

1 Ocak – 21 Kasım 2025 AI Gelişmeleri: Kapsamlı Bir Değerlendirme

1. Öncülerin Yeni Nesil Dil Modelleri (LLM) ve AI Ajanları

1.1. OpenAI'nin Yenilikleri

1.1.1. GPT–5.1 Ailesi: Konuşma ve Akıl Yürütme Odaklı Modeller

2025 yılı boyunca OpenAI, yapay zekâ alanındaki liderliğini pekiştirmek için bir dizi yenilikçi model geliştirdi. Bu modeller, hem genel kullanıcı deneyimini iyileştirmeyi hem de karmaşık problem çözme yeteneklerini artırmayı hedefliyordu. Özellikle Kasım 2025 itibarıyla, OpenAI'nin **GPT–5.1 model ailesi**, yapay zekâ topluluğunda önemli bir dönüm noktası olarak kabul edildi. Bu model ailesi, farklı kullanım senaryolarına yönelik olarak optimize edilmiş çeşitli varyantlar sunarak, dil modellerinin sadece metin üretmekten daha fazlasını yapabileceğini gösterdi. GPT–5.1 ailesi, daha doğal ve akıcı konuşmalar kurabilen, aynı zamanda derinlemesine analiz ve mantık yürütme gerektiren görevlerde de üstün performans sergileyen modellerden oluşuyordu. Bu çift yönlü yaklaşım, OpenAI'nin hem günlük kullanıcılar hem de araştırmacılar ve geliştiriciler için kapsayıcı bir ekosistem oluşturma stratejisinin bir parçası olarak görülebilir. Ayrıca, bu modellerin piyasaya sürülmesiyle birlikte, OpenAI'nin güvenlik ve hizalama konularına verdiği önem de artarak, daha güvenilir ve öngörülebilir yapay zekâ sistemlerinin geliştirilmesine yönelik adımlar atıldı.

1.1.2. GPT–5.1 Instant: Hızlı ve Doğal Etkileşim

OpenAI'nin Kasım 2025'te piyasaya sürdüğü **GPT–5.1 Instant** modeli, kullanıcı etkileşimlerinde hız ve doğallığı ön planda tutan bir varyanttı. Bu model, özellikle gerçek zamanlı diyaloglar ve hızlı yanıt gerektiren uygulamalar için optimize edilmişti. GPT–5.1 Instant, önceki modellere kıyasla daha akıcı ve insansı bir konuşma tarzı benimsedi, bu da kullanıcıların modelle daha doğal bir şekilde iletişim kurmasını sağladı. Ayrıca, bu modelin talimatları daha iyi anlama ve yerine getirme yeteneği, onu çeşitli görevlerde daha güvenilir bir yardımcı haline getirdi. OpenAI, GPT–5.1 Instant'ın kendi başına daha derin düşünme gerektiren karmaşık sorularla karşılaştığında bunu otomatik olarak algılayıp, daha kapsamlı bir değerlendirme sürecine geçebildiğini belirtti. Bu özellik, modelin hem hızlı hem de derinlemesine analiz gerektiren durumlarda uygun şekilde davranabilmesini sağlayarak, kullanıcı deneyimini önemli ölçüde iyileştirdi. Bu yenilik, yapay zekâ asistanlarının günlük yaşamda daha yaygın ve etkili bir şekilde kullanılmasının önünü açtı.

1.1.3. GPT–5.1 Thinking: Karmaşık Görevler için Derin Akıl Yürütme

GPT–5.1 ailesinin bir diğer önemli üyesi olan **GPT–5.1 Thinking** modeli, özellikle karmaşık problem çözme ve derin akıl yürütme görevleri için tasarlandı. Bu model, daha zorlu sorunlara yönelik olarak daha ayrıntılı ve yapılandırılmış bir düşünme süreci benimsiyordu. OpenAI'nin açıklamalarına göre, GPT–5.1 Thinking, kullanıcının sorusunun karmaşıklığını analiz ederek, gerektiğinde daha uzun süreli ve derinlemesine bir değerlendirme yapabiliyordu. Bu, modelin özellikle bilimsel araştırma, veri analizi ve stratejik planlama gibi alanlarda daha tutarlı ve güvenilir sonuçlar üretmesini sağladı. Ayrıca, GPT–5.1 Thinking'in yanıtları, daha az teknik jargon içeriyor ve daha anlaşılır bir dil kullanıyordu. Bu durum, modelin sadece uzman kullanıcılar değil, genel kullanıcılar tarafından da etkili bir şekilde kullanılmasını kolaylaştırdı. GPT–5.1 Thinking'in bu özellikleri, yapay zekânın eğitim, araştırma ve profesyonel hizmetler gibi alanlarda daha yaygın bir şekilde benimsenmesine katkı sağladı.

