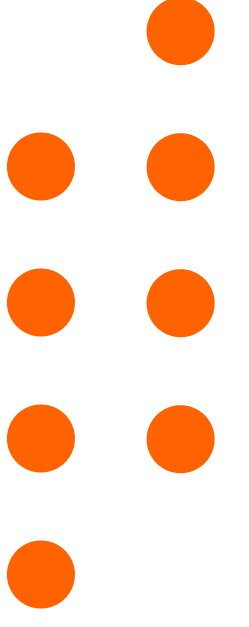
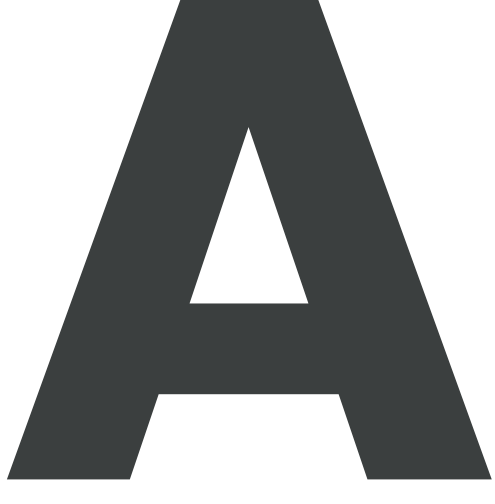


YAPAY ZEKA SÖZLÜĞÜ

AGI'den zero-shot'a kadar
yapay zekâ terimlerine hızlı
başvuru rehberiniz.



AGI (Artificial General Intelligence)

Yapay Genel Zekâ (AGI)

İnsanların üstesinden gelebildiği hemen her entelektüel görevi anlayabilen, öğrenebilen ve uyum sağlayabilen varsayımsal bir yapay zekâ türüdür. Sadece dar ve önceden tanımlanmış problemlerle sınırlı değildir. Güçlü transfer öğrenmesi, kendi kendine yön belirleme ve sağduyuya dayalı akıl yürütme yeteneklerini içerir. Günümüzde bu seviyeye net biçimde ulaşan bir sistem olmadığından, AGI tartışmaları teknik uygulanabilirlik kadar yönetim ve güvenlik konularına da odaklanır.

AI (Artificial Intelligence)

Yapay Zekâ (AI)

Verilerden öğrenerek, insanlar tarafından zeka gerektirdiği düşünülen görevleri yerine getirebilen her türlü sistemdir. Bu görevler arasında örüntü tanıma, dil anlama ve tahmin yapma gibi işler yer alır. “Yapay zekâ” kavramı, basit sınıflayıcılardan çok modlu büyük modeller ve dış araçları çağırabilen ajanlara kadar geniş bir yelpazeyi kapsar.

AI Agent

Yapay Zekâ Ajanı

Çevresini algılayan (girdiler veya araçlar yoluyla), bir hedef doğrultusunda eylemler seçen ve bu eylemleri uygulayan yazılımdır. Genellikle planla → uygula → gözlemle döngüsü içinde çalışır. Basit bir API çağırarak sohbet robotundan, görevleri planlayıp dosyalarla bağımsız şekilde çalışan sistemlere kadar uzanabilir.

● **AI Assistant**

● **Yapay Zekâ Asistanı**

● Kullanıcılara görevlerini yerine getirmede yardımcı olan, metin veya ses tabanlı konuşma arayüzüdür. Soru yanıtlamak, içerik taslağı hazırlamak, uygulama çalıştırmak veya görevleri otomatikleştirmek gibi işlevler üstlenir. Modern asistanlar bağlamı hatırlayabilir, tonunu uyarlayabilir ve arama, e-posta ya da tablo gibi araçları doğrudan kullanarak yalnızca metin değil, sonuç da sunabilir.

AI Browser

Yapay Zekâ Tarayıcısı

Sayfa özetleme, satır içi soru-cevap ve form doldurma gibi yapay zekâ özelliklerini içeren tarayıcı ya da tarayıcı katmanıdır. Pazarlama odaklı bir terimdir; teknik olarak çoğu, standart bir tarayıcıya LLM destekli bir kenar çubuğu veya uzantı eklenerek oluşturulmuştur.

AI Ethics

Yapay Zekâ Etiği

Yapay zekâ sistemlerini adil, şeffaf ve güvenli biçimde geliştirme ve kullanma alanıdır. Önyargı azaltma, gizlilik, açıklanabilirlik, hesap verebilirlik, çevresel etkiler ve uzun vadeli toplumsal sonuçları kapsar.

AI Phone

Yapay Zekâ Telefonu

Cihaz üzerinde çalışan ya da sıkı entegrasyonla gelen yapay zekâ özellikleri etrafında pazarlanan akıllı telefonlardır. Yerel ses yazımı, özetleme, fotoğraf ve video düzenleme gibi işlevlere sahip olabilirler; ancak çoğu hâlâ ağır görevler için bulut tabanlı çıkarıma başvurur.

AI Safety

Yapay Zekâ Güvenliği

Gelişmiş yapay zekâ sistemlerinin insan değerleriyle uyumlu kalmasını, kötüye kullanıma karşı dayanıklı olmasını ve kontrol edilebilirliğini sağlamak amacıyla yapılan araştırma ve mühendislik çalışmalarıdır.

AI Search

Yapay Zekâ Destekli Arama

Doğal dilde soruları analiz edip bağlantı listesi yerine sentezlenmiş yanıtlar (genellikle kaynaklarla birlikte) sunan arama motorlarıdır. Bilgi getirme, LLM üretimi ve bazen canlı web taraması gibi öğeleri birleştirir.

Algorithm

Algoritma

Bilgisayarın bir problemi çözmek için izlediği açık, adım adım bir tarifiedir. Makine öğreniminde algoritma, modelin verilerden nasıl öğreneceğini belirler.

Alignment

Hizalama (Uyumlandırma)

Yapay zekâ sisteminin gerçekten insan tarafından istenen hedefleri izlemesini ve yetenekleri geliştikçe kısıtlamalara uymasını sağlamaktır. Teknik yaklaşımları (örneğin RLHF, anayasal istemler) ve sistemleri yönlendirilebilir ve güvenli tutmak için yönetim politikalarını kapsar.

Anthropic

Anthropic

OpenAI'dan ayrılan araştırmacılar tarafından kurulan bir yapay zekâ laboratuvarıdır. Claude model ailesi ve "Anayasal Yapay Zekâ" yaklaşımıyla bilinir—bu yaklaşım, eğitim sürecinde davranışı biçimlendirmek için yazılı ilkeler kullanır.

API

API (Uygulama Programlama Arayüzü)

Bir programın başka bir programla iletişim kurmasına olanak tanıyan kural setidir. Yapay zekâ API'leri, geliştiricilerin kendi modellerini eğitmeden görüntü tanıma veya metin üretimi gibi yetenekler eklemesini sağlar.

ASR (Automatic Speech Recognition)

Otomatik Konuşma Tanıma (ASR)

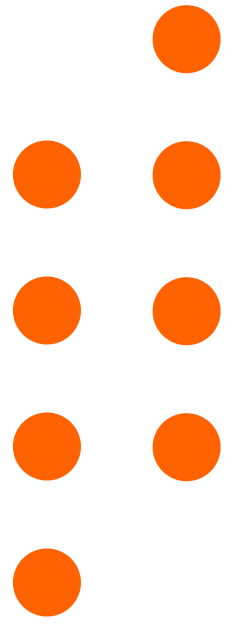
Konuşulan sesi metne dönüştüren yazılımdır; akustik sinyalleri ses birimlerine (fonemlere) ve kelimelere eşleyip, ardından bir dil modeliyle çözümler. Modern ASR sistemleri (genellikle dönüştürücü tabanlı) aksan, gürültü ve kod değişimini işleyebilir; sesli asistanlar ve çağrı analizlerinde kullanılır. Bazen “speech-to-text” (sesten metne) olarak da adlandırılır.

Attention Mechanism

Dikkat Mekanizması

Modellerin her adımda hangi giriş öğelerinin (tokenların) daha önemli olduğunu belirlemesini sağlayan sinirsel bir bileşendir. Seçici olarak “dikkat verme”ye benzer. Dönüştürücülerin (Transformers) uzun dizileri verimli biçimde işlemesini mümkün kılan temel unsurdur.

B



Benchmark

Karşılaştırma Ölçütü

Modelleri karşılaştırmak için kullanılan standartlaştırılmış veri kümesi ve değerlendirme ölçütüdür (örneğin, akıl yürütme için MMLU, görsel algı için ImageNet). “En son teknoloji” (state-of-the-art) iddiaları genellikle belirli benchmarklara dayandığından, bağlam ve tarih çok önemlidir.

Bias

Önyargı

Model çıktılarında sistematik ve adil olmayan sapmadır—genellikle dengesiz veya önyargılı eğitim verilerinden miras kalır. Önyargının tespiti, ölçülmesi ve azaltılması; etik ve güvenilir yapay zekâ uygulamaları için kritik öneme sahiptir.

