A
C
D
G

Ole_CNX / iStock

Yapay Zekâya "Merhaba" Dediğinizde Aslında Kiminle Konuşuyorsunuz?

Doç. Dr. Erkan Er [*Orta Doğu Teknik Üniversitesi, Eğitim Fakültesi, Bilgisayar ve Öğretim Teknolojileri Eğitimi Bölümü*

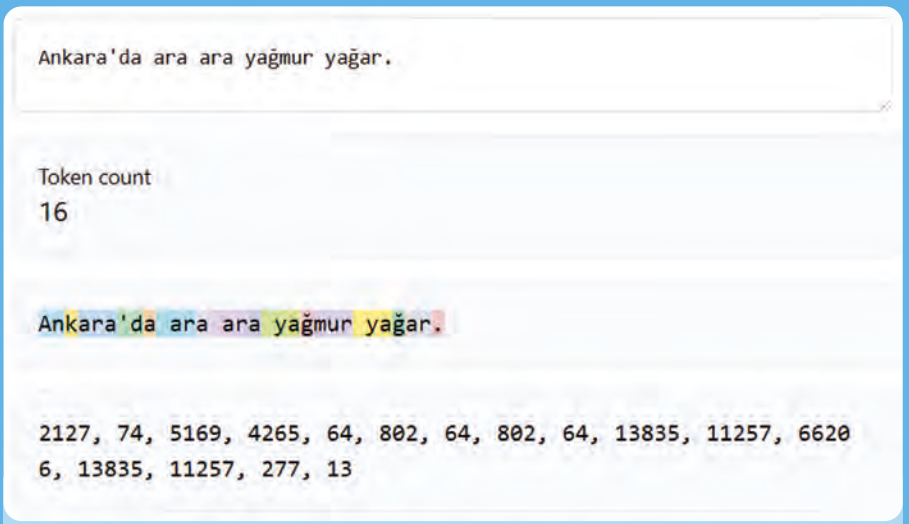
Chatgpt'nin Metin Kutusunun Ardındaki Gerçek

Yapay zekâ sohbet robotlarında bir metin kutusu bulunuyor ve bu kutu içine istediğiniz konuda ne yazarsanız yazın size cevap veriyor. "Bana bir iş başvurusu mektubu yaz" diyorsunuz, size ayrıntılı bir metin hazırlıyor. "Kuantum fiziğini beş yaşındaki bir çocuğa anlat" diyorsunuz, deneyimli bir öğretmen gibi konuyu açıklıyor. Bu örneklerin sonu yok. Peki bu cevaplar nasıl üretiliyor? Gerçekten zeki bir sistem ile mi konuşuyorsunuz yoksa bu çok karmaşık bir illüzyon mu? Cevap, tahmin edebileceğinizden hem daha basit hem daha etkileyici.

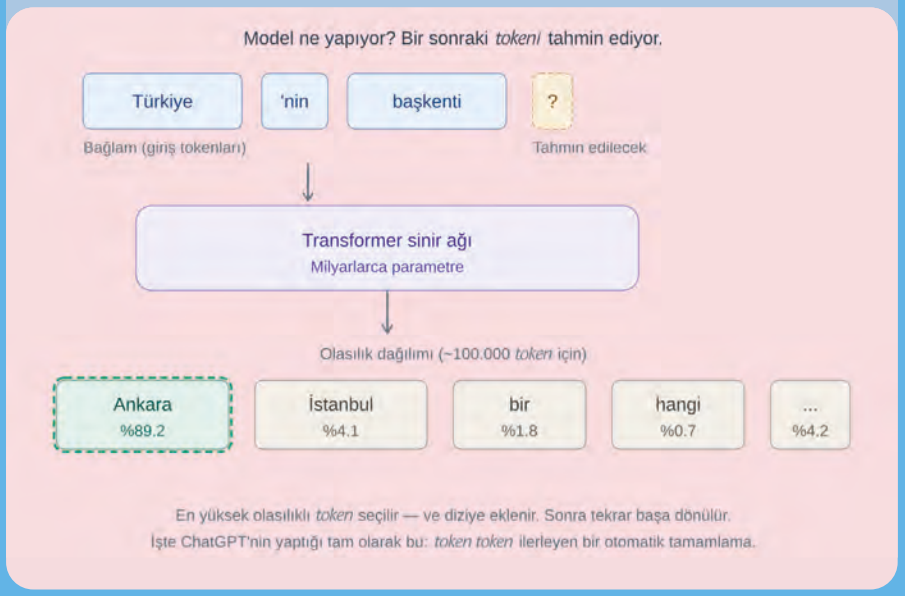
Dünyanın En Pahalı Otomatik Tamamlama Sistemi