1.1.4. Gelişmiş Güvenlik ve Hizalama Teknikleri

OpenAI, 2025 yılı boyunca modellerinin güvenliği ve hizalama konusunda önemli adımlar attı. Kasım 2025 itibarıyla, şirket, yapay zekâ sistemlerinin daha güvenilir, öngörülebilir ve etik kurallara uygun şekilde çalışmasını sağlamak için bir dizi yeni teknik ve araç geliştirdi. Bu çalışmalar kapsamında, modellerin "**halüsinasyon**" olarak adlandırılan, gerçek dışı veya yanıltıcı bilgi üretme eğilimlerini azaltmaya yönelik yöntemler geliştirildi. Ayrıca, modellerin potansiyel olarak zararlı veya istenmeyen içerik üretmesini engellemek için daha güçlü reddetme mekanizmaları oluşturuldu. OpenAI, aynı zamanda modellerinin karar alma süreçlerini daha şeffaf hale getirmek için çalışmalar yürüttü. Bu kapsamda, modelin bir sonuca nasıl ulaştığını açıklayan "**chain of thought**" (**zincirleme düşünce**) benzeri tekniklerin geliştirilmesi, kullanıcıların modelin güvenilirliğini daha iyi değerlendirmesine olanak tanıdı. Bu güvenlik ve hizalama çalışmaları, OpenAI'nin sorumlu yapay zekâ geliştirme ilkelerine bağlılığını göstererek, toplumun bu teknolojiye olan güvenini artırmayı hedefliyordu.

1.2. Google DeepMind'in AI Ajansı ve Model Gelişmeleri

1.2.1. SIMA 2: 3D Sanal Ortamlarda Oyun Oynayan ve Öğrenen AI Ajansı

Google DeepMind, 13 Kasım 2025 tarihinde, yapay zekâ ajanları alanında çığır açan bir gelişme olan **SIMA 2**'yi tanıttı. SIMA (Scalable Instructable Multiworld Agent) projesinin devamı niteliğindeki bu yeni ajan, **Gemini model ailesinin** güçlü yeteneklerini kullanarak, 3D sanal ortamlarda sadece talimatları takip etmekle kalmıyor, aynı zamanda etkileşime

geçebiliyor, kullanıcılarla sohbet edebiliyor ve zamanla kendini geliştirebiliyor . Bu özellikler, SIMA 2'yi sadece bir komut uygulayıcısı olmaktan çıkarıp, **etkilesimli bir oyun arkadaşı** ve öğrenen bir varlık haline getiriyor. DeepMind yetkilileri, bu gelişmenin yapay genel zekâ (AGI) yolculuğunda önemli bir kilometre taşı olduğunu ve robotik ile gömülü yapay zekâ sistemlerinin geleceği açısından büyük önem taşıdığını vurguladı . SIMA 2'nin, sanal dünyalarda insan dilini anlayarak anlamlı eylemlere dönüştürme yeteneği, yapay zekânın gerçek dünya ile etkileşim kurma kapasitesini temelden geliştirmeyi hedefliyor. Bu, hem eğlence sektöründe hem de sanal eğitim ve simülasyon uygulamalarında devrim yaratma potansiyeline sahip.