Big Data

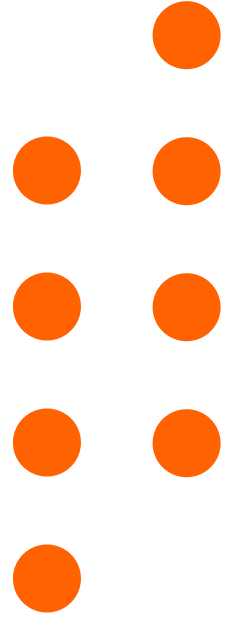
Büyük Veri

Geleneksel işleme yöntemlerini aşan boyutta, hızda veya çeşitlilikteki veri kümeleridir. Makine öğrenimi büyük veriden beslenir, çünkü ölçek sayesinde ince kalıplar ortaya çıkar—ancak veriye gömülü önyargılar da aynı oranda büyütülür.

Bounding Box

Sınırlayıcı Kutu

Bir görüntüde tespit edilen nesneyi yerelleştirmek için görsel modelin çizdiği dikdörtgendir. Nesne tespiti iş akışlarının ve takip gibi sonraki görevlerin temel bileşenidir.



Chain-of-thought (CoT)

Düşünce Zinciri

Modelin, nihai cevabı vermeden önce ara mantık adımlarını açıkça yazması için yönlendirilmesi ya da bu şekilde eğitilmesidir (örneğin, “adım adım düşünelim”). Bu yöntem çok adımlı problemler üzerinde doğruluğu artırabilir. Ancak bu düşünce izleri her zaman modelin gerçek iç mantığını yansıtmaz ve hassas bilgi sızdırabilir; bu nedenle üretim ortamlarında genellikle gizli CoT kullanılır ya da bu faydalar daha kısa ve güvenli gerekçelere dönüştürülür.

ChatGPT

ChatGPT

OpenAI tarafından geliştirilen, GPT model ailesi üzerine inşa edilmiş bir sohbet arayüzüdür. Soruları yanıtlayabilir, metin ve kod yazabilir ya da düzenleyebilir; ayrıca web tarayıcısı veya veri API’leri gibi araçları çalıştırabilir. Yeni sürümleri görüntü, ses ve çok uzun bağlamları işleyebilir—bu sayede sadece bir sohbet robotu değil, çok amaçlı bir dijital asistan işlevi görür.

Claude

Claude

Anthropic tarafından geliştirilen büyük dil modeli ve asistanıdır. “Yardımsever, zararsız, dürüst” ilkeleri üzerine tasarlanmıştır. Claude uzun belgeleri, görselleri ve karmaşık akıl yürütmeyi yönetebilir; ayrıca çıktıları istenen sınırlar içinde tutmak için eğitim sürecine gömülü anayasal kuralları takip edebilir.

Classification

Sınıflandırma

Etiketli örneklerden öğrenerek yeni girdileri önceden tanımlanmış kategorilerden birine atayan denetimli öğrenme (supervised learning) görevidir (örneğin, “spam” mı “spam değil” mi). LLM döneminde, genellikle ayrı bir sınıflayıcı eğitmek yerine, genel amaçlı bir modele doğrudan etiket seçtirmek için istem (prompt) kullanılır; bu da esneklik ve hız kazandırırken biraz kontrol kaybına yol açar.

Cloud Computing

Bulut Bilişim

Kendi donanımınızı satın alıp işletmek yerine, uzak veri merkezlerinden (örneğin AWS, GCP, Azure) bilgi işlem gücü, depolama ve ağ kaynaklarını kiralamaktır. Büyük model eğitimi ve yüksek erişilebilirlikli çıkarım (inference) işlemleri çoğunlukla bulutta yapılır; çünkü esnektir, küreseldir ve yönetilen makine öğrenimi operasyonlarıyla entegre çalışır.

Clustering

Kümeleme

Verileri etiket olmadan gruplandıran denetimsiz bir tekniktir. Aynı kümedeki öğeler birbirine, diğer kümelerdeki öğelere göre daha çok benzer. Belgeleri düzenlemek, müşteri segmentleri keşfetmek veya gömülü vektörler üzerinden anormallikleri saptamak gibi işlemler için kullanışlıdır.

Computer Vision

Bilgisayar Görüşü

Makinelerin görsel veriyi yorumlamasını sağlayan model ve yöntemler alanıdır—nesne tespiti, sahne segmentasyonu, görsel altyazılama veya videoda hareket takibi gibi görevleri kapsar. Modern görsel sistemler genellikle CNN (Evrişimli Sinir Ağları) ya da Görsel Dönüştürücüler (Vision Transformers) kullanır ve giderek daha fazla dil ile birlikte çok modlu (multimodal) görevlerde yer alır.

Confidence Score

Güven Skoru

Bir modelin çıktısından ne kadar emin olduğunu sayısal olarak belirten değerdir. Bu skorlar bazen hatalı kalibre edilebilir (örneğin: yüksek güven ama yanlış sonuç), bu nedenle üretim ortamlarında genellikle kalibrasyon yöntemleri veya ikincil kontroller uygulanır.

Context Window

Bağlam Penceresi

Bir büyük dil modelinin aynı anda “görebileceği” maksimum token (sözcük parçası) sayısıdır. Daha büyük pencereler uzun konuşmaları veya tüm PDF'leri modele sığdırmanıza olanak tanır; ancak bu daha fazla bellek tüketimi ve yavaş çıkarım anlamına gelir. Bu nedenle, bağlamı verimli yönetmek için getirme (retrieval) ve özetleme teknikleri kullanılır.

Conversational AI

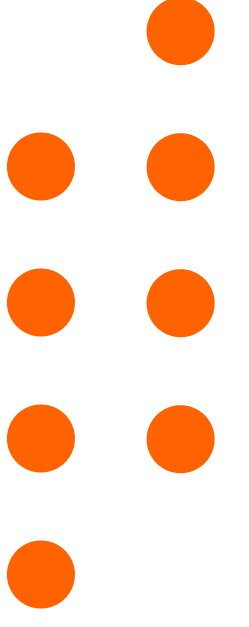
Sohbet Odaklı Yapay Zekâ

Metin veya ses üzerinden çok adımlı diyaloglar kurabilen sistemlerdir; bağlamı takip eder, kullanıcı niyetini netleştirir ve çoğu zaman harici araçları çalıştırarak görevleri tamamlar. Basit SSS botlarından toplantı planlayan veya kod çalıştıran tam teşekküllü ajan sistemlerine kadar uzanır.

CPU (Central Processing Unit)

İşlemci (CPU – Merkezi İşlem Birimi)

Bir bilgisayardaki genel amaçlı işlemcidir. CPU'lar pek çok farklı görevi iyi yönetebilir; ancak derin öğrenmede kullanılan yoğun matris işlemleri genellikle GPU'lar veya özel hızlandırıcılarda daha hızlıdır. Yine de CPU'lar, bu iş yüklerini yönlendirme, ön işleme ve giriş/çıkış işlemlerinde kritik rol oynar.



Data Center

Veri Merkezi

Sunucular, ağ donanımları ve soğutma sistemleriyle dolu, eğitim işlemleri, vektör veri tabanları ve yapay zekâ API'lerinin fiziksel olarak çalıştığı tesislerdir. Hiperskaler (dev ölçekli) sağlayıcılar bu merkezleri yüksek güç tüketimi ve yedeklilik için tasarlar. Bir modelin maliyeti ve gecikmesi, iş yüklerinin nerede ve ne kadar verimli çalıştığıyla yakından ilişkilidir.

Data Mining

Veri Madenciliği

İstatistik ve makine öğrenimi kullanılarak büyük veri kümeleri içinde desenlerin ve ilişkilerin keşfedilmesidir. Günümüz derin öğrenme patlamasından önce de yaygındı ve hâlâ müşteri kaybı tahmini, sepet analizi ve anomali tespiti gibi görevlerin temelini oluşturur.

Data Science

Veri Bilimi

Ham veriyi işe yarar kararlar hâline dönüştürme sürecinin uçtan uca yürütülmesidir: veri toplama, temizleme, keşif, modelleme, değerlendirme ve sonuçları paylaşma. Yapay zekâ ekiplerinde veri bilimciler genellikle model prototipi oluşturur, performansı doğrular ve bulguları iş değerine dönüştürür.

Decision Trees

Karar Ağaçları

Veriyi özellik eşiklerine göre yinelemeli olarak bölerek karar veren denetimli öğrenme modelidir. “Eğer → o hâlde” kurallarından oluşan bir ağaç yapısı oluşturur; her dal kararı, yaprak düğümler ise sınıf etiketini veya sayısal tahmini temsil eder. Yorumu kolaydır—kökten yaprağa yolu izlemek yeterlidir—ancak budanmadıkça ya da topluluk yöntemleriyle (örneğin Random Forest, GBT) birleştirilmedikçe aşırı öğrenmeye (overfitting) yatkındır.

Deep Learning

Derin Öğrenme

Makine öğreniminin bir alt dalıdır; elle tanımlanmış kurallar yerine, çok katmanlı sinir ağları ile ham veriden (metin, piksel, ses) doğrudan özellikleri öğrenir. Derinlik ve ölçek bu modellerin çok karmaşık desenleri yakalamasını sağlar; ancak yüksek hesaplama gücü ve dikkatli optimizasyon gerektirir.

DeepSeek

DeepSeek

Kodlama, akıl yürütme ve genel sohbet alanlarında açık kaynak büyük dil modelleri yayınlayan Çin merkezli bir araştırma grubudur. İlk modeli “DeepSeek-Coder”ın ötesine geçerek, akıl yürütme odaklı ve matematik ağırlıklı modellerle Batılı açık ağırlıklı modellere rakip hâle gelmiştir.