ChatGPT ve benzer yapay zekâ araçlarının temelinde yatan mekanizma, telefonunuzdaki bir sonraki kelimeyi tahmin etme özelliğinden kavramsal olarak farksızdır. "Bugün hava çok" yazdığınızda telefonunuzun "güzel" kelimesini önermesi gibi ChatGPT de bir sonraki kelimeyi tahmin ederek yanıt üretir ama internet ölçeğinde, milyarlarca parametre ile.

Süreç şöyle işliyor: Önce internetteki büyük miktardaki verinin bir kısmı indiriliyor. Daha sonra spam siteler, zararlı içerikler, reklam sayfaları eleniyor. Geriye kalan onlarca terabaytlık metinsel veri, "token" denilen birimlere ayrılıyor. Bu token'lar harf ya da hece değil, kelime parçalarıdır ve metnin hangi bölümlere ayrılacağını belirlemede çeşitli algoritmalar kullanılıyor. Bunun pratikte nasıl uygulandığını somut bir örnekle gösterelim. Aşağıdaki görselde, "Ankara'da ara ara yağmur yağar." cümlesinin GPT-4 modelinin tokenizer'ı (metni parçalara ayıran araç) tarafından nasıl bölündüğünü görebilirsiniz. Altı kelimelik sıradan bir cümle gibi görünen bu metin, model tarafından 16 ayrı token olarak algılanıyor. Her renk farklı bir token'ı temsil ediyor: Örneğin "Ankara" kelimesi tek bir token değil, birden fazla parçaya ayrılıyor, "yağmur" kelimesi iki ayrı token'a bölünüyor; kesme işareti



"Ankara'da ara ara yağmur yağar." cümlesinin GPT-4 modelinin tokenizer'ı ile token'lara ayrılması. Her renk ayrı bir token'ı, alttaki sayılar ise her token'ın benzersiz kimlik numarasını gösteriyor.



da-kuk / iStock

ve ekler kendi başlarına birer token oluşturuyor. En altta ise modelin gerçekte aldığı girdiler var: rastgele gibi görünen bir sayı dizisi: 2127, 74, 5169, 4265, 64, 802, 64, 802, 64, 13835, 11257, 6620, 6, 13835, 11257, 277, 13. İşte model; harfleri, kelimeleri ve hatta noktalama işaretlerini sadece bu sayısal kimlikler üzerinden algılıyor.

Sonra Transformer adlı verilen ve milyarlarca parametreden oluşan matematiksel bir fonksiyon, bu token dizileri üzerinden eğitilerek bir beceri kazanıyor: Bir sonraki token'ın ne olacağını tahmin etmek. "Türkiye'nin başkenti" yazdığınızda "Ankara" token dizisinin gelme olasılığının yüksek olduğunu öğreniyor çünkü eğitim verisinde bu kalıp binlerce kez geçiyor.

Sürecin en etkileyici özelliği şu: Model, bu basit tahmin işlemini 15 trilyon token üzerinde tekrarlayarak dünyanın coğrafyasını, tarihin akışını, programlama dillerinin söz dizimini, hatta insan ilişkilerinin dinamiklerini parametrelerinde yani milyarlarca sayısal değerde kodluyor. Bunu internetteki verilerin sıkıştırması gibi düşünebilirsiniz. Yani bir ZIP dosyası gibi. Ancak verileri birebir kopyalamıyor bunun yerine verilerdeki istatistiksel örüntüleri parametrelerine yansıtıyor.