1.2.2. Gemini Model Ailesindeki İlerlemeler

Google'ın **Gemini model ailesi**, 2025 boyunca sürekli olarak güncellenerek daha güçlü ve çok yönlü hale getirildi. **Gemini 2.0**, 2024 yılında piyasaya sürüldükten sonra, 2025 yılında performans ve akıl yürütme yeteneklerinde önemli iyileştirmelerle güncellendi. Bu güncellemeler, modelin metin, görsel ve sesli verileri daha etkili bir şekilde işlemesini ve sentezlemesini sağladı. Ayrıca, **Gemini 2.5** sürümü, özellikle **"Computer Use" (Bilgisayar Kullanımı)** yetenekleriyle dikkat çekti. Bu yetenek, AI ajanlarının doğrudan bilgisayar arayüzlerini kullanarak çeşitli yazılımlarda görevleri yerine getirebilmesini mümkün kılıyor. Bu gelişmeler, Gemini'nin sadece bir dil modeli olmanın ötesinde, dijital dünyada eyleme geçebilen bir ajan olarak konumlandırılmasını sağlıyor. Bu strateji, Google'ın AI ajanları pazarında rekabet gücünü artırmayı ve kullanıcılara daha pratik ve entegre bir AI deneyimi sunmayı hedefliyor.

1.2.3. Gemini 3 ve AI Ajan Entegrasyonu

Google CEO'su Sundar Pichai, Dreamforce 2025 etkinliğinde **Gemini 3.0**'ın 2025 yılı sonuna kadar piyasaya sürüleceğini ve **"daha da güçlü bir ajan"** olacağını doğruladı . Bu açıklama, Gemini 3'ün sadece bir dil modeli güncellemesi olmayacağını, aynı zamanda otonom görev tamamlama ve karmaşık akıl yürütme yeteneklerinin önemli ölçüde geliştirileceğini gösteriyor. Piyasadaki spekülasyonlar, Gemini 3'ün özellikle çoklu modalite (metin, görsel, ses, video) işleme yeteneklerinde ve bağlam penceresi boyutunda büyük sıçramalar yapacağı yönünde. Ayrıca, modelin Gmail ve Docs gibi Google hizmetleriyle daha derin bir entegrasyon sunması bekleniyor. Bu entegrasyon, kullanıcıların günlük iş akışlarında AI ajanlarından daha etkili bir şekilde yararlanmasını sağlayacak. Gemini 3'ün tanıtımı, Google'ın AI stratejisinde ajan odaklı bir dönüşümün tamamlanması anlamına gelebilir ve bu da şirketin pazardaki konumunu güçlendirecektir.

1.3. Anthropic'in Güvenli ve Hizalanmış Modelleri

1.3.1. Claude Sonnet 4.5: Hızlı Yanıt ve Gelişmiş Bellek

Anthropic, Kasım 2025'te **Claude Sonnet 4.5** modelini piyasaya sürdü. Bu yeni model, önceki sürümlere kıyasla daha hızlı yanıt verme yeteneği ve gelişmiş bir bellek kapasitesi sunuyor. Claude Sonnet 4.5, özellikle uzun ve karmaşık diyaloglarda bağlamı koruma ve daha tutarlı yanıtlar üretme konusunda önemli iyileştirmeler gösteriyor. Anthropic, bu modelin geliştirilmesinde **güvenlik ve hizalama (alignment)** konularına özel bir önem verdiğini belirtiyor. Bu sayede, modelin zararlı veya yanıltıcı içerik üretme olasılığı minimize edilmeye çalışılıyor. Claude Sonnet 4.5'in tanıtımı, Anthropic'in "AI güvenliği" odaklı yaklaşımını sürdürdüğünü ve kurumsal müşterilerin güvenini kazanmaya yönelik stratejisini pekiştirdiğini gösteriyor. Modelin yüksek performansı ve güvenli yapısı, yasal, finansal ve sağlık gibi hassas sektörlerde kullanım için ideal bir seçenek haline getiriyor.

1.3.2. Ön-Eğitim Verisi Filtreleme ile Güvenlik

Anthropic, model güvenliğini artırmak için **ön-eğitim verilerini sistematik bir şekilde filtreleme** yaklaşımını benimsedi. Bu süreçte, özellikle kimyasal, biyolojik, radyolojik ve nükleer (CBRN) tehditlerle ilgili hassas bilgilerin filtrelenmesi, modellerin bu tür zararlı içerikleri üretmesini önlemeye yönelik önemli bir adım olarak öne çıktı. Bu filtreleme teknikleri, sadece açıkça zararlı verileri değil, aynı zamanda potansiyel olarak kötüye kullanılacak dolaylı veya teknik bilgileri de kapsayacak şekilde tasarlandı. Bu yaklaşım, "**garbage in, garbage out**" (**çöp içerik girerse, çöp içerik çıkar**) ilkesine dayanarak, modelin eğitim aşamasında güvenli ve faydalı bilgilerle beslenmesini sağlıyor. Bu sayede, Claude modellerinin, kullanıcı sorgularına verdiği yanıtlarda güvenlik ve etik kurallara daha uygun bir şekilde hizalanması hedefleniyor. Bu veri temizleme süreci, aynı zamanda modelin genel kalitesini ve güvenilirliğini de artırarak, hem bireysel kullanıcılar hem de kurumsal müşteriler için daha sağlam bir temel sunuyor.

1.3.3. CBRN (Kimyasal, Biyolojik, Radyolojik, Nükleer) Bilgilerin Filtrelenmesi

Anthropic'in güvenlik stratejisinin merkezinde, **CBRN (Chemical, Biological, Radiological, Nuclear)** olarak adlandırılan ve toplumsal güvenlik açısından yüksek risk taşıyan alanlardaki bilgilerin kontrollü bir şekilde filtrelenmesi yer alıyor. Bu filtreleme süreci, sadece açıkça yasa dışı veya zararlı talimatları değil, aynı zamanda bu tür faaliyetlerin gerçekleştirilmesinde kullanılacak temel bilimsel ve teknik verileri de kapsayacak şekilde tasarlanmıştır. Amaç, Claude gibi güçlü dil modellerinin, bu hassas

alanlarda bilgi paylaşımını sınırlandırarak, kötü niyetli aktörlerin bu teknolojiyi zararlı amaçlarla kullanmasını önlemektir. Bu yaklaşım, AI güvenliği alanında "**sorumlu açıklık**" (**responsible disclosure**) ilkesiyle örtüşmekte, faydalı bilgilerin yayılması ile potansiyel risklerin minimize edilmesi arasında bir denge kurmayı amaçlamaktadır. Bu filtreleme, hem eğitim verilerinin hazırlanması aşamasında hem de modelun çıktılarının gerçek zamanlı olarak denetlenmesi sırasında uygulanarak, çok katmanlı bir savunma mekanizması oluşturur.

1.4. Meta AI'nin Açık Kaynak ve Uyum Odaklı Gelişmeleri

1.4.1. Llama 4 Ailesi: Scout ve Maverick Modelleri

Meta AI, Nisan 2025'te **Llama 4 model ailesini** tanıttı. Bu aile, farklı ihtiyaçlara yönelik olarak optimize edilmiş üç ana modelden oluşuyor: **Scout, Maverick ve Behemoth** . Llama 4 Scout, 17 milyar aktif parametreye ve 10 milyon token'lık geniş bir bağlam penceresine sahip. Bu model, özellikle uzun dokümanları analiz etme ve özetleme gibi görevler için ideal. Llama 4 Maverick ise aynı aktif parametre sayısına sahip olmakla birlikte, daha büyük bir toplam parametre havuzuna (400 milyar) ve 1 milyon token'lık bir bağlam penceresine sahip. Bu model, daha karmaşık ve çok yönlü görevler için tasarlandı. Her iki model de "**mixture-of-experts**" (**uzman karışımı**) mimarisi kullanıyor, bu da hesaplama verimliliğini artırıyor. Llama 4 Behemoth ise henüz yayınlanmamış, olmakla birlikte, 288 milyar aktif parametreye ve 2 trilyon toplam parametreye sahip olacak şekilde tasarlandı. Bu modellerin tamamı, metin, görsel ve video verileri içeren çok büyük ve çeşitli veri setleri üzerinde eğitildi ve 200'den fazla dili destekliyor .

1.4.2. Otomatik Uyumluluk Sertleştirme (ACH) Aracı

Meta AI, 2025 yılında yapay zekâ modellerinin güvenli ve etik kurallara uygun şekilde çalışmasını sağlamak için **Otomatik Uyumluluk Sertleştirme (Automated Compliance Hardening – ACH)** adını verdiği yeni bir araç geliştirdi. Bu araç, LLM'lerin ürettiği çıktıları otomatik olarak analiz ederek, potansiyel olarak zararlı, yanıltıcı veya yasa dışı içerikleri tespit edebiliyor ve bu içerikleri filtreleyebiliyordu. ACH, özellikle regülasyonlara tabi olan sektörlerde (örneğin, finans, sağlık) kullanılan yapay zekâ uygulamaları için kritik önem taşıyordu. Bu araç sayesinde, şirketler yapay zekâ sistemlerini daha güvenli ve yasalara uygun bir şekilde dağıtabiliyordu. Meta'nın bu tür uyumluluk araçlarını geliştirmeye yönelik çalışmaları, şirketin sorumlu yapay zekâ geliştirme konusundaki kararlılığını yansıtıyordu.

1.4.3. Yazılım Testi ve Uyumluluk için LLM Kullanımı

Meta AI, büyük dil modellerini (LLM) yazılım testi ve uyumluluk görevleri için kullanmaya başladı. Şirket, Llama gibi modelleri kullanarak yazılım kodlarını analiz ediyor, test senaryoları oluşturuyor ve potansiyel hataları tespit ediyor. Bu yaklaşım, yazılım geliştirme sürecini hızlandırmayı ve kaliteyi artırmayı hedefliyor. Ayrıca, Meta AI, bu modelleri kullanarak yazılımın belirli düzenlemelere ve standartlara uygunluğunu da kontrol ediyor.

1.5. Mistral AI'nin Akıl Yürütme Odaklı Modeli

1.5.1. Magistral Modeli: Zincirleme Düşünce (Chain-of-Thought) Yetenekleri

Mistral AI, Haziran 2025'te **Magistral** adlı yeni bir dil modelini tanıttı. Bu model, özellikle **"Chain-of-Thought" (Zincirleme Düşünce)** yetenekleriyle öne çıkıyor. Magistral, karmaşık problemleri çözerken, çözüm adımlarını açık bir şekilde ortaya koyarak ve mantıksal bir sıra izleyerek ilerliyor. Bu yaklaşım, modelin yalnızca doğru cevabı bulmasını değil, aynı zamanda bu cevaba nasıl ulaştığını da açıklamasını sağlıyor. Bu özellik, özellikle eğitim, bilimsel araştırma ve teknik destek gibi alanlarda, AI'nın karar alma süreçlerinin şeffaf ve anlaşılır olmasını gerektiren uygulamalar için büyük bir avantaj sağlıyor. Mistral AI, Magistral'ın bu yeteneğinin, özel olarak tasarlanmış eğitim verileri ve ince ayar teknikleri kullanılarak geliştirildiğini belirtiyor. Bu gelişme, AI modellerinin sadece "ne" bildiğini değil, "nasıl" düşündüğünü de anlamamıza yardımcı olacak önemli bir adım olarak değerlendiriliyor.

1.5.2. Magistral Small 1.2 Sürümü

Mistral AI, Magistral modelinin daha hafif ve hızlı bir versiyonu olan **"Magistral Small 1.2"** sürümünü de yayınladı. Bu model, daha düşük hesaplama kaynaklarına sahip cihazlar için optimize edildi ve hızlı yanıt gerektiren uygulamalarda kullanılmak üzere tasarlandı.

2. AI Ajanları ve Otomasyon Sistemlerindeki Uygulamalar

2.1. Kurumsal ve Ticari AI Ajan Platformları

2.1.1. Freshworks Freddy AI Agent: Kendi Kendine Hizmet ve Görüşme Analizi

Freshworks, Kasım 2025'te **Freddy AI Agent** adlı yeni bir AI ajanı platformunu tanıttı. Bu platform, özellikle çalışan deneyimi (EX) ve müşteri deneyimi (CX) alanlarında kendi kendine hizmet ve otomasyon çözümleri sunuyor. Freddy AI Agent, Service Portallar ve Microsoft 365 Copilot gibi platformlara yerleşik olarak çalışarak, çalışanlara

kisşelleştirilmiş ve srtnmesiz destek saęlıyor. Ajan, Google Drive, Confluence ve Sharepoint gibi kaynaklardan bilgi ekebilir ve hatta grntleri tanıyarak, insan mdahalesi olmadan alıřanların ihtiya duyduęu yanıtları sunabilir. Ayrıca, **Freddy AI Insights** adlı bir başka zellik, liderlerin ve ekiplerin hizmet verilerini analiz ederek anlamlı grřler elde etmelerini saęlar. Bu platform, AI ajanlarının sadece otomasyon deęil, aynı zamanda veri odaklı karar verme srelerinde de nemli bir rol oynayabileceęini gsteriyor .

2.2. AI Ajanlarının Kullanım Alanları

2.2.1. E–posta ve Takvim Ynetimi

2025 yılı boyunca AI ajanları, e–posta ve takvim ynetimi gibi gnlk dijital grevlerin otomasyonunda nemli bir rol stlenmeye bařladı. Google'ın Kasım 2025'te tanıttıęı Gemini Agent, kullanıcının adına e–posta kutusunu dzenleme, nemli mesajları nceliklendirme ve hatta toplantıları takvimlemek gibi grevleri yerine getirebiliyor. Bu ajanlar, doęal dil komutlarını anlayarak kullanıcının dijital yařamını kolaylařtırmayı hedefliyor.

2.2.2. Seyahat Planlama ve Rezervasyon

AI ajanları, seyahat planlama ve rezervasyon srelerini de byk lde kolaylařtırmaya bařladı. Kullanıcılar, "nmzdeki hafta sonu iin Paris'e ucuz bir uus bul" gibi doęal dil komutları vererek, ajanların uus arama, otel rezervasyonu ve hatta gezi planı oluřtırma gibi tm sreleri otomatik olarak yrtmesini saęlayabiliyor. Bu geliřme, seyahat endstrisinde kisşelleştirilmiş ve srtnmesiz bir mřteri deneyimi yaratma potansiyeline sahip.

2.2.3. Mřteri Hizmetleri ve Destek Otomasyonu

2025 yılı itibarıyla, AI ajanlarının en yaygın ve etkili kullanım alanlarından biri **mřteri hizmetleri ve destek otomasyonu** oldu. Upwork platformunda yapılan bir analiz, AI ajanı geliřtirme projelerinin **%14.8'inin** doęrudan mřteri desteęi otomasyonuna ynelik olduęunu gsterdi . Bu ajanlar, 7/24 kesintisiz hizmet sunarak, mřteri sorularına anında yanıt verme, sorunları nceliklendirme ve gerekli durumlarda insan temsilciye ynlendirme gibi grevleri stleniyor. E–posta, sohbet, sesli arama ve WhatsApp gibi ok kanallı iletiřimi destekleyerek, mřteri deneyimini btnsel bir řekilde iyileřtiriyorlar. Bu sistemler, genellikle byk dil modelleri (LLM) ile vektr veritabanları ve CRM sistemlerinin entegrasyonuyla oluřturuluyor. Bu sayede, ajanlar bilgi tabanlarından anında bilgi ekebilir, mřteri gemiřini analiz edebilir ve kisşelleştirilmiş zmler

sunabilir. Sonuç olarak, işletmeler hem maliyetlerini önemli ölçüde düşürürken hem de müşteri memnuniyetini artırarak rekabet avantajı elde ediyor. Bu ajanlar, insan temsilcilerin rutin görevlerle uğraşmasını engelleyerek, onların daha karmaşık ve değer odaklı sorunlara odaklanmasına olanak tanıyor .

2.3. AI Ajanı Geliştirme Eğitimleri

2.3.1. Google'ın "5-Day AI Agents Intensive" Kursu

Yapay zekâ ajanlarının giderek artan önemi, bu alanda yetkinlik kazanmak isteyen geliştiriciler için eğitim ihtiyacını da beraberinde getirdi. Google ve Kaggle, bu ihtiyaca cevap vermek üzere 10–14 Kasım 2025 tarihlerinde **"5-Day AI Agents Intensive"** adlı yoğun bir eğitim programı düzenledi . Bu kurs, temel yapay zekâ kavşarlarını bilen ancak ajan geliştirme konusunda deneyim sahibi olmayan katılımcılar için tasarlandı. Programın amacı, katılımcılara Python ve Google Colab gibi araçları kullanarak kendi AI ajanlarını geliştirme becerisini kazandırmaktı. Kursun sonunda, katılımcılar ajan mimarisi, araç kullanımı, bağlam yönetimi ve değerlendirme gibi konularda pratik deneyim kazanarak, üretime hazır ajanik sistemler geliştirebilecek düzeye geliyorlardı. Bu tür eğitimler, AI alanındaki hızlı gelişmelere paralel olarak, nitelikli insan kaynağı ihtiyacını karşılamak adına oldukça değerlidir.

3. Açık Kaynak Topluluğundaki Gelişmeler

3.1. PyTorch Ekosistemindeki İlerlemeler

3.1.1. PyTorch Conference 2025: Kasım 12–13 Tarihli Etkinlik

PyTorch topluluğu, 12–13 Kasım 2025 tarihlerinde **PyTorch Conference 2025** etkinliğini düzenledi. Bu konferans, derin öğrenme araştırmacıları, geliştiricileri ve endüstri profesyonellerini bir araya getirerek PyTorch ekosistemindeki en son gelişmeleri tartışmak için bir platform sağladı. Konferansta, PyTorch framework'ünün yeni sürümlerindeki güncellemeler, performans optimizasyonları ve yeni özellikler detaylı bir şekilde ele alındı. Ayrıca, PyTorch'un endüstrideki kullanım durumları, büyük ölçekli model eğitimi ve dağıtımı konularında da oturumlar düzenlendi. Bu etkinlik, PyTorch'un sadece bir açık kaynak kütüphanesi olmanın ötesinde, sürekli olarak gelişen ve genişleyen bir topluluk ve ekosistem etrafında şekillenen bir platform olduğunu vurguladı.

3.1.2. Ray'in PyTorch Foundation'a Katılımı

PyTorch Conference 2025'te önemli bir duyuru da, dağıtık hesaplama framework'ü **Ray'in PyTorch Foundation'a** resmi olarak katılacağı oldu. Bu gelişme, PyTorch ve Ray toplulukları arasındaki işbirliğini resmileştiriyor ve büyük ölçekli AI iş yüklerinin yönetiminde daha entegre ve güçlü bir ekosistem yaratmayı hedefliyor. Ray, özellikle dağıtık eğitim ve çıkarım görevlerinde kaynak yönetimi ve ölçeklenebilirlik sağlayarak, geliştiricilerin büyük modelleri daha verimli bir şekilde eğitmesine ve dağıtmasına olanak tanıyor. Bu entegrasyon, AI araştırmacılarının ve mühendislerinin, daha büyük ve daha karmaşık modelleri daha verimli bir şekilde geliştirmelerini ve dağıtmalarını sağlayarak, açık kaynak AI alanındaki ilerlemeyi hızlandıracaktır.

3.1.3. Dağıtık Hesaplama ve AI İş Yükleri İçin İşbirliği

PyTorch ve Ray toplulukları arasındaki işbirliği, dağıtık hesaplama ve AI iş yüklerinin yönetiminde önemli bir adım olarak değerlendiriliyor. Bu entegrasyon sayesinde, geliştiriciler PyTorch modellerini geliştirirken ve dağıtırken karşılaştıkları karmaşıklığı azaltabilecek ve üretime geçiş sürecini hızlandırabilecek. Ray'in PyTorch ekosistemine katkısı, özellikle büyük dil modelleri (LLM) ve kompleks AI sistemlerinin eğitimi ve servis edilmesinde kritik önem taşıyor. Ray, kaynak yönetimi, görev zamanlama ve hata toleransı gibi dağıtık sistemlerin temel zorluklarını soyutlayarak, geliştiricilerin AI modellerine odaklanmasına olanak tanır. PyTorch ile birlikte kullanıldığında, Ray, birden fazla GPU veya makine üzerinde verimli bir şekilde paralel eğitim yapmayı, büyük veri kümelerini işlemeyi ve gerçek zamanlı çıkarım hizmetleri sunmayı kolaylaştırır.

3.2. Hugging Face ve Diğer Açık Kaynak Katkıları

3.2.1. Parametre Verimli İnce Ayar (PEFT) Tekniklerindeki Araştırmalar

Hugging Face, büyük dil modellerini (LLM) özel görevler için verimli bir şekilde uyarlamak amacıyla geliştirilen **Parametre Verimli İnce Ayar (PEFT – Parameter-Efficient Fine-Tuning)** tekniklerindeki araştırmalarına 2025 yılında da devam etti. PEFT yöntemleri, büyük modellerin tüm parametrelerini güncellemek yerine, sadece küçük bir alt kümesini eğiterek, hem hesaplama maliyetini hem de gereken bellek miktarını önemli ölçüde azaltır. Hugging Face, bu alandaki çalışmalarını **PEFT** kütüphanesi aracılığıyla topluluğa sunarak, geliştiricilerin LoRA (Low-Rank Adaptation), Prefix Tuning ve AdaLoRA gibi popüler PEFT yöntemlerini kolayca uygulamasını sağladı. Bu araştırmalar, özellikle kaynakları sınırlı olan araştırmacılar ve girişimler için büyük dil modellerini özel görevler için uyarlamayı daha erişilebilir hale getirdi.