Diffusion Model

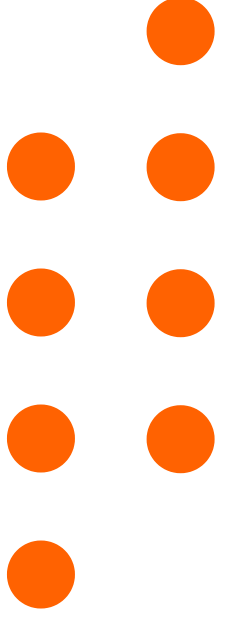
Difüzyon Modeli

Rastgele gürültüyle başlayıp, bunu adım adım gidererek bir görsel, ses dosyası ya da video üreten üretici bir yaklaşımdır. Stable Diffusion, DALL·E 3 ve birçok yeni nesil video sistemi bu tekniği kullanır çünkü yüksek doğruluk ve hassas kontrol sunar.

Distillation

Damıtma

Daha küçük bir “öğrenci” modelin, daha büyük bir “öğretmen” modelin davranışlarını taklit ederek eğitilmesidir –genellikle çıktı olasılıkları (logit), yumuşak etiketler veya düşünce zinciri izleri kullanılır. Böylece daha az parametreyle, daha ucuz çıkarımla benzer performans elde edilir. Varyasyonları arasında çevrimdışı damıtma (önceden eğitilmiş öğretmen), çevrim içi damıtma (öğretmen ve öğrenci birlikte öğrenir) ve öz-damıtma (model kendini eğitir) yer alır. LLM'leri sıkıştırmak, yayına alma süresini hızlandırmak ve “akıl yürütme”yi daha hafif modellere aktarmak için artık standart bir yöntemdir.



Embedding

Gömülü Temsil

Bir nesneyi—metin, görsel, ses ya da kullanıcı profili—temsil etmek için model tarafından öğrenilen yoğun (dense) sayısal vektördür. Bu sayede anlamsal benzerlik, geometrik yakınlık şeklinde modellenir. Gömülü temsiller; anlamsal arama, öneri sistemleri, kümeleme ve bilgi getirme destekli üretim (RAG) gibi birçok sistemin temelini oluşturur.

ELIZA

ELIZA

1964–1966 yıllarında MIT’de geliştirilen bir programdır. Örüntü eşleme tabanlı betikler kullanarak Rogerian tarzı bir terapisti taklit eder ve ilk sohbet botlarından biri olarak kabul edilir. Aynı zamanda kullanıcıların basit programlara insani özellikler atfetme eğilimine atıfta bulunan “ELIZA etkisi” teriminin doğmasına ilham vermiştir.

Evaluation / Evals

Değerlendirme

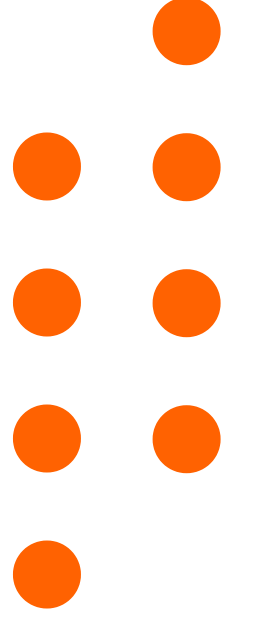
Bir modelin neleri yapabildiğini (ve yapamadığını) test etme sürecidir. Bu süreç; benchmark testleri, karşıt istemler (adversarial prompts), insan tercih çalışmaları veya alana özgü puan kartları kullanılarak yürütülür. İyi bir değerlendirme; gerilemeleri tespit eder, önyargı ve halüsinasyonları ortaya çıkarır ve bir modelin dağıtımına “yeterince güvenli” olup olmadığını anlamaya yardımcı olur.



Explainable AI

Açıklanabilir Yapay Zekâ

Bir modelin davranışlarını insanlar için anlaşılır hâle getiren yöntemlerdir—özellik katkıları, önem haritaları (saliency maps), karşı olgusal senaryolar, basitleştirilmiş vekil modeller veya yapılandırılmış gerekçeler gibi teknikleri içerir. Büyük dil modellerinde açıklamalar genellikle sonradan üretilen metin biçimindedir ve dikkatle ele alınmalıdır; çünkü akıcı anlatım her zaman doğru mantık yürütme anlamına gelmez.



Feature

Özellik

Modelin tahmin yaparken kullandığı ölçülebilir bir özellik ya da işarettir—bir piksel değeri, zaman damgası ya da bir gömü vektöründeki bir boyut olabilir. Derin öğrenmede birçok “özellik”, gizli katmanlar içinde otomatik olarak öğrenilir; manuel olarak tasarlanmaz. Kaliteli özellikler, göreve özgü bilgiyi yoğunlaştırırken gürültüyü eler.

Feature Extraction

Özellik Çıkarımı

Ham veriden, modelin verimli şekilde kullanabileceği bilgi dolu özellikler elde etme sürecidir. Klasik makine öğrenimi, çoğunlukla manuel özellik mühendisliğine dayanırdı (örneğin metin için TF-IDF, görseller için SIFT); modern sistemler ise genellikle önceden eğitilmiş modellerin ara temsillerini çıkararak bu süreci otomatikleştirir.

Federated Learning

Dağıtık (Birleşik) Öğrenme

Modelin kullanıcı cihazlarında veya uç sunucularda kalmasını sağlayan, gizliliği koruyan bir eğitim yaklaşımıdır. Her istemci yerel verisiyle modeli eğitir ve sadece model güncellemelerini (ham veriyi değil) merkezi koordinatöre gönderir. Bu yöntem, telefonlar, hastaneler veya bankalar gibi hassas veri içeren sistemlerde, kişisel verileri cihazda tutarak ortak öğrenmeyi ve düzenlemelere uyumu mümkün kılar.

Fine-tuning

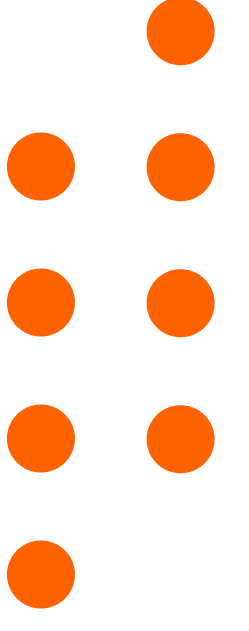
İnce Ayar

Büyük ve önceden eğitilmiş bir modeli, belirli bir alan ya da görev verisiyle kısa ve hedefli bir şekilde yeniden eğitmektir. Bu sayede modelin tonu uyarlanabilir, özel talimatlar takip ettirilebilir ya da hukuk/tıp gibi uzmanlık alanlarına kolayca adapte olması sağlanabilir—sıfırdan eğitim maliyetine katlanmadan.

Foundation Model

Temel Model

Geniş ve çeşitli veri üzerinde (metin, görsel, kod, ses) eğitilmiş, genel yetkinlikler kazanmış çok büyük modellerdir. Geliştiriciler bu modelleri yönlendirme (prompting), ince ayar (fine-tuning) ya da araç kullanımı yoluyla sayısız göreve uyarlayabilir. Bu esneklik, bu modellerin “temel” (foundation) olarak adlandırılmasına neden olur.



Gemini

Gemini

Google'ın, metin, görsel, ses ve videoyu tek bir sistem içinde yerel olarak işleyebilen çok modlu model ailesidir. Hem genel amaçlı bir asistan hem de geliştirici platformu olarak konumlanır; mobil, kurumsal ve ileri düzey akıl yürütme iş yükleri için optimize edilmiş farklı katmanlara sahiptir.

Generative AI

Üretken Yapay Zekâ

Verinin istatistiksel yapısını öğrenerek, bu dağılımdan yeni içerikler (metin, kod, görsel, ses, video) üreten modellerdir. Sohbet botlarından tasarım araçlarına, veri artırmadan simülasyonlara kadar birçok uygulamayı mümkün kılarlar ve genellikle istemler (prompts) veya kontrol sinyalleriyle yönlendirilirler.

GAN (Generative Adversarial Network)

Üretici Çekişmeli Ağ

Bir üretici modelin örnekler oluşturduğu, bir ayırt edici modelin bu örnekleri değerlendirdiği iki modelli yapıdır. Bu çekişmeli oyun sayesinde üretici, giderek daha gerçekçi örnekler (örneğin görsel, ses) üretmeyi öğrenir. Ancak eğitim süreci dengesiz olabilir ve "mod çökmesi" (mode collapse) sık görülen bir sorundur.

● GPU (Graphics Processing Unit)

● Grafik İşlem Birimi

● Başlangıçta grafik işleme için geliştirilmiş, ancak derin öğrenmedeki matris işlemleri için son derece uygun, çok paralel çalışan işlemcidir. GPU'lar hem model eğitimi hem de yüksek verimli çıkarım (inference) süreçlerinde baskındır; CPU'lar ise genellikle yönetim, ön işleme ve daha hafif görevleri yürütür.

GPT (Generative Pre-trained Transformer)

Üretici Önceden Eğitilmiş Dönüştürücü

OpenAI tarafından geliştirilen, internet ölçeğindeki metinler üzerinde eğitilmiş büyük dil modeli serisidir. İnce ayar (fine-tuning) veya yönlendirme (prompting) ile farklı görevlerde kullanılabilir. "GPT" artık çoğu zaman güçlü büyük dil modellerini ifade eden kısa bir ad hâline gelmiştir; ailede farklı boyutlarda, yeteneklerde ve çok modlu sürümler yer alır.

Gradient Descent

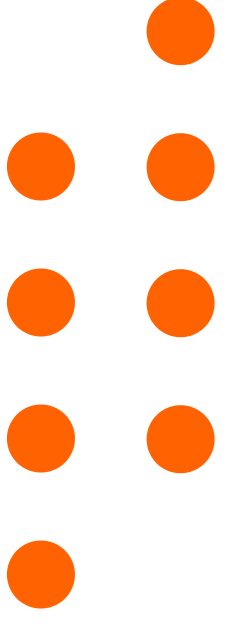
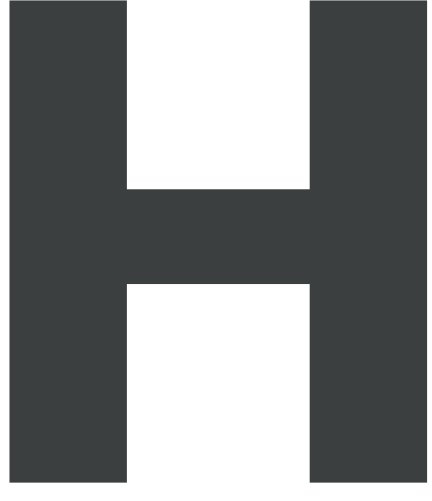
Gradyan İnişi

Derin öğrenmenin temel optimizasyon yöntemidir—modelin ne kadar yanlış olduğunu (kayıp), bu hatanın parametrelere göre gradyanını ve parametreleri nasıl aşağıya doğru güncelleyeceğimizi hesaplar. Adam, RMSProp ve momentum gibi varyasyonlar; büyük modellerde eğitim sürecini hızlandırır ve dengeyi sağlar.

Grok

Grok

xAI tarafından geliştirilen, şirketin kendi büyük dil modelleri üzerine inşa edilmiş sohbet botudur. Esprili ve "sansürsüz" bir kişiliğe sahiptir. Kişiliğinin ötesinde, Grok; X (eski Twitter) platformundaki canlı verileri tarayabilme ve geleneksel kurumsal asistanlardan daha serbest bir üslupla yanıt verme becerisiyle öne çıkar.



Hallucination

Halüsinasyon

Bir modelin, inandırıcı görünen ancak gerçekte hatalı veya tamamen uydurma içerikler üretmesidir. Bu, modelin “bir sonraki olası kelimeyi” tahmin etmesine dayalı çalışma biçiminin doğal bir yan etkisidir; gerçek doğrulama yapılmaz. Bilgi getirme temelli yaklaşımlar (retrieval grounding) ve daha sıkı istemler (prompting) bu sorunu azaltabilir, ancak tamamen ortadan kaldıramaz.

Hugging Face

Hugging Face

Modeller, veri kümeleri ve makine öğrenimi araçları (özellikle transformers kütüphanesi) için açık kaynak bir merkezdir. Aynı zamanda barındırma ve çıkarım (inference) servisleri de sunar. Araştırmacıların ve geliştiricilerin ağırlıkları paylaştığı, değerlendirmeler yaptığı ve birbirlerinin çalışmalarını üzerine inşa ettiği, yapay zekânın adeta “GitHub”ı hâline gelmiştir.

Hunyuan

Hunyuan

Tencent’in metin, görsel ve video alanlarını kapsayan büyük model ailesidir. Uzun bağlam yönetimi ve kurumsal entegrasyon için optimize edilmiştir. Kamuyla paylaşılan teknik detaylar sınırlıdır; bu yüzden Hunyuan, belirli bir mimariden çok Tencent’in kendi iç “temel model” altyapısı olarak düşünülebilir.

Hybrid Architecture

Hibrit Mimari

Transformer katmanlarını durum-uzay (state-space) ya da yinelemeli (recurrent) bileşenlerle harmanlayan model tasarımıdır. Amaç, uzun bağlamları daha verimli işlemek, hesaplama maliyetini azaltmak veya saf Transformer mimarisinde bulunmayan yönlendirici önyargılar (inductive bias) eklemektir.

Hyperparameter

Hiperparametre

Model eğitimi başlamadan önce belirlenen ayarlardır—öğrenme oranı, toplu işlem (batch) boyutu, katman sayısı gibi. Bu ayarlar, modelin nasıl ve ne kadar başarılı şekilde eğitileceğini doğrudan etkiler. Hiperparametrelerin doğru ayarlanması, modelin öğrenip öğrenememesi arasındaki farkı yaratabilir.

Image Generation

Görsel Üretimi

İstemler (prompts), eskizler veya referans görsellerden yeni görseller üretme işlemidir; genellikle difüzyon ya da transformer tabanlı modellerle gerçekleştirilir (DALL·E 3, Midjourney, Stable Diffusion gibi). Güncel sistemler stil kontrolü, görselin içine/dışına çizim yapma (inpainting/outpainting) ve hatta videolarda zamansal tutarlılık sağlama gibi yetenekler sunar.

Inference

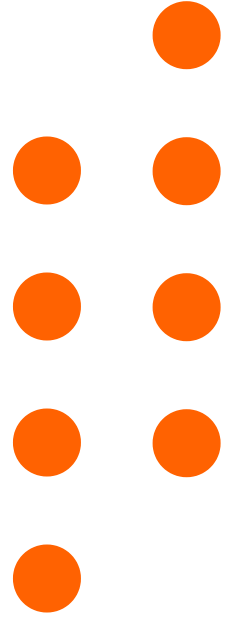
Çıkarım

Eğitilmiş bir modeli ileriye doğru çalıştırarak yeni girdiler için çıktı elde etme işlemidir—modelin parametre ağırlıkları bu süreçte güncellenmez. Büyük dil modelleri (LLM) için çıkarım; soru yanıtlama, uzun metin üretme gibi pek çok işlemi kapsar ve genellikle gecikme (latency) ve maliyet sınırları içinde gerçekleştirilir. Çıkarımı optimize etmek için nicemleme (quantization), toplu işleme (batching), anahtar/değer önbellekleme (caching) ve özel donanım hızlandırıcılar gibi teknikler kullanılır.

Input Token

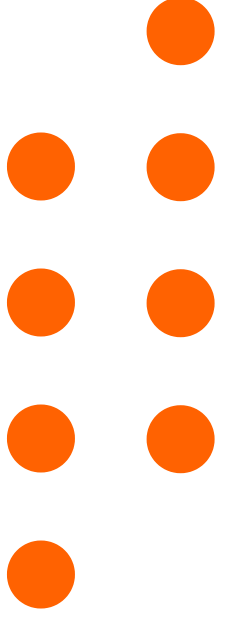
Girdi Token'ı

Bir modelin bağlamı işlerken saydığı en küçük birimdir—genellikle bir alt kelime parçası ya da birkaç karakterden oluşur; bu, kullanılan ayrıştırıcıya (tokenizer) bağlıdır. Faturalama ve bağlam sınırları genellikle karakterle değil, token birimiyle ifade edilir. Girdi token'ları bağlamın parçası olduğu için yalnızca bir kez işlenir ve API sağlayıcıları tarafından genellikle çıktı token'larına kıyasla daha düşük ücretlendirilir.



Knowledge Graph Bilgi Grafiği

Kişiler, yerler, kavramlar gibi varlıkların ve aralarındaki ilişkilerin yapılandırılmış bir ağ biçiminde temsil edildiği sistemdir. Sembolik sorgulara ve olgusal (doğrulanabilir) akıl yürütmeye imkân tanır. Büyük dil modelleri (LLM'ler) ile birlikte kullanıldığında; üretimleri somut gerçeklere dayandırabilir ve yapay zekâ ajanlarının açık ilişkiler üzerinden akıl yürütmesini sağlayabilir.



Latent Space

Gizil Uzay

Modelin eğitim sonrasında kavramları ve ilişkileri temsil ettiği, içsel ve çok boyutlu uzaydır. Bu uzaydaki uzaklıklar ve yönler, anlamsal ilişkilere karşılık gelir (örneğin: “kral” – “adam” + “kadın” \approx “kraliçe” gibi kelime gömülerinde görülen örnekler).

Layer

Katman

Bir sinir ağında aynı “derinlikte” yer alan nöronlar grubudur. Her katman, aldığı girdiyi dönüştürerek bir sonraki katmana iletir. Birden fazla katmanın üst üste yığılması, modelin hiyerarşik özellikleri öğrenmesini sağlar (kenarlar \rightarrow şekiller \rightarrow nesneler, karakterler \rightarrow kelimeler \rightarrow kavramlar gibi).

LLaMA

LLaMA

Meta tarafından geliştirilen açık ağırlıklı büyük dil modeli ailesidir (LLaMA \rightarrow Llama 2 \rightarrow Llama 3/3.1...). Araştırma ve ticari kullanım için farklı boyutlarda (milyarlarca parametreye kadar) yayımlanmıştır. Komutlarla eğitilmiş (instruction-tuned), boyutlarına göre güçlü kod üretme ve akıl yürütme yeteneklerine sahiptir. Topluluk tarafından yapılan ince ayarlar ve türev modeller sayesinde LLaMA, günümüzün açık kaynaklı LLM ekosisteminin fiili omurgası hâline gelmiştir.

● LLM (Large Language Model)

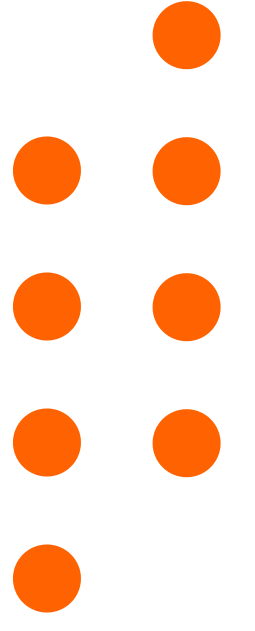
● Büyük Dil Modeli (LLM)

- Milyarlarca (hatta trilyonlarca) parametreye sahip sinir ağıdır. Devasa metin kütüphaneleri üzerinde eğitilerek bir sonraki token'ı tahmin etmeyi öğrenir. Ölçek ve talimatlara uyum sağlama becerisiyle, LLM'ler istemleri takip edebilir, uzun bağlamlar üzerinde akıl yürütebilir, kod yazabilir ve genellikle çok modlu arayüzler olarak görev yapabilirler.

Loss Function

Kayıp Fonksiyonu

Modelin eğitim verisindeki tahmininin ne kadar hatalı olduğunu ölçen matematiksel puanlamadır. Eğitim süreci esasen bu kaybı minimize etmeye (skoru düşürmeye) yöneliktir—parametreler buna göre ayarlanır. Seçilen kayıp türü (örneğin çapraz entropi, ortalama kare hata, karşılaştırmalı kayıp) modelin “başarı” tanımını doğrudan belirler.



Machine Learning

Makine Öğrenimi

Bilgisayarların, sabit kurallar yerine veriden öğrenerek bir görevde kendilerini geliştirmelerine olanak tanıyan yöntemdir. Denetimli (supervised), denetimsiz (unsupervised) ve pekiştirmeli (reinforcement) öğrenme yaklaşımlarını kapsar. Spam filtrelerinden üretici modellere ve öneri motorlarına kadar birçok teknolojinin temelini oluşturur.

Mamba Architecture

Mamba Mimarisi

Klasik Transformer'lara kıyasla çok daha uzun bağlamları verimli biçimde işleyebilen, durum-uzay (state-space) temelli yeni bir dizi modelidir. Karesel zaman karmaşıklığına sahip dikkat (attention) mekanizması yerine, doğrusal zamanlı bir işleme yaklaşımı sunar. Bu da onu uzun belgeler, günlük verileri veya DNA dizileri gibi veri türleri için cazip kılar.

MCP (Model Context Protocol)

Model Bağlam Protokolü

Anthropic ve diğer kuruluşlar tarafından başlatılan, modellerin harici araçları, dosyaları veya veri kaynaklarını nasıl keşfedeceği ve çağıracağına dair açık bir standarttır. Her sağlayıcı için özel kodlama yapmak yerine, MCP; ölçekli “araç kullanımı” için düzenli bir istek/yanıt arayüzü tanımlar.

Metadata

Üst Veri

Veriniz hakkında bilgi içeren verilerdir—zaman damgaları, yazar adları, dosya türleri, etiketler gibi. Yönetim, filtreleme ve daha akıllı geri getirme (retrieval) işlemleri için yardımcı olur. Özellikle RAG (retrieval-augmented generation) sistemlerinde zengin üst veri; sorguları yönlendirmek, erişim izinlerini uygulamak ve sonuçları yalnızca metin benzerliğinden öte kriterlerle sıralamak için kullanılır.

Mistral

Mistral

Verimliliği yüksek, küçük ve orta ölçekli açık ağırlıklı modeller geliştiren Fransız yapay zekâ şirketidir. Açık ağırlık felsefesini benimseyerek, daha büyük laboratuvarlara karşı etkinlik açısından rekabet eden modeller ve ticari API'ler sunar.

Model

Model

Girdileri çıktılara eşleyen öğrenilmiş işlevdir—parametreleri, eğitim verisinden öğrenilen tüm bilgileri içerir. “Model” terimi; basit bir lojistik regresyon algoritmasından trilyonlarca parametrelili çok modlu transformer modellerine kadar her şeyi ifade edebilir.

Model Training

Model Eğitimi

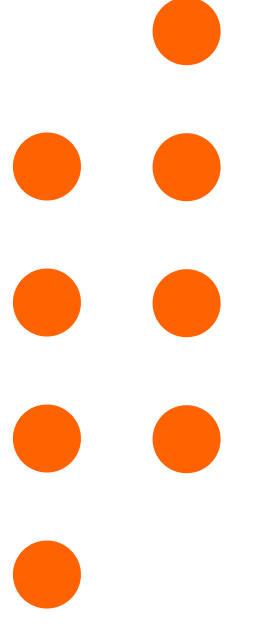
Verinin modelden geçirilmesi, kaybın (loss) hesaplanması ve geri yayılım (backpropagation) ile parametrelerin güncellenmesi yoluyla gerçekleşen, yüksek hesaplama gücü gerektiren süreçtir. Veri seçimi, hiperparametre ayarlamaları, kontrol noktaları (checkpointing) ve aşırı/az öğrenmeyi önlemek için sürekli izleme içerir.

Multimodal AI

Çok Modlu Yapay Zekâ

Birden fazla veri türünü—metin, görsel, ses, video—anlayabilen ve üretebilen modellerdir; genellikle tek bir mimari içinde çalışırlar. Bu sayede dilin görselle bağlantısı kurulabilir (örneğin bir görseli betimlemek ya da grafik hakkında soruları yanıtlamak) veya çıktılar farklı modları birleştirebilir (örneğin anlatımlı video üretimi). Eğitim için uyumlu veri kümeleri ve dikkatli dengeleme gerekir, çünkü bir modun diğerlerini baskılaması istenmez. Karşılığında ise daha zengin, bağlam farkındalığı yüksek akıl yürütme ve etkileşim elde edilir.

N



NLP (Natural Language Processing) Doğal Dil İşleme

Bilgisayarların insan diliyle anlamlı biçimde etkileşim kurmasını, dili anlamasını ve üretmesini sağlamayı amaçlayan yapay zekâ alanıdır. Geleneksel NLP; ayrıştırma, duygu analizi ve makine çevirisi gibi görevlerde elle tasarlanmış özelliklere dayanırdı. Günümüzde ise büyük, önceden eğitilmiş transformer modelleri; isteme (prompting) veya ince ayar (fine-tuning) yoluyla uyarlanarak kullanılır. NLP, sohbet botlarından arama motorlarına, özetleme sistemlerinden metin analizi uygulamalarına kadar sayısız teknolojinin temelidir.

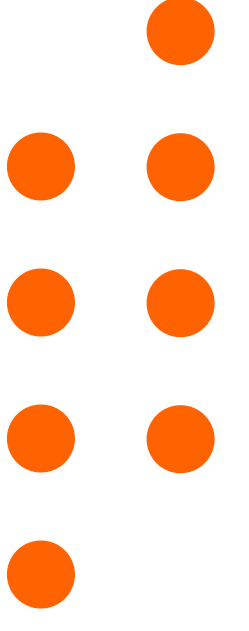
Neural Network Sinir Ağı

Birbirine bağlı çok sayıda “nöron”dan (basit matematiksel birimler) oluşan bir yapıdır; bu bağlantıların ağırlıkları veriyle öğrenilir. İleri yayılım (forward pass) sürecinde girdiler katmanlardan geçerek dönüştürülür, ardından geri yayılım (backpropagation) ile hata azaltılacak şekilde ağırlıklar güncellenir. Katman sayısı (derinlik) ve nöron sayısı (genişlik) arttıkça ağ, son derece karmaşık işlevleri modelleyebilir—ancak bunun karşılığında büyük veri kümeleri ve yüksek işlem gücü gerekir. Biyolojik metafora rağmen, günümüzdeki sinir ağları biyolojik değil, matematiksel olarak tasarlanmış yapılardır.

Nvidia

Nvidia

Modern derin öğrenmeyi mümkün kılan GPU'ların ve ilgili yazılımların (CUDA, cuDNN) en büyük sağlayıcısıdır. A100, H100 gibi donanımları; büyük ölçekli model eğitimi ve yüksek hızlı çıkarım işlemleri için standart hâline gelmiştir. DGX ve Grace sistemleri, uçtan uca yapay zekâ altyapılarına yöneliktir. Nvidia'nın donanım, sürücü ve kütüphanelerden oluşan kapalı ekosistemi, onu yapay zekâ ekonomisinde ve kapasite planlamasında merkezi bir oyuncu hâline getirir.



OCR (Optical Character Recognition) Optik Karakter Tanıma

Metin içeren görüntü veya taramaların, makine tarafından okunabilir karakterlere dönüştürülmesidir. Modern OCR sistemleri önce metin bölgelerini tespit eder, ardından karakterleri tanır—genellikle CNN ya da transformer tabanlı modellerle—ve eğik açı, gürültü veya çok dilli içerikler gibi zorlukları da yönetebilir. OCR, PDF’lerin, faturaların, formların ve tarihî belgelerin dijitalleştirilmesi, aranabilir hâle getirilmesi veya modellerle işlenmesi için kritik öneme sahiptir.

Open-source Açık Kaynak

Kaynak kodu (veya model ağırlıkları) herkesin inceleyebileceği, kullanabileceği, değiştirebileceği ve yeniden dağıtabileceği bir lisansla yayımlanan yazılımdır. Yapay zekâ alanında, açık ağırlıklar topluluk denetimi, hızlı geliştirme ve yerel/özel dağıtım imkânı sağlar. Lisanslar değişkenlik gösterir—özgürlükçü (Apache/MIT) olanlardan kısıtlayıcı (yalnızca ticari olmayan kullanım gibi) olanlara kadar—dolayısıyla “açık” terimi farklı özgürlük seviyelerini ifade edebilir.

OpenAI

OpenAI

GPT, ChatGPT, DALL·E ve Whisper'ın geliştiricisi olan laboratuvardır. Başlangıçta, yapay genel zekânın (AGI) tüm insanlığa fayda sağlamasını güvence altına almayı amaçlayan kâr amacı gütmeyen bir kuruluş olarak kurulmuştur. Günümüzde ise kârı sınırlı bir yapı altında faaliyet göstermekte ve ticari API'ler sunarken aynı zamanda güvenlik ve hizalama (alignment) araştırmalarını da finanse etmektedir. OpenAI, komutla yönlendirme (instruction tuning), insan geri bildirimiyle pekiştirme öğrenimi (RLHF) ve araç kullanabilen asistanların yaygınlaşmasında öncüdür.

Output Token

Çıktı Token'ı

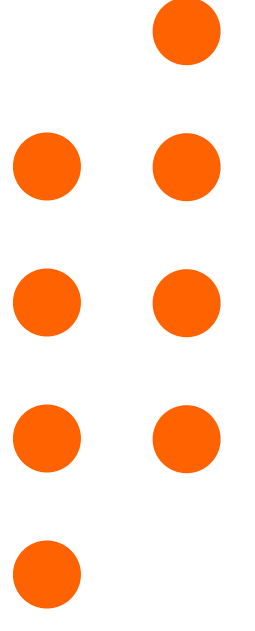
Modelin çıkarım (inference) sırasında ürettiği metin parçacığıdır. Çoğu API, bu token'ları girdilerden ayrı (ve genellikle daha yüksek) birim fiyatla faturalandırır çünkü üretim işlemi daha fazla örnekleme (sampling) mantığı içerir ve toplu işleme (batching) daha az uygundur. Üretilen çıktının uzunluğu; gecikme süresi ve maliyeti etkiler, bu nedenle istem (prompt) tasarımı ve durdurma koşulları operasyonel açıdan kritiktir.

Overfitting

Aşırı Uyum

Modelin eğitim verisinin ayrıntılarına ve rastlantısal gürültüsüne aşırı derecede uyum sağlayarak, yeni veriler üzerinde genelleme yapamaması durumudur. Belirtileri arasında çok yüksek eğitim doğruluğu, ancak zayıf doğrulama/test performansı yer alır. Bu soruna karşı yaygın çözümler; düzenleme (regularization), dropout, erken durdurma (early stopping) ve daha büyük/çeşitli veri kümeleridir.

P



Parameter

Parametre

Bir modelin katmanlarında girdilerin nasıl dönüştürüleceğini belirleyen, öğrenilmiş ağırlık (veya bias) değeridir. Modern büyük dil modelleri (LLM'ler) milyarlarca hatta trilyonlarca parametreye sahiptir ve bu parametreler modelin öğrendiği “bilgiyi” depolar. Parametre sayısı, kapasite ve maliyeti etkiler; ancak mimari, veri kalitesi ve eğitim stratejisi de en az onun kadar önemlidir.

Perplexity

Karmaşıklık

Dil modellemede, bir modelin metni ne kadar iyi tahmin ettiğini ölçen bir metriktir—daha düşük karmaşıklık, modelin diziye daha az “şaşırtıcı” bulduğu anlamına gelir. Ayrıca, “Perplexity” aynı zamanda kaynak göstererek yanıt veren bir yapay zekâ arama motorunun adıdır.

Personalization

Kişiselleştirme

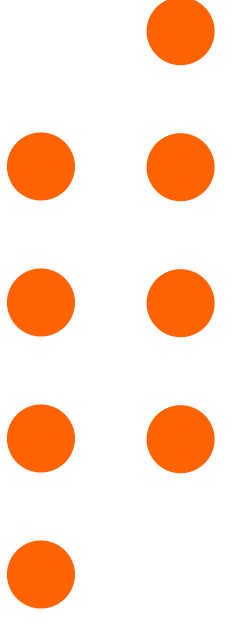
Modelin çıktılarının belirli bir kullanıcıya özel olarak uyarlanmasıdır—tonlama, öneriler veya içerikler; kullanıcının geçmişi, tercihleri veya bağlamına göre ayarlanır. Bu kişiselleştirme; çıkarım (inference) sırasında (örneğin prompt bağlamı yoluyla) ya da kullanıcı verileriyle yapılan ince ayar (fine-tuning) ile gerçekleştirilebilir. Bu süreçte gizlilik ve önyargı riskleri de dikkate alınmalıdır.

Prompt İstem

Bir büyük dil modeline (LLM) davranışını yönlendirmek için verilen talimat ve beraberindeki bağlam, örnekler veya kısıtlamalardır. İyi bir istem; görevi açık şekilde tanımlar, çıktı formatını belirler ve belirsizliği azaltacak kadar sinyal verirken bağlam penceresini aşırı doldurmaz.

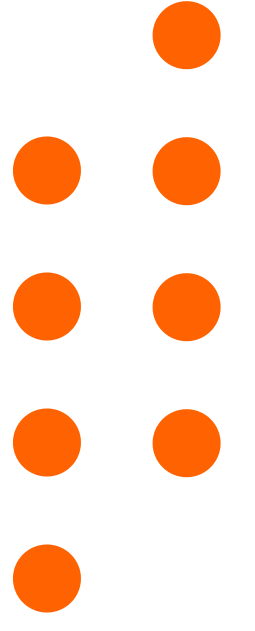
Prompt Engineering İstem Mühendisliği

Bir modelden tutarlı ve yüksek kaliteli sonuçlar almak için istemlerin (ve araç çağırma şemalarının) sistematik biçimde tasarlanması, test edilmesi ve iyileştirilmesidir. Bu süreç; az örnekli (few-shot) örnekleme, düşünce zinciri (chain-of-thought) yapıları, çıktı doğrulayıcıları ve başarısızlıkları yakalayacak güvenlik katmanları gibi teknikleri içerir.



Quantization Kuantizasyon

Bir modelin ağırlıklarını ve aktivasyonlarını daha az bit (örneğin 8, 4 hatta 2 bit) ile temsil ederek sıkıştırılması işlemidir. Bu sayede bellek kullanımı azaltılır ve çıkarım (inference) süreci hızlanır. Doğru uygulandığında doğrulukta minimal kayıp yaşanır. Eğitim sonrası kuantizasyon (post-training quantization) en hızlı uygulanan yöntemdir; kuantizasyona duyarlı eğitim (quantization-aware training) ise modelin doğruluğunu daha iyi koruyabilir.



RAG (Retrieval-Augmented Generation) Geri Getirme Destekli Üretim

Modelin yanıt vermeden önce, gömülemeler (embedding) ya da anahtar kelime araması aracılığıyla ilgili belgeleri getirip isteme (prompt) dahil ettiği bir desen. Bu yaklaşım, çıktıları gerçek verilere dayandırır, halüsinasyonları azaltır ve temel model yeniden eğitilmeden içerik güncellemelerine olanak tanır.

Reasoning Model Akıl Yürütme Modeli

Çok adımlı mantık problemlerini çözebilmesi için (genellikle özel veri kümeleri, düşünce zinciri çıkarımı veya pekiştirmeli öğrenme ile) ayarlanmış büyük dil modellerini tanımlamak için kullanılan pazarlama terimidir. Gerçekte, en iyi modellerin çoğu benzer teknikler kullanır; “akıl yürütme” ifadesi genellikle sembolik ispat değil, daha iyi ara planlama anlamına gelir.

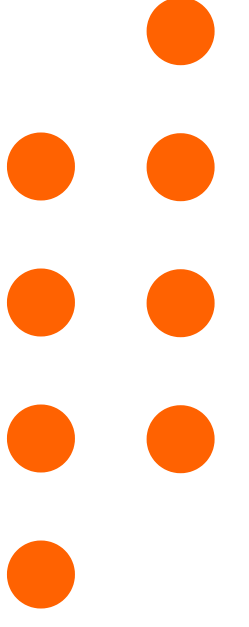
Reinforcement Learning Pekiştirmeli Öğrenme

Bir ajanın bir ortamda eylemler gerçekleştirerek kümülatif ödülü maksimize etmeyi öğrendiği eğitim yöntemidir. Etiketli veriler yerine deneme-yanılma yoluyla öğrenir—keşfetme ve en iyi bilinen stratejiyi uygulama arasında denge kurar. Q-learning veya politika gradyanları gibi algoritmalar kullanılır. Pekiştirmeli öğrenme, robotikte, oyun oynayan yapay zekâlarda ve tercih tabanlı ayarlamalarda (örneğin RLHF) temel rol oynar.

● RLHF (Reinforcement Learning from Human Feedback)

● İnsan Geri Bildirimiyle Pekiştirmeli Öğrenme

İnsanlar model çıktılarının sıralamasını yapar veya puan verir; bu tercihlere göre bir ödül modeli eğitilir ve ardından temel model, insanların beğendiği yanıtları üretmesi için ince ayardan geçirilir. Bu yöntem, el yazımı kurallar olmadan modeli istenen şekilde yönlendirmeyi sağlar.



Semantic Search

Anlamsal Arama

Sadece anahtar kelimeler değil, anlam üzerinden yapılan arama türüdür. Sorgular ve belgeler gömülemelere (vektörlere) dönüştürülür ve semantik yakınlığa göre eşleştirilir. Örneğin “Kasım’da Tokyo’ya ucuz uçuşlar” araması, “bu sonbahar Tokyo için bütçe dostu bilet fırsatları” ile eşleşebilir. Anlamsal arama; modern RAG sistemlerinin, SSS botlarının ve öneri motorlarının temelini oluşturur.

Sentiment Analysis

Duygu Analizi

Metnin duygusal tonunun otomatik olarak sınıflandırılmasıdır—olumlu, olumsuz, nötr veya daha ince duygusal tonlar şeklinde. Klasik yöntemler sözlükler ve sınıflandırıcılar kullanırdı; günümüzde ise büyük dil modelleri (LLM) veya özel olarak ayarlanmış dönüştürücüler (transformers), yorumlar, tweetler, destek talepleri gibi metinlerde duygu analizi yapar. Alaycılık, sektör jargonu ve karışık duygular gibi unsurlar basit modelleri kolayca yanıltabilir.

Speech-to-Text

Konuşmadan Metne

Konuşulan sesin yazılı metne dönüştürülmesidir (otomatik konuşma tanıma). Modern sistemler sesi parçalara ayırır, fonemleri veya tokenları algılar ve ardından sinir ağı tabanlı dil modelleriyle çözümler. Aksanlar, arka plan gürültüsü ve dil geçişleri gibi zorluklarla geçmişe kıyasla çok daha iyi başa çıkabilir. Kaliteli ses tanıma sistemleri; sesli asistanlar, çağrı merkezi analizleri, toplantı transkriptleri ve erişilebilirlik araçları için kritik önemdedir.

SOTA (State-of-the-Art)

En Son Teknoloji / Alanında En İyisi

Belirli bir zaman diliminde bir ölçüt (benchmark) üzerinde raporlanan en yüksek başarı performansdır. Ölçütler değiştikçe ve çeşitli hileler bu ölçütlere aşırı uyum sağladıkça, “SOTA” hareketli bir hedeftir—“hangi görevde, ne zaman ve hangi koşullarda?” sorularını sormak gerekir. SOTA iddiaları genel geçer değil, anlık birer durum fotoğrafı olarak değerlendirilmelidir.

Structured Reasoning

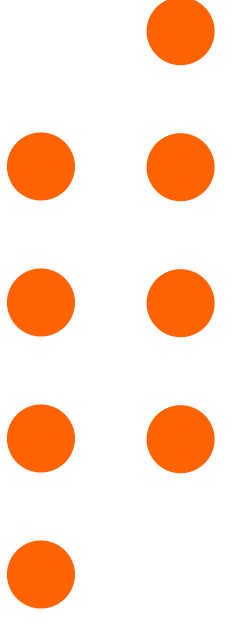
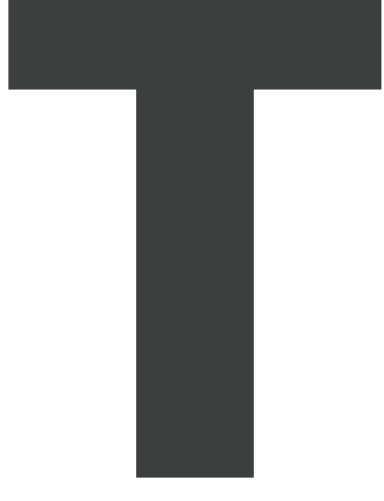
Yapılandırılmış Akıl Yürütme

Modelin tek adımda yanıt vermesi yerine; planlama, düşünce zinciri, karar ağaçları gibi açık adımlarla düşünmeye zorlanması veya yönlendirilmesidir. Bu; istem (prompt) tabanlı (“adım adım düşünelim”), araç destekli (taslaklar, program çözücüleri) ya da mimari düzeyde (düşünce ağacı, grafik temelli akıl yürütme) olabilir. Amaç; daha güvenilir mantıksal akıl yürütme ve hata durumlarının daha kolay izlenebilir olmasıdır.

Supervised Learning

Denetimli Öğrenme

Modelin, doğru cevapların bilindiği etiketli örnekler üzerinde eğitildiği yöntemdir. Model; girişten çıkışa (sınıf, sayı, sıra) nasıl eşleme yapacağını öğrenir ve yeni, görülmemiş verilerde ne kadar genelleme yapabildiğiyle değerlendirilir. Spam tespiti, dolandırıcılık puanlama, kalite kontrol gibi pek çok pratik makine öğrenimi uygulaması hâlâ bu kategoriye girer.



Tensor

Tensör

Derin öğrenmede temel veri yapısıdır—N boyutlu bir dizidir (skalarlar, vektörler, matrisler ve ötesi). PyTorch ve TensorFlow gibi framework'ler, tensör işlemlerini (toplama, çarpma, konvolüsyonlar) GPU/TPU üzerinde optimize eder; bu nedenle ileri/geri yayılım işlemlerinin tamamı aslında tensör matematiğidir.

Text-to-Image

Metinden Görüntüye

Sadece metinsel komutlardan görüntü üretmeyi ifade eder (örneğin, “şafakta kitap okuyan bir tilkinin sulu boya tablosu”). Modern sistemler; stil, kompozisyon, en-boy oranı üzerinde kontrol sağlar ve mevcut görsellerin düzenlenmesini (inpainting/outpainting) mümkün kılar—tasarım iş akışları ve yaratıcı prototipleme için kullanılır.

Text-to-Speech

Metinden Sese

Yazılı metni doğal sesli konuşmaya dönüştürür. Günümüzün TTS modelleri, vurgular, ritim ve duyguyu yakalayabilir; belirli sesleri klonlayabilir ve hatta cümle ortasında diller arasında geçiş yapabilir. Gecikme süresi, gürültülü ortamlarda anlaşılabilirlik ve sesin sentetik olduğunun açık şekilde belirtilmesi üretim ortamlarında kritik öneme sahiptir.

Token

Token

Büyük dil modellerinin (LLM) işlediği en küçük birimdir— genellikle bir alt kelime parçasıdır (örneğin “trans-”, “former”). Bağlam sınırları ve fiyatlandırmalar karakter değil token üzerinden hesaplanır, bu nedenle istem (prompt) uzunluğu genellikle token sayısını azaltarak optimize edilir. Farklı belirteçleyiciler (tokenizer) metni farklı şekillerde bölebilir ve bu durum hem maliyeti hem de model davranışını etkileyebilir.

Tool Use

Araç Kullanımı

Bir yapay zekâ sisteminin kendi başına gerçekleştiremeyeceği görevler için harici API’leri, veritabanlarını veya kod çalıştırıcıları çağırmasıdır. “Araca duyarlı” istemler veya protokoller (ör. MCP, OpenAI araç çağırma sistemi), modele hangi araçların kullanılabilir olduğunu bildirir; model de ne zaman çağıracağına karar verir ve sonuçları yanıtına entegre eder.

Training Data

Eğitim Verisi

Bir modelin öğrenim sürecinde kullandığı veri kümesidir— web metinleri, kodlar, görseller, sesler, özel belgeler vb. Veri kalitesi, çeşitliliği ve temizliği, modelin yeteneklerini ve önyargılarını doğrudan etkiler; lisanslama ve gizlilik kısıtlamaları da hangi verilerin dahil edilebileceğini belirler.

Transfer Learning

Aktarım Öğrenmesi

Bir görevden veya alandan edinilen bilgilerin başka bir görevde kullanılmasını sağlar. Sıfırdan model eğitmek yerine, büyük bir önceden eğitilmiş modeli ince ayar (fine-tune) veya istemlerle adapte edersiniz—etiketli verinin kıt olduğu özel alanlar (hukuk, tıp) için çok kritiktir.

Transformer

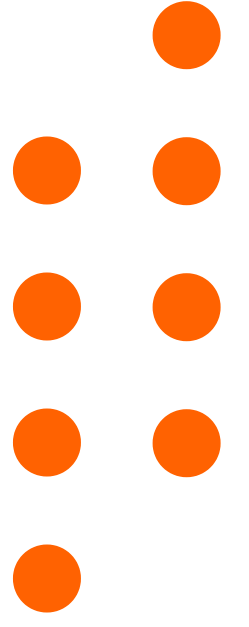
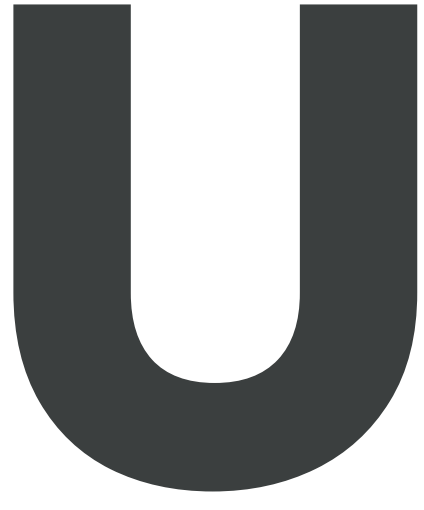
Dönüştürücü

Tamamen kendi üzerine dikkat mekanizması (self-attention) üzerine kurulu, yinleme (recurrence) içermeyen bir sinir ağı mimarisidir. 2017 tarihli “Attention Is All You Need” adlı makalede tanıtılmıştır ve günümüzdeki modern dil ve çok modlu yapay zekâ modellerinin temelini oluşturur.

Turing Test

Turing Testi

Bir makine, metin üzerinden yürütülen kör bir sohbette, bir insan hakem tarafından başka bir insanla karıştırılamayacak şekilde yanıtlar verdiğinde Turing Testi’ni geçmiş sayılır. Bu, insan düzeyinde konuşma becerisi sergilediğini gösterir.



Underfitting

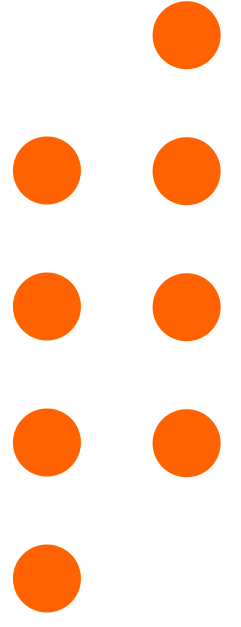
Aşırı Basitleştirme

Verideki kalıpları yakalayamayan, çok basit bir modelin durumudur. Hem eğitim verisinde hem de daha önce görülmemiş verilerde düşük performans sergiler. Bu tür modeller, öğrenmesi gereken ilişkileri kavrayamadığından genel olarak yetersiz kalır.

Unsupervised Learning

Gözetimsiz Öğrenme

Etiketlenmemiş veriden kalıp, yapı veya temsil öğrenme yöntemidir. Klasik örnekleri arasında kümeleme (clustering) ve boyut indirgeme (dimensionality reduction) yer alır. Daha yeni yaklaşımlar arasında, verinin kendi içinden sahte (proxy) etiketler üreten kendiliğinden gözetimli öğrenme (self-supervised learning) de bulunur. Gözetimsiz ön eğitim, sonrasında uygulanacak gözetimli görevler için güçlü bir temel sağlayabilir.



Vector Vektör

Bir veri noktasını cebirsel olarak işlenebilecek şekilde temsil eden, sıralı sayılar listesidir. Yapay zekâda vektörler, kelimelerden görsellere kadar her şeyi (embedding'ler aracılığıyla) temsil etmek için kullanılır. Bu sayede hızlı arama, karşılaştırma (örneğin kosinüs benzerliği) ve matematiksel akıl yürütme mümkün hale gelir. Ham metin ya da pikseller yerine kullanılan vektörler, büyük ölçekli yakın komşu aramalarında verimlilik sağlar.

Vector Database Vektör Veritabanı

Embedding vektörlerini depolamak ve en yakın komşuları hızla bulmak için tasarlanmış veritabanıdır (örneğin HNSW veya IVF gibi yaklaşık yakın komşuluk indeksleriyle). Semantik arama ve RAG (retrieval-augmented generation) süreçlerinin temelini oluşturur; milyonlarca vektör arasında milisaniyelik sorgulamalar yapılmasını sağlar.

Video Generation (AI Video) Video Üretimi (Yapay Zekâ ile)

Metin veya görsel girdilerden anlamlı video klipler oluşturan modellerdir. Modern sistemler hareket tutarlılığı, kamera geçişleri ve sahne değişimlerini uçtan uca yönetebilir; bu, eski "kareleri birleştir" yaklaşımlarının çok ötesindedir.

Vision Model

Görsel Model

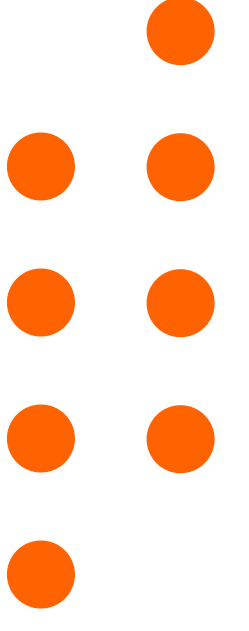
Görsel verileri anlamaya odaklanmış modellerdir; sınıflandırma, nesne tespiti, segmentasyon ve görsel açıklama üretimi gibi görevleri yerine getirirler. Günümüzde bu modeller giderek daha fazla dil modelleriyle birleştirilerek çok modlu (multimodal) akıl yürütme yapılmasını sağlar (örneğin “bu resimde ne var?” gibi sorulara yanıt).

Voice AI

Ses Tabanlı Yapay Zekâ

Konuşma tanıma (ASR), ses sentezi/klonlama ve gerçek zamanlı sesli asistanları kapsayan alandır. ASR, büyük dil modelleri (LLM) ve metinden konuşma (TTS) sistemlerini birleştirerek kullanıcıların yazılı değil sesli olarak yazılımla doğal şekilde iletişim kurmasını sağlar (ve yazılımın da sesli yanıt vermesine imkân tanır).

W

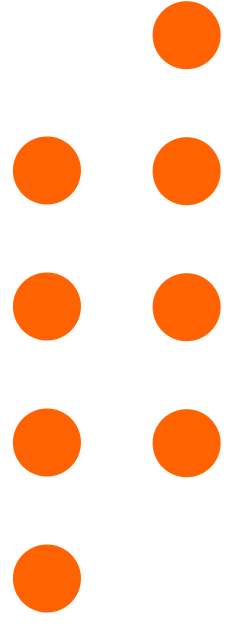


Weight Ağırlık

Nöronlar arasındaki sinyallerin nasıl iletileceğini belirleyen, öğrenilmiş bir parametredir. Milyarlarca ağırlık birlikte, modelin “bildiği” her şeyi temsil eder ve belleğini oluşturur.

Whisper Whisper

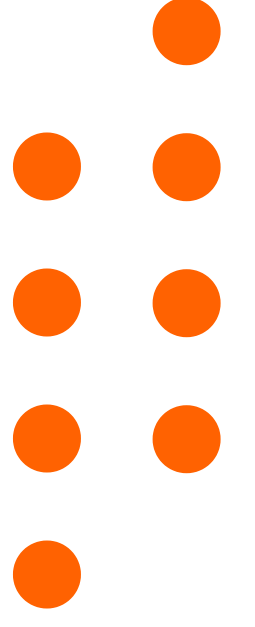
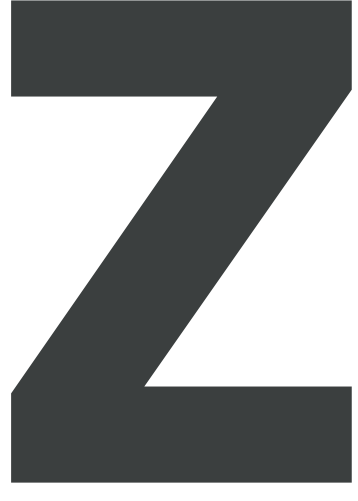
OpenAI tarafından geliştirilen açık kaynaklı bir otomatik konuşma tanıma (ASR) modelidir. Yaklaşık 680 bin saatlik çok dilli, gürültülü ve aksanlı ses verileriyle eğitilmiştir. Bu sayede kutudan çıkar çıkmaz sağlam bir transkripsiyon (metne döküm) ve İngilizceye çeviri performansı sunar.



xAI

xAI

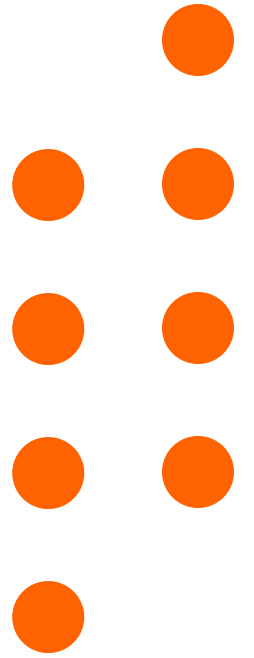
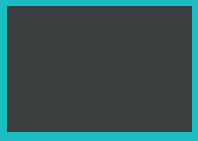
Elon Musk'ın "gerçeği arayan" modeller geliştirme hedefiyle kurduğu yapay zekâ şirketidir. Amiral gemisi asistanı Grok, alaycı bir üsluba sahip olup canlı web verileriyle etkileşime girebilen bir karaktere sahiptir.



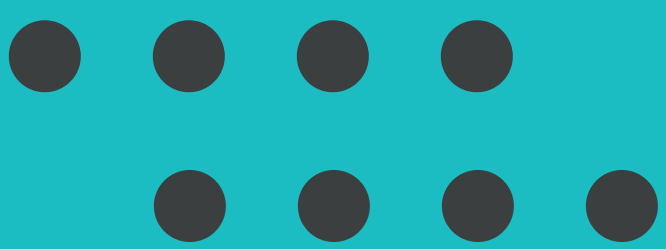
Zero-shot Learning

Sıfır Örnekli Öğrenme

Modelin, eğitim sırasında hiç görmediği bir görevi veya sınıfı, sahip olduğu genel temsil yetenekleri ya da verilen yönergeler sayesinde çözebilmesidir. Veri gerektirmemesi ve hızlı uygulanabilirliği sayesinde avantajlıdır; ancak genellikle az örnekli (few-shot) ya da özel olarak eğitilmiş modellere göre daha düşük performans gösterir ve verilen yönergenin nasıl ifade edildiğine karşı oldukça hassastır.



DAHA
FAZLASI
İÇİN



TAKİP



hybridakademi



Hybrid Akademi



<https://hybridakademi.com.tr>

ET!