Aylarca süren bu eğitim süreci binlerce özel GPU çipinde milyonlarca dolar harcanarak gerçekleştiriliyor. Modellerin eğitiminde kullanılan çiplerin tek bir görevi var: Bir sonraki *token*'ı tahmin etmeyi öğrenmek. Tahmin sürecinin nasıl işlediği sol üstteki şekilde görselleştirilmiştir.

NicoElNino / iStock

Yapay Zekâ Asistanını İnsana Benzeten Süreç

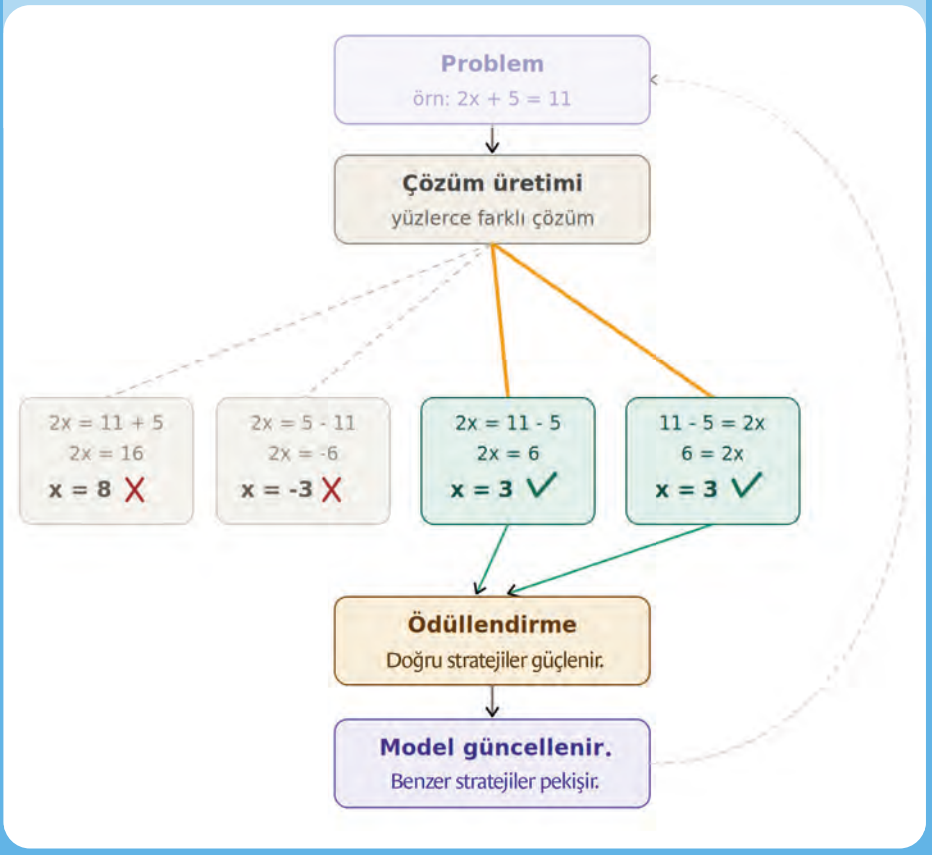
Peki büyük ölçüde internet verileriyle eğitilen bir model nasıl oluyor da sorularınıza kibarca cevap veren dijital bir asistana dönüşüyor? İşte sürecin en az bilinen ama en önemli kısmı burası. Temel model aslında internetteki bilgileri doğrudan yansıtan veya kullanıcıyla doğrudan iletişim kuran bir araç olarak tasarlanmıyor. Modele "İstanbul'da görülecek yerler" diye sorsanız bir forum tartışmasının devamını, bir blog yazısının parçasını ya da tamamen alakasız bir metni üretebilir. Çünkü sadece internetteki verilere göre bu kelimelerden sonra gelebilecek kelimeleri tahmin ederek çalışıyor.