3.2.2. Açık Kaynak Model ve Veri Seti Paylaşımı

2025 yılı boyunca, açık kaynak topluluğu model ve veri seti paylaşımında büyük bir artış yaşadı. Hugging Face, Model Hub ve Datasets kütüphaneleri aracılığıyla, araştırmacıların ve geliştiricilerin kendi modellerini ve veri setlerini kolayca paylaşmasına ve keşfetmesine olanak tanıdı. Bu paylaşım kültürü, AI alanındaki yeniliği hızlandırmaya ve tekrar kullanılabilirliği artırmaya yardımcı oldu. Özellikle, Llama 4 gibi açık kaynak modellerin yayınlanmasıyla birlikte, topluluk tarafından bu modellerin çeşitli görevler için ince ayarlanmış versiyonları ve sentetik veri setleri hızla paylaşıldı. Bu durum, açık kaynak AI ekosisteminin dinamik ve işbirliğine dayalı doğasını vurguladı.

4. Veri Seti Geliştirme ve Akıl Yürütme Tekniklerindeki Akademik ve Teknik İlerlemeler

4.1. Sentetik Veri Seti Oluşturma ve Düzenleme

4.1.1. Causal AI ile Sentetik Tüketici Araştırma Verisi Üretimi

2025 yılında, özellikle tüketici araştırma alanında, **Causal AI (Nedensel AI)** kullanarak sentetik veri üretimi önemli bir gelişme gösterdi. Simulacra'nın kurucusu ve CEO'su Jason Cohen, Kasım 2025'te IFT FIRST etkinliğinde yaptığı sunumda, bu yaklaşımın geleneksel yöntemlerin sınırlamalarını nasıl aşabileceğini açıkladı. Geleneksel tüketici araştırmaları, pahalı olmaları, sınırlı örneklem büyüklüğüne sahip olmaları ve çoğu zaman istatistiksel olarak geçerli sonuçlar döndürmemeleri gibi zorluklarla karşı karşıyadır. Sentetik veri, bu boşlukları doldurarak, kohort analizi, çapraz tablolulu içgörülerini güçlendirme veya yönetsel sonuçları daha fazla istatistiksel güçle keşfetme olasılığı sunar. Cohen, büyük dil modellerinin (LLM) bu tür yapılandırılmış tablolar için uygun bir araç olmadığını vurguladı. LLM'lerin çoğu, internetten taranmış yapılandırılmamış metinler üzerinde eğitilir ve bu veriler, pazar araştırmasında önemli olan demografik veya davranışsal özniteliklerle neredeyse hiç bağlantılı değildir. Bu da çeşitli önyargılara (demografik, zamansal, kültürel) yol açar. Alternatif olarak, Simulacra'nın geliştirdiği "causal AI" modeli, yalnızca kullanıcı tarafından sağlanan yapılandırılmış araştırma veri seti üzerinde eğitilir. Bu model, veri seti içindeki her bir kohort ve değişken arasındaki nedensel ilişkileri öğrenir ve ardından bu ilişkileri koruyarak yeni sentetik veriler üretir. Bu yöntem, mevcut veri setini genişletirken istatistiksel özelliklerini ve içsel ilişkilerini korur, halüsinasyon veya yanlış üretim riskini ortadan kaldırır.

4.1.2. Kullanıcı Tarafından Sağlanan Verilerle Eğitim

2025 yılında, sentetik veri üretiminde kullanıcı tarafından sağlanan verilerin (user-provided data) kullanımı önemli bir eğilim haline gelmiştir. Bu yaklaşım, özellikle Causal

AI gibi tekniklerde öne çıkmakta ve veri gizliliği ile modelin özgünlüğü arasında bir denge kurmaktadır . Kullanıcılