANCAK (1) P doğrudur, (2) S, P'nin doğru olduğuna inanır ve (3) S, P'nin doğru olduğuna inanmak için bir gerekçesi vardır*” (Gettier, 1963, s. 121). Daha önce dile getirildiği gibi, bu makale sadece anlamanın ve bilmenin simülasyonu ile ilgilenmektedir ve dil modellerinin bilinç veya inanç gibi zihinsel durumlara sahip olup olamayacağına dair soruları ele almaz. Bilmenin simüle edilmesinin analizi, yukarıdaki analizden hareketle şöyle ifade edilebilir: *S, P'nin doğru olduğunu bilme durumunu simüle eder ANCAK VE ANCAK (1) P doğrudur, (2) S, P'nin doğru olduğuna inanıyormuş gibi davranır ve (3) S, P'nin doğru olduğuna inanmak için bir gerekçe sunabilir*. Burada kritik olan şart (3) yani gerekçe sunabilme şartıdır ve yakın dönem çalışmaları büyük dil modellerinin gerekçe sunabildiklerine işaret etmektedir (Ma, vd., 2024). Büyük dil modelleri $E=mc^2$ 'ye inanıyormuş gibi davrandığı ve bunun için gerekçeler sunabildiği için, bilinç ve inanma gibi bileşenler göz ardı edildiğinde, onların bilmeyi simüle edebildiği sonucuna ulaşılabilir.

Felsefi çalışmalarda anlamanın bilmeye göre daha fazla çaba gerektirdiği ve daha büyük bir epistemik değere sahip olduğu konusunda bir fikir birliği varmış gibi görünmektedir (Lear, 1988; Kvanvig, 2003; Pritchard, 2009). Linda Zagzebski (2019) anlamanın nesnesinin mantıksal-dilbilimsel ilişkiler (Kim, 1994) ya da nedensel ilişkiler değil, örüntü ya da yapılar (structures) olduğunu ve bilginin özel bir anlama türü olduğuna inanmaktadır. Ancak bilgiyi anlamanın bir türü sayan bu sıra dışı anlama ontolojisini savunan Zagzebski bile, bilme yoluyla elde edilen anlamanın çoğu durumda diğer anlama türlerine göre daha zayıf olduğunu kabul etmektedir.

Linda Zagzebski'ye göre anlama süreci, kavradığı şeyi basitleştirir:

Bir coğrafi bölgenin haritası, anlama yetisi hakkındaki bu noktayı iyi bir şekilde örneklendirir. Bir harita, haritalanan fiziksel bölgenin büyük bir kısmını çıkarır. Büyük bir bölgenin haritası, küçük bir bölgeninkinden daha fazla detayı çıkarır çünkü çok fazla detayın dahil edilmesi, zihnin kavrayabileceğinden daha fazlasını içerir ve bu da anlamaya yardım etmez. (Zagzebski 2019, 125)

Öyleyse anlamak dünyanın karmaşıklığının sonucu olarak gereklidir ve



1) *anlamanın bileşenlerinden birisi karmaşık dünyayı basitleştirmektir.*

Zagzebski'nin altını çizdiği bir diğer konu da anlamanın tekrar eden örüntüleri keşfetmekle ilgili olduğudur. İnsan anlayışı bir yapıyı bambaşka iki alanda bulup bunları eşleştirebilmektedir. Bunu göstermek için Zagzebski altın orandan bahsetmektedir: Altın oran, doğanın farklı düzeylerdeki çok sayıda parçasında, “taç yaprak ve yaprakların gelişiminde, Samanyolu Galaksisinde, Leonardo'nun Vitruvius Adam'ında ve Fibonacci dizilerinde” görülmektedir. Öyleyse iyi bir anlama için ortaklıkları keşfetmek önemlidir.

Zagzebski'nin bu savı, Michael Friedman (1974) ve Philip Kitcher'in (1981) analiziyle de uyumludur: Michael Friedman (1974) ve Philip Kitcher'in (1981) *birleştirme* kavramının analizindeki ayrılıklarına rağmen, anlama ile bilimsel açıklama arasındaki ilişkiyi ortaya koyarken, anlamanın bilimsel açıklamanın birleştirme fonksiyonunu desteklediğini öne sürmüşlerdir (Kim 1994). Bu iki filozof için, epistemik başarının ölçütü holistiktir: tüm sistemin nasıl organize edildiği ve birleştirici bir yapı oluşturulup oluşturulmadığı ile son derece ilişkilidir (Kim, 1994, s. 64). Öyleyse 2) *anlamak tüm bilgiyi organize etmeyi ve farklı alanlardaki ortaklıkları birleştirebilme yetisini gerektirir.*

Zekâ Bilinç ve Duyguları Gerektirmekte midir?

Robert Pepperell (2022) üç tür anlamadan bahsetmektedir: *Doğal Anlama*, *Yapay Anlama* ve *Makine Anlaması*. Doğal anlama basitçe insan anlamasıdır, sinir sisteminde gerçekleşmektedir ve Pepperell'e göre anlama algoritmalarının nihai hedefidir. Yapay anlama bugünün yapay zekâ sistemlerinin gerçekleştirmeye çalıştığı bu doğal anlama çıktılarını taklit etmek işidir. Ancak bugünün yapay zekâ sistemleri, hiç değilse Pepperell'e göre bu taklit etme işini layığıyla yapamamaktadır ve bunun bir sonucu olarak doğal anlamayı da gerçekleştirememektedirler. Yani, bu sistemler ne yapay zayıf genel zekânın ne de Yapay Güçlü Genel Zekânın örneği değildir. Makine anlaması ise, Pepperell'e göre, biyolojik olmayan bir maddesel temelde gerçekleşecek olan insansı anlama kapasitesidir. Bu kavramsal zeminde, yapay anlamanın makine anlaması yolunda sadece bir adım olduğu söylenebilir; çünkü makine anlaması yapay anlamadan farklı olarak insan kadar ve insan gibi anlamadır.



Pepperell'e göre yapay anlamanın ve doğal anlamanın bazı ortak özellikleri (öğrenme, tanıma, ayırma, birleştirme, bağlam belirleme, akıl yürütme, öngörme) varsa da sadece doğal anlamanın gösterdiği üç ayrııcı özellik de bulunmaktadır: Bunlar bilinç, kavrayış (insight) ve ödüldür. Pepperell kavrayışı şöyle tanımlıyor: "Aha! anı veya bir uyarının algılanışında aniden meydana gelen bir değişim ve bunun sonucu olarak, önceden mevcut olmayan yeni bir anlamın keşfiyle ilgili bir aydınlanmayı içeren durum." (Pepperell, 2022, s. 5) Peki bu Aha! anı neden bugünün öğrenme algoritmalarında bulunmamaktadır? Öyle ya, bugünün makine öğrenmesi algoritmaları da hiç değilse basit bir öğrenme sürecini simüle ederek, yeni önermelere inanmayı hiç değilse simüle etmektedirler. Pepperell'e göre yapay zekâ algoritmalarında eksik olan tam olarak nedir? Bir kere, Pepperell'in kullandığı anlamda "kavrayış"ın bir deneyim durumu olduğunun ve ödül duygusuyla ilişkili olduğunun altını çizmek gerekiyor. Dahası, Pepperell'in YSA'larda eksik bulduğu ikinci bir bileşen ödül de açıkça duygularla ilgilidir. Bu çeşit duygusal deneyimlere sahibi olmak için yapay zekâya sistemleri oluşturmaya ötesinde, yapay *duygu* sahibi sistemlerin oluşturulmasını gerekmektedir. Dahası duyguların da bilinci gerektirdiği savı üzerinden (Nahmias, Allen & Loveall, 2019), Pepperell'in *kavrayışının* bilinç gerektirdiği söylenebilir. Bu durumda YSA'larda eksik olanın bilinç ve duygular ya da duygular bilinci gerektiriyorsa sadece bilinç olduğu söylenebilir.

Öte yandan, yapay zekâ sistemlerinin bilinçli ya da duygu sahibi (sentient) olduğu zaten oldukça savunulması zor bir görüştür. Yapay duygular ya da yapay bilinç hakkında ortaklaşılacak teorilere ulaşmak için daha yolumuzun olduğu söylenebilir. Neredeyse hiçbir felsefeci ya da bilim insanı, problem çözmek için tasarlanmış bu sistemlerde bilinç ya da duyguların zühur ettiği savına inanmamaktadır. Genel yapay zekâya ya da anlama kabiliyetine, çözümü konusunda hatırı sayılır adımlar atmadığımız yapay bilinç ya da yapay duygular gibi şartlar koymak, bizi YGZ'lerin öngörülebilir bir zamanda çözüm üretilemeyeceği, bugünün yapay zekâ sistemlerinin genel zekâ ya da anlama açısından değerlendirilemeyeceği kati sonucuna götürecektir. Ve bu durumda, Pepperell gibi düşünürlerin yaptığı aslında malumu ilam etmek olur. Öte yandan, Pepperell günümüz YSA'larının sorunlarını ortaya koyarak, sadece bu makinelerin anlamadığını, ya da genel zekâya sahip olmadığı değil anlamayı ya da genel zekâyı simüle etmekte de yetersiz



kalacağını savunuyor gibi görünmektedir. Bu yazıda, bunun aksi savunulacaktır. Yapay duyu ve bilinç problemleri çözülmeden de anlamayı simüle etmek mümkündür.

Zekânın tanımlanması işi, psikologlar, filozoflar ve yapay zekâ araştırmacıları için çözümlenmesi güç bir zorluktur. Farklı zekâ kavramlarının tam bir analizini vermek, geniş kapsamlı bir analiz gerektirir ki, bu yazıda bunu gerçekleştirmek mümkün değildir. Problem çözme, çeşitli ortamlarda adaptif davranışlar gösterebilme ve hedeflere ulaşma fonksiyonlarına odaklanan tanımlardan; öğrenme, yaratıcılık, akıl yürütme, bellek gibi bilişsel yeteneklere odaklanana kadar çeşitli tanımlar öne sürülmüştür (Gottfredson, 1997, 1998; Neisser, vd., 1996; Binet & Theophile, 1973). Zekânın birçok tanımı duygular ya da bilince referans vermez; ancak bunun sebebi duygular ve bilinci varsayımları ve zekanın yeterli koşuluyla değil gerekli koşuluyla ilgileniyor olmaları olabilir. Çünkü genellikle, hem yapay zekâ sistemleri de içinde olmak üzere insan-dışı zeki sistemlere hem de insan zekâsına uygulanabilecek evrensel bir zekâ kavramı yerine insan zekâsıyla ilgililerdir.

Burada, genellikle insan zekâsına yoğunlaşan ve var olan kavramı analiz etmeye çalışan zekâ tanımlarına (*lexical definition*) çok fazla odaklanılmayacaktır, çünkü mevcut YGZ tartışmasına katkıda bulunmak yapmamız gereken, temelde evrensel zekâyı odaklanan spekülâtif tanımlardır. Bir yapay zekâ araştırmacısı perspektifinden, Shane Legg ve Marcus Hutter (2006, s. 2) zekânın “...bir ajanın [agent] geniş bir çevre yelpazesinde hedeflere ulaşma yeteneği...” olduğunu speküle etmiştir. Alexander D. Wissner-Gross (2014) halka açık bir konuşmasında, “zekâyı *gelecekteki eylem özgürlüğünü maksimize etme gücü* belirler diyerek” evrensel bir tanım vermede bir adım daha öteye gitmiştir, ki bu da bilinç ya da duyguları gerektirmeyen bir başka tanımdır.

Pek tabii, bu tanımlarda geçen *amaç*, *özgürlük* gibi kavramların istekleri ve dolayısıyla duyguları gerektirdiği bir haklılık payıyla savunulabilir; ancak zekâ birçok durumda duyguları göz önünde bulundurularak karar verse de duygulardan kaynaklanmak zorunda değildir. Hiç değilse bu yazının kavramsal varsayımı budur. Zekâ pek tabii problem çözmek için istek, duyu, amaç oluşturmaya karar verebilir. Söz gelimi, “Bir işi başarmak istiyorsan önce onu istemelisin” önermesi zeki bir sistem tarafından üretilebilir.



Öyleyse zekâ bazen duyguları üretir ve duygular bazen zekânın bir parçasıdır ama bu, duygu sahibi olmanın veya duygu üretme işinin zekânın zorunlu bir bileşeni olduğunu göstermez. Legg, Hutter ve Wissner-Gross'un bahsettiği amaç ya da özgürlüğün zeki sisteme içsel olması gerekmez. Yani yapay zekâ sistemleri kendileri yerine insanların özgürlüğünü ya da amaçlarına ulaşmadaki başarılarını maksimize etmeleri onların zeki olmadığı anlamına gelmez. Dahası hiçbir duygusu olmayan bir robot, kendisine böyle bir amaç da atanmamışken, varlığını korumak için zekâsını kullanıyor, söz gelimi şarjı azaldığında pilini şarj ediyorsa, zeki olduğu söylenebilir.

Öyleyse zekâyı fonksiyonu ile el alan bir kavramsal zeminde, bilinçsiz ve duygusuz zekâ mümkündür. Duyguların problem çözmede bir kısa yol fonksiyonu olsa da (Damasio, 1994, Antoine, Damasio & Damasio, 2000), birçok durumda çözümün değil problemin parçasıdır. Dahası duyguların bir problemi çözmeye katkı sağladığı her durumda, aynı problem duyguların yoksunluğunda, duygusuz akıl yürütmeyle de çözülebilir. Söz gelimi doğada, bir bitkinin tadı bize rahatsızlık hissi verdiği için ondan uzak duruyorsak, bu kimyasal bazı özelliklerinden kaynaklanıyordur ve bu problemi duyguların aracılığı olmadan, biyokimya bilgisiyle de çözmek de mümkündür. Öyleyse, duygular olmadan problem çözen bir sistem tahayyül edilebilir. Dahası bilinç olmadan da problem çözen sistemler tahayyül edilebilir. Nitekim günümüz yapay zekâ sistemleri tam da bu türden sistemlerdir.

Son olarak, bilinç ya da duyguları gerektirmeyen bir zekâ kavramı tutarlı olmasına ek olarak faydalıdır da: Açıkça görülmektedir ki, gün ve gün yapay zekâ sistemleri daha geniş alanlarda daha genelleştirilebilir çıkarımlar yapabilmektedir ve bu yeni sistemler, geçmişteki ilkel sistemlerden epistemik işlevleri ve altlarında yatan algoritmalar açısından farklılıklar göstermektedir. Öyleyse buradaki gelişmeleri nasıl incelememiz gerekir? Yapay zekâ sistemlerinin zayıf anlamda ama genel zekâyı simüle edip edemediğiyle ilgilenen araştırmacıların bu dönüşümü değerlendirmek için işe yarar bir kavrama ihtiyaçları vardır. İşte bu kavram, tam da bilinç ve duygular gibi içeriklerinden ayrıştırılmış zayıf bir zekâ kavramıdır.

Bugünün Yapay Sinir Ağlarının Genel Zekâ Açısından Değerlendirilmesi

Bugünün YSA'larının genel zekânın bütün kıvılcıklarını gösterip göstermediğini tartışmadan önce, iki noktanın belirlenmesinde fayda vardır:



Bunların birincisi bu yazının savunduğu bir pozisyonu tanımlamak için kullanacağım ve “zayıf fonksiyonalizm” diye anacağım tezdır. Fonksiyonalizm, zihin felsefesinde, zihinsel durumların öznenin fiziksel bileşenlerine direkt bağlı olmadığını, bunun yerine işlevleriyle ilişkili olduğunu ileri süren bir savdır. Bir fonksiyonalist, zihinsel durumların içsel özellikleriyle değil, hangi fonksiyonu gerçekleştirdiğiyle tanımlandığını savunur. Savunduğum “zayıf fonksiyonalizm,” fonksiyonalizmin aksine zihinsel durumları işlevsel rolleriyle eşleştirmemektedir. Bunun yerine, zayıf fonksiyonalizm teziyle, zekâ, anlama, düşünme gibi becerileri *simüle etmek* için en iyi yöntemlerden birisinin bu yetilerin altında yatan bilişsel becerilerin işlevlerini taklit etmek olduğunu önerir. Zayıf fonksiyonalizm tezi fonksiyonalizm tezinden çok daha pratik ve yapay zekâ mühendisliğiyle ilişkili bir kavramdır.

Tartışmaya başlamadan önce ikinci vurgulanacak nokta da bu yazıda insan-merkezci zekâyâ karşıt olarak evrensel zekâ kavramının neden tercih edildiğidir. Bir uzaylı ya da bir makine fiziksel mekanizması açısından, hem de düşünme ya da problem çözme fonksiyonları açısından çok farklı yollar izliyor olabilir. Bunu göz önünde bulundurarak, makinelerin zekâsını speküle ederken evrensel bir zekâ tanımı tercih edilmelidir. Bunu yapmadığımız sürece makineler düşünebilir mi ya da bir insanın yaptığı işleri yapabilir mi sorusuna yanlı bir biçimde başlamış oluruz ve aşağıdaki türden, bana göre, üretken olmayan argümanların önünü açmış oluruz.

İnsan bilişinin *X* özelliği var.

X özelliği makinelerde yok.

Öyleyse makineler düşünemez.

Söz gelimi insanın duyguları var. Makinelerin yok. Öyleyse makineler düşünemez. İnsanın yönelimselliği var. Makinelerin yok. Öyleyse makineler düşünemez. John Searle’ün (1980) ifade ettiği gibi, insan beyni organik, makineler değil. Öyleyse makineler düşünemez. Bu türden argümanlar hiç şüphesiz savunulabilir; ancak bu argümanların odaklandığı duygu, organik olma, yönelimsellik gösterebilme gibi özelliklerin insan zekâsının birer bileşeni olduğu savı, tek başına çok önemli bir akıl yürütme değeri taşımamaktadır. Bu bileşenlerin insan zekâsının bir parçası olması onların (evrensel) zekânın bir şartı olduğunu göstermez. İnsanların duygular, yönelimsellik, ya da bilinçle gerçekleştirdiği fonksiyonları söz gelimi bir makine ya da



uzaylı tarafından başka bileşen ve yollarla gerçekleştiriyor olabilir. Bu tür argümanları savunabilmek için, duygu, bilinç ya da yönelimselliğin evrensel bir zekâ için neden gerekli olduğunun ortaya konulması ve bunun ya da onun yerini tutar (fonksiyonunu yerine getiren) bir bileşenin makinede neden var olamayacağının gösterilmesi gerekmektedir.

Söz gelimi Searle (1980) beynin nedensel özelliklerin sahip olmayan bir makinenin düşünemeyeceğini iddia ediyor ama bu nedensel özelliklerin ne ya da ne türden özellikler olduğunu ikna edici bir biçimde ortaya koyamıyor. Bu özellikleri tam olarak ortaya koymadığımız sürece bunların fonksiyonlarının bir silikon bazlı makine tarafından gerçekleştirilip gerçekleştirilemeyeceğini sorgulamak dolayısıyla da makinelerin düşünemeyeceğini temellendirmek mümkün olmayacaktır.

Öte yandan duyguları tartışacak olursak, yukarıda da ortaya koyduğumuz gibi gerçekten de duygular insanların doğru karar vermesine birçok durumda olanak sağlamaktadır ve makinelerin duygu sahibi olması bugün mümkün değildir. Öyleyse şöyle özensiz bir argüman ortaya atılabilir:

İnsan duyguları vardır.

Makinelerin duyguları yoktur.

Öyleyse makineler düşünemez.

Ancak böyle bir argümanın savunulması için öncelikle duygu sahibi olmanın *evrensel* bir zekânın neden şartı olduğu ortaya konulmalıdır. Daha sonra pekiştirmeli öğrenmede kullanılan *ödül* kavramının ya da makine öğrenmesi modellerinin tahminlerinin gerçek değerden ne kadar farklı olduğunu gösteren *kayıp* kavramının neden duyguların fonksiyonunu gerçekleştiremeyeceğini ortaya konulmalıdır. Aksi takdirde, duygu sahibi olmayan makinelerin insan gibi zeki olamayacağı ya da anlayamayacağı sonucuna gitmek aceleci bir akıl yürütme olacaktır.

Şimdi bütün bu tartışmaların ardından, daha somut bir biçimde makinelerin genel zekâ niteliklerini gösterip gösteremeyeceği tartışılabilir. Daha önce ifade edildiği gibi, bugün popüler olan büyük dil modellerinin eğitim ya da öğrenme süreci, mevcut söz dizimine göre söz diziminin nasıl devam edeceğinin tahmin edilmesine dayanır ve işlevi de tam olarak budur. Peki, bütün işlevi daha önce gördüğü söz dizilerini göz önünde bulundurarak sonraki kelimeyi tahmin etmek olan bir sisteme zeki demek ne kadar



doğrudur? Emily M. Bender ve Alexander Koller (2020) anlama dilsel formla (linguistic form) ulaşamayacağını ifade ederken; Melanie Mitchell (2019) YSA'ların anlam bariyerine çarptığını, bu sınırı geçmek için ise, mevcut makine öğrenmesi algoritmalarını küçük adımlarla geliştirilmesinin yerine, insanın bilişsel özelliklerinin daha iyi anlaşılmasının gerektiğini ifade etmiştir. Bu ve benzeri karşı çıkışların birçoğunun arkasında geçersiz olduğunu savunacağım şu argüman vardır: “YSA'ların yaptığı, çok başarılı bir biçimde bile olsa, istatistiki bir yakınsamadan ibarettir. Sadece bir sözdiziminin ardından bir kelimenin geliyor olduğunu tahmin etmeniz de sizi o söz dizimini anlıyor yapmaz.”

YSA'ların sadece istatistiki yakınsamalar yapıp yapmadığı bir yanda dursun, en azından YSA'ların yaptığı her işin istatistiki yakınsamalara dayandığı doğrudur. Öte yandan bunun insan sinir sistemi için de geçerli olduğu savunulabilir: İnsan sinir sistemi sinir hücresi boyutunda basit neden-sonuç ilişkileriyle çalışıyorsa, nasıl olup da anladığını açıklayamayacak hiç de kolay bir iş değildir. Ancak nöronlar arasındaki etkileşimlerin bir sonucu olarak insan bir biçimde holistik bir perspektiften anlama kabiliyeti göstermektedir. Bunun YSA'lar için de böyle olması mümkündür. Mesele tek tek istatistiki işlemlerin ne kadar mekanistik ya da basit olduğu değil, tüm bu işlemlerin nasıl organize olduğu olabilir. Eğer YSA'lar her durum için basit bir istatistiki hesap yapıp yazıyı geçmişte gördüğü yazılara bire bir ve basit bir yolla benzetiyorsa ve anlamayı simüle edemiyorsa, sadece eğitilirken gördüğü metinlerin benzerlerini başarıyla tamamlayacak ancak benzer semantik yapıları içeren yeni söz dizilerinde başarısız olacaktır (aşırı öğrenme). Öte yandan bugünün büyük dil modelleri eğitilirken görmediği giridilleri işlemede hatırı sayılır başarı göstermektedir.

Yapay sinir ağlarını diğer makine öğrenmesi algoritmalarından daha ilginç yapan bir niteliği gerektiğinde gizli katmanlarda yeni değişken (feature) üretebilmesidir. Yani YSA tahminde bulunmak için eğitilirken, sadece son katmandaki çıktı nöronunu değil ara, gizli katmanlardaki nöronların da anlamlı değişkenler haline gelebilmektedir. Bu ara katmanlarda neler öğrenildiği sorusuna cevap vermek kara kutu problemi nedeniyle zor bir iştir; ancak geçmiş söz dizilerine bakarak eldeki sözdizimini tamamlamak, dilbilgisine, kültüre, anlama dair birçok bilgiyi gerektirdiği ortadadır. Söz gelimi, “Onların ahlak anlayışları bizimkinden farklıdır. Kendilerine



yapılmasını istemedikleri bir şeyi başkasına...” sözünü “yaparlar.” sözüyle tamamlamak işinin salt bir istatistik yakınsamasıyla açıklanması çok zor bir iştir. Bu görev, genel geçer ahlak anlayışına dair bir fikir sahibi olmayı, noktalama işaretlerini doğru kullanmayı ve şu anda sayamayacağımız çok sayıda bilgiyi holistik bir biçimde bir araya getirerek sonraki kelimeyi tahmin etmeyi gerektirir. Bu da istatistiki yakınsama indirgenebileceği; ama istatistikle açıklanması son derece zahmetli niteliklerin zuhur ettiği bir sisteme işaret eder. Öyleyse bir sonraki kelimeyi tahmin etmek gibi hafife alarak ifade edebileceğimiz bir görev son derece incelikli çok sayıda görevin bir araya gelmesini gerektirmektedir ve büyük dil modelleri de bu alanda hatırı sayılır beceri sergilemektedir.

Son olarak anlamının daha önceki bölümlerde belirlenmiş olan iki niteliğinin yapay sinir ağlarında görülüp görülmediğini değerlendirilecektir:

- 1) *Anlamanın bileşenlerinden birisi karmaşık dünyayı basitleştirmektir.*
- 2) *Anlamak tüm bilgiyi organize etmeyi ve farklı alanlardaki ortaklıkları birleştirebilme yetisini gerektirir.*

Öncelikle yapay sinir ağlarının (YSA) 1. koşulun nasıl sağladığına bir bakalım: YSA'lar bu basitleştirme işini en az iki yolla yapmaktadır. İlk olarak, tahmin doğruluğuna dayalı olarak ağırlıkların ayarlanması sürecinde yani geriye doğru yayılım sürecinde, yapılan tahminle ilişkisiz olan değişkenlerin önemsizleşmesi sağlanır. Bunu bir çeşit önemsiz değişkenlerin filtrelenmesi süreci olarak görülebilir ve dolayısıyla bir *basitleştirme* sürecidir.

İkincil olarak, daha önce de ifade edildiği gibi, YSA'ların her katmanı, önceki katmandaki verinin daha üst düzey temsillerini oluşturur ve böylece karmaşık girdilerin basitleştirir. Her katman, girdiden yararlı bilgileri bulur, birleştirir ve yeni değişkenler oluşturur. Bu süreç, değerli bir madde elde etmek için birçok karışımı damıtmak ve elde edilen saf maddeleri kimyasal reaksiyonlara sokmaya benzetilebilir. Dolayısıyla çok sayıda gürültülü (noisy) bir değişken kümesini az sayıda sade bir değişken kümesine çevirmesi anlamında bir *basitleştirmedir*.

YSA'ların paradigmatik örneklerinden birisi olan rakam tanıma görevi bu üst düzey temsil oluşturmaya bir örnek olarak verilebilir. Bir yapay sinir ağı el yazısı ile yazılmış rakamları tanımaya çalışırken, ham girdi, tekil piksellerin sayısal değerleri olacaktır. Ancak rakamı doğru bir şekilde



sınıflandırabilmek için, yapay sinir ağı, girdi resminden daha anlamlı özellikler çıkarmalıdır. Bunlar, rakamın şekli (yuvarlak mı, düz mü, kuyruğu var mı), ve çizgilerin konumu gibi özellikler olabilir. Bu arada sinyaller bir sonraki katmana taşınırken, el yazısının kişiye özgü farklılıkları ya da el yazılarının kenarındaki bir damla mürekkep gibi ilgisiz değişkenler de elimine edilecektir. Öyleyse ağırlık ayarlamalarının ve üst düzey temsillerin oluşturulmasının kombinasyonu, YSA'nın gerçek dünya verilerinin karmaşıklığını basitleştirmesi yeteneğini sağlar ve bu özellikleri, örüntü tanıma ve tahmini modellemedeki başarısına önemli ölçüde katkıda bulunur.

2. koşula gelindiğinde, derin öğrenmenin “kara kutu” problemi nedeniyle başlangıçta bir sonuca ulaşmanın daha zor olduğu düşünülebilir. Bu “kara kutu” problemi, YSA'ların kararlarını nasıl verdiklerini anlamamanın zorluğunu ifade eder. Ancak, giderek daha net bir şekilde görülüyor ki, YSA'lar karmaşık veriyi sadeleştirmenin yanı sıra, tek bir ağırlık sınırları içinde farklı problemler veya alanlar arasındaki ortaklıkları bulma ve bilgiyi organize etme yeteneğini de gösterir. Bu özellikle, multimodal ağların ortaya çıkışı ve transfer öğrenme tekniklerinin başarısıyla belirgin hale gelmiştir. Multimodal ağlarda, bir YSA, farklı türlerdeki giriş verileri (görsel, dilsel, işitsel, vs.) üzerinde eğitilir. Bu nedenle çeşitli modaliteler arasında paylaşılan temsilleri öğrenip kullandığı speküle edilebilir.

Ayrıca, transfer öğrenme yoluyla, bir görev üzerinde eğitilmiş bir sinir ağı, kazandığı bilgiyi farklı ancak ilgili görevlere uygulayabilir. Bu, YSA'ların görevler arasında ortak olan alt düzey özellikleri “yakalayıp,” ilk problemde öğrenilen oluşan “sinir yollarını” yeniden kullandığına işaret sayılabilir. Söz gelimi obje tanıma oldukça başarılı olan bir temel model, farklı yüz ifadelerini tanımak için ek katmanlarla eğitilebilir ve bu, sıfırdan öğrenmeye kıyasla çok daha hızlı ve başarılı sonuçlar verir; çünkü yeni model yani yüz ifadesi tanıma modeli, obje tanıma temel modelinin alt düzey örüntülerini kullanır.

Sonuç

YSA'ların bir problemde diğerine genelleme yapabilme yetisi, insanların bir alanındaki bilgilerini başka bir alanda kullanabilme yetisine karşılık gelmektedir. Dolayısıyla, YSA'ların, tüm bilgileri düzenleme ve farklı problemler veya alanlar arasında ortak örüntüleri bulma yeteneğini



göstererek anlama koşulunun ikincisini de karşıladığı iddia edilebilir.

Bu çalışma, YSA'ların zekânın temel bileşen ve gerekliliklerini karşılayıp karşılamadığı sorusuna bir cevap aramıştır. Öncelikle insan-merkezci değil, evrensel bir zekâ tanımına ihtiyaç duyulduğu savunulmuştur. Bu evrensel zekâ bakış açısı makinelerin insanlar gibi genel zekâyâ sahip olamayacağı savlarının bir kısmını çürütmektedir. Ayrıca zayıf yapay genel zekâ sorusunun, tıpkı güçlü yapay genel zekâ sorusu gibi felsefe ve yapay zekâ mühendisliği açısından dikkati hak ettiği ifade edilmiştir. Duyguların ve bilincin kavramsal olarak zekânın zorunlu şartları olmadığı ve alandaki gelişmeleri değerlendirmek için de zekânın zorunlu şartı sayılmamaları gerektiği savunulmuştur. Son olarak, anlamanın bazı şartlarının YSA'lar tarafından gösterildiği ortaya konularak, duyguları ve bilinci dışlayan bir zayıf genel zekâ sistemine günümüz YSA'larıyla ulaşılmaması için bir sebep olmadığı temellendirilmeye çalışılmıştır.

Bu akıl yürütmeler, YSA'ların herhangi bir devrimsel gelişmenin yokluğunda bile genel zayıf zekâ hedefine ulaşabileceğini göstermeyi amaçlamaktadır. Buradan mevcut YSA algoritmalarının kusursuz ve tastamam olduğu sonucuna varmamak gerekir. Ancak mevcut yöntemlerle, YSA'lar genel zayıf zekâyâ ulaşacaktır. Daha büyük ağlar, daha optimize algoritmalar ya da donanımlar, daha temiz ya da geniş veri setleri ve çoklu-modaliteli eğitim süreçleri, çok geçmeden makinelerin genel zekâ konusunda yeterlilikleri konusunda en kararlı düşünürleri bile hiç değilse bu makinelerin genel zekâyı simüle etmeye başladığı konusunda ikna edecektir.

Kaynaklar

- Allen, N. J. & Eroglu, C. (2017). Cell Biology of Astrocyte-Synapse Interactions. *Neuron*, 96(3), 697-708.
- Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral cortex*, 10(3), 295-307.
- Bender, E. M. & Koller, A. (2020). Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data. İçinde *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, (ss. 5185-5198). Association for Computational Linguistics.
- Binet, A., & Simon, T. (1973). *The Development of Intelligence in Children*. Arno Press.



- Chalmers, D. J. (2010). The Singularity: A Philosophical Analysis. *Journal of Consciousness Studies*, 17(9-10), 7-65.
- Damasio, A. R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. Avon Books.
- Fields, R. D., Woo, D. H., & Basser, P. J. (2015). Glial Regulation of the Neuronal Connectome through Local and Long-Distant Communication. *Neuron*, 86(2), 374-386.
- Friedman, M. (1974). Explanation and Scientific Understanding. *The Journal of Philosophy*, 71(1), 5-19.
- Gettier, E. L. (1963). Is Justified True Belief Knowledge?. *Analysis*, 23(6), 121-123.
- Good, I. J. (1966). Speculations Concerning the First Ultraintelligent Machine. İçinde F. Leopard Alt & M. Rubino (Eds.), *Advances in Computers*. (6. Cilt; ss. 31-88). Academic Press Inc.
- Gottfredson, L. S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history and bibliography. *Intelligence*, 24(1), 13-23.
- Gottfredson, L. S. (1998). The general intelligence factor. *Scientific American Presents*, 9, 24-29.
- Kaku, M. (1999). *Visions: How Science Will Revolutionize the Twenty-first Century*. Oxford University Press.
- Kim, J. (1994). Explanatory Knowledge and Metaphysical Dependence. *Philosophical Issues*, 5, 51-69.
- Kitcher, P. (1981). Explanatory Unification. *Philosophy of Science*, 48(4), 507-531.
- Kurzweil, R. (2005). *The Singularity is Near: When Humans Transcend Biology*. The Viking Press.
- Kvanvig, J. L. (2003). *The Value of Knowledge and the Pursuit of Understanding*. Cambridge University Press.
- Lear, J. (1988). *Aristotle: The Desire to Understand*. Cambridge University Press, 1988.
- Legg, S., & Hutter, M. (2006). A Formal Measure of Machine Intelligence. *ArXiv, abs/cs/0605024*.



- Ma, W., Wu, D., Sun, Y., Wang, T., Liu, S., Zhang, J., Xue, Y. & Liu, Y. (2025). Combining Fine-Tuning and LLM-based Agents for Intuitive Smart Contract Auditing with Justifications. *İçinde 2025 IEEE/ACM 47th International Conference on Software Engineering (ICSE)*, (ss. 1742-1754), IEEE.
- Mitchell, M. (2019). Artificial Intelligence Hits the Barrier of Meaning. *Information*, 10(2), 51.
- Moore, G. E. (1965). Cramming More Components onto Integrated Circuits. *IEEE Solid-State Circuits Society Newsletter*, 11(3), 82-85.
- Nahmias, E., Allen, C. H., & Loveall, B. (2020). When do robots have free will? Exploring the relationships between (attributions of) consciousness and free will. *İçinde B. Feltz, M. Missal, & A. Sims (Eds.), Free will, causality, and neuroscience* (pp. 57-80). Brill Rodopi.
- Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. J., & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, 51(2), 77-101.
- Pepperell R. (2022). Does Machine Understanding Require Consciousness?. *Frontiers in systems neuroscience*, 16, 1-13.
- Price, R. B., & Duman, R. (2020). Neuroplasticity in cognitive and psychological mechanisms of depression: an integrative model. *Molecular psychiatry*, 25(3), 530-543.
- Pritchard, D. (2009). Knowledge, Understanding and Epistemic Value. *Royal Institute of Philosophy Supplement*, 64, 19-43.
- Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417-424.
- Vinge, V. (2003, Ocak). *Technological Singularity*. <http://www8.cs.umu.se/kurser/5DV084/HT10/utdelat/vinge.pdf> (erişildi: Ekim 12, 2019).
- Zagzebski, L. (2019). Toward a Theory of Understanding. *İçinde S. R. Grimm (Ed.), Varieties of Understanding: New Perspectives from Philosophy, Psychology and Theology* (ss. 123-136). Oxford University Press.



Öz: Bugünün büyük ölçekli dil modelleri, yapay sinir ağı tabanlıdır. 2020 yılından bu yana, GPT serisi gibi büyük ölçekli dil modelleri, akademik dünya, iş dünyası ve kamuoyunun dikkatini çekmiş ve ilgisini yapay sinir ağı tabanlı doğal dil işleme teknolojilerine yönlendirmiştir. Makine öğrenmesi alanındaki araştırmacılar, GPT serisinin doğal dil işleme alanındaki başarısı ve Yapay Genel Zekâ için önemi konusunda farklı görüşlere sahip olsalar da GPT serisinin dil işleme yeteneklerinin beklentileri aştığı konusunda bir fikir birliği vardır. Bu çalışma, yapay sinir ağlarının genel zekâyâ ulaşip ulaşamayacağı sorusunu tartışmaktadır. Bu soruyu ele alırken genel zekâyâ ulaşmak için gerekli olan bazı temel bilişsel becerilerin, mevcut halleriyle yapay sinir ağları tarafından simüle edilip edilemediği değerlendirilecektir. Bunun yanıtını vermek için ise yapay sinir ağlarının nasıl istatistiksel süreçlerle çalıştığı ve bu istatistiksel süreçler ile genel zekâ ve anlamanın bazı temel bileşenleri arasındaki ilişki tartışılacaktır.

Anahtar Kelimeler: Yapay Sinir Ağları, Yapay Genel Zekâ, Doğal Dil İşleme, Yapay Anlama, Fonksiyonalizm.