Asistana dönüşüm, denetimli ince ayar adı verilen ikinci aşamada gerçekleşiyor. Bu aşamada OpenAI, Google, Anthropic gibi şirketlerin istihdam ettiği yüzlerce uzman insan, modellerin çıktılarını değerlendiriyor ve etiketliyor. Bu kişilerin işi şu: binlerce farklı soru için "İdeal bir asistan bu soruya nasıl cevap verirdi?" sorusunun cevabını, geliştirici tarafından belirlenen yönergelere uygun şekilde yazmak. Model bu yanıtlar üzerinde eğitildiğinde istatistiksel olarak bu değerlendiricilerin davranışını taklit etmeyi öğreniyor. Yani ChatGPT'ye soru sorduğunuzda, aslında uzman bir insanın o soruya vereceği cevabın yapay sinir ağı tarafından üretilmiş bir karşılığını alıyorsunuz. Aslında ChatGPT yapay bir zekâ değil son derece sofistike bir insan taklidi.

Bu, modelin yetersiz olduğu anlamına gelmiyor. Aksine, çok sayıda uzman kişinin kolektif bilgi birikimine erişme imkânı sağlıyor.

Dolayısıyla "Gerçekten zekâsı olan bir sistemle mi konuşuyorum?" sorusunun doğru cevabı: "Büyük veri setlerinden öğrendiği örüntüleri, insan geri bildirimleriyle şekillendirerek kullanan istatistiksel bir model ile konuşuyorum." olmalı.

Düşüncenin Doğuşu: Pekiştirmeli Öğrenme



Yapay zekâ alanındaki en güncel ve çarpıcı gelişmelerden biri, "akıl yürüten modeller" olarak adlandırılan yeni nesil sistemlerdir. 2025 yılının başında Çinli teknoloji şirketi DeepSeek tarafından kamuoyuyla paylaşılan R1 modeli, bu alanda bir dönüm noktası niteliği taşıyor. R1'in önemi, pekiştirmeli öğrenme yönteminin dil modellerinin performansını nasıl etkilediğini ortaya koymasından ve eğitim parametrelerinin açık bir şekilde paylaşılmasından kaynaklanıyor.

Yöntemin özü şu: Modele bir matematik problemi sunulur ve bu probleme yönelik yüzlerce farklı çözüm üretmesi istenir. Üretilen çözümlerin bir kısmı doğru sonuca ulaşırken bir kısmı ulaşamaz. Doğru sonuca ulaşan çözümler ödüllendirilerek modelin parametreleri, ileride benzer akıl yürütme adımlarını daha yüksek olasılıkla üretecek biçimde güncellenir. Bu döngü, binlerce farklı problem üzerinde yinelendiğinde sistemin genel problem çözme kapasitesi kademeli olarak güçlenir.

Bu sürecin ortaya çıkardığı sonuçlar oldukça dikkat çekicidir. Model, mühendisler tarafından adım adım programlanmadan kendiliğinden düşünme stratejileri geliştirmeye başlar. "Bir adımı yeniden doğrulamak", "Sorunu farklı bir perspektiften ele almak" ya da "Alternatif bir yöntemle teyit etmek" gibi akıl yürütme örüntüleri, eğitim sürecinde kendiliğinden gelişebilir. Yani bu örüntüler modele doğrudan programlanmamıştır, pekiştirmeli öğrenme sürecinde doğru sonuca ulaşma olasılığını artırdıkları için model tarafından optimizasyon sürecinde daha fazla tercih edilir. Bu mekanizma yukarıdaki şekilde görselleştirilmiştir.

Pekiştirmeli öğrenme yaklaşımı sadece matematiksel problemlerde değil, tek bir cevabı olmayan açık uçlu düşünme ve problem çözme süreçlerinde de uygulanır. Bu durum yapay zekânın geleceğine dair köklü bir değişim potansiyeline işaret eder: hiçbir insanın kurmadığı analojiler, hiçbir insanın düşünmediği çözüm stratejileri ve hatta belki de doğal bir dille ifade edilemeyen, modelin kendiliğinden geliştirdiği ve içsel olarak kullandığı özgün bir düşünce yapısı.

da-kuk / iStock

Halüsinasyonlar: Model Neden Yalan Söylüyor?

ChatGPT'ye "Serhan Kumralı kimdir?" diye sorduğunuzu düşünün (bu isim tamamen uydurmadır). Eski büyük dil modelleri size kesin bir dille "Serhan Kumralı, 1970'lerde TRT'de program yapan ödüllü bir belgesel yönetmenidir." şeklinde cevap verebilirdi. Tekrar sorsanız "Galatasaray altyapısından yetişmiş eski bir milli futbolcu." derdi. Bir daha sorsanız cevabı "Ankara Üniversitesinde görev yapmış bir hukuk profesörü." olabilirdi.

Neden? Çünkü eğitim verisinde "X kimdir?" formatındaki sorulardan sonra çoğunlukla biyografik açıklamalar gelir, bu kalıbı öğrenen modeller gerçekte var olmayan isimler için de benzer metinler üretebilir. Çünkü eski modeller, belirsizlik durumlarında cevap üretmek yerine yanıt vermekten kaçınacak şekilde eğitilmemiştir.

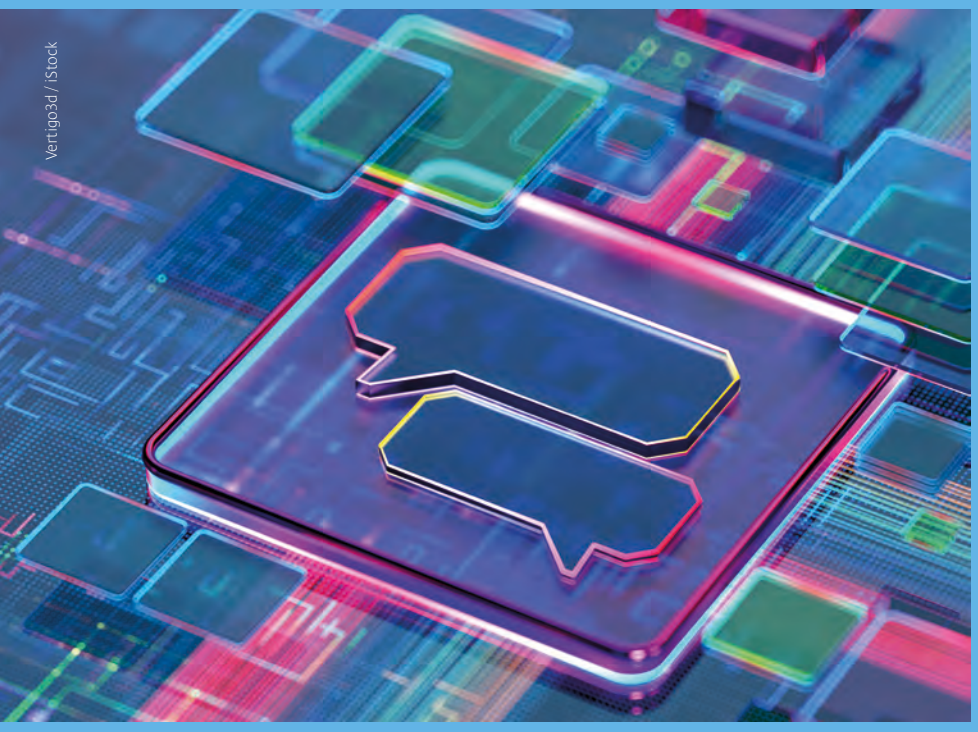
Modern modeller bu sorunu kısmen çözdü. Nasıl mı? Yapay zekâ modelini tabiri caizse "sorguya çekerek." Modele sorulan soruların cevapları kontrol ediliyor. Model aynı soruya birbirinden farklı ve yanlış cevaplar verirse denetimli ince ayar sürecinde eğitim verisine "Bu sorunun doğru cevabı: Bilmiyorum" şeklinde örnekler ekleniyor. Böylece modelde zaten var olan ve modelin belirsiz tahminler yaptığını gösteren nicel özellikler, metinsel çıktı şeklinde ifade edilmiş oluyor.

İkinci çözüm ise araç kullanımı: Model cevabından emin olmadığında özel token'lar üreterek bir web araması başlatabiliyor. Arama sonuçları "bağlam penceresi" olarak adlandırılan modelin çalışma belleğine ekleniyor. Bu sayede eğitim verilerinin güncelliğini yitirdiği durumlarda model doğrudan erişilebilen güncel bilgiyle cevap veriyor.

Etkileyiciliğin Ötesine Geçmek

Büyük dil modellerinin tüm bu etkileyici yeteneklerine rağmen öngörülemeyen açıkları var. Genel olarak çok başarılı performanslar gösterebiliyor ancak çok basit sorularda tutarsız veya hatalı cevaplar verebiliyorlar. Örneğin matematik olimpiyatı soruları çözebilen bir model "9,11 mi büyük yoksa 9,9 mu?" şeklindeki basit bir matematik problemine "9,11" diye cevap verebiliyor. Ya da doktora seviyesinde fizik problemlerini adım adım çözerken bir cümledeki harf sayısını yanlış sayabiliyor. Bu tür tutarsızlıklar modellerin sınırlılıklarını gösteriyor.

Sonuç olarak bu yapay zekâ araçlarını yanılmaz bilgi kaynakları olarak değil, yardımcı araçlar olarak değerlendirmek gerekiyor. Fikir üretme, taslak oluşturma ve alternatif bakış açıları keşfetme gibi alanlarda olağanüstü faydalar sağlıyorlar. Ancak kritik konularda çıktıların mutlaka doğrulanması, matematiksel hesaplama gerektiren görevlerde kod çalıştırarak kesin sonucu hesaplayan kod yorumlayıcısının kullanılması ve güncel bilgi içeren konularda web aramasıyla cevabın kontrol edilmesi büyük önem taşıyor. Ve belki de en temel ilke şu: Bir yapay zekâ uygulamasının ürettiği çıktının doğruluğunu ve niteliğini değerlendirme sorumluluğu, her zaman onu kullanan kişiye aittir.



Vertigo3d / iStock

Metin kutusunun arkasında ne etkileyici bir zekâ var ne de karmaşık bir illüzyon. Bu sistemler, büyük veri setlerindeki istatistiksel örüntülerin milyarlarca parametreye yansıtılması sayesinde çalışıyor.

Son yıllarda yapay zekâ alanında çok hızlı gelişmeler yaşanıyor. Bugün etkileyici olarak nitelenen yeteneklerin birkaç yıl sonra geriye dönüp bakıldığında sadece bir başlangıç noktası olarak hatırlanması kuvvetle muhtemel.

Şunu asla unutmamalıyız: Yapay zekâ ne kadar güçlenirse güçlensin onu anlamlı kılan şey hâlâ onu yönlendiren, sorgulayan ve sonuçlarının sorumluluğunu üstlenen insan zihnidir.

Bilim ve Teknik podcast yayını kapsamında Doç. Dr. Erkan Er ile gerçekleştirdiğimiz "Eğitimde Yapay Zekânın Kullanımı: ChatGPT ve Diğer Sohbet Robotları Eğitimde Nasıl Faydalı Şekilde Kullanılabilir?" konulu programı dinlemek için kare kodu akıllı cihazınıza okutabilirsiniz. ■

Kaynak

https://www.youtube.com/watch?v=7xTGNNLPyMI

Not:
Bu yazı, Andrej Karpathy'nin büyük dil modelleri üzerine kapsamlı dersinden esinlenerek hazırlanmıştır.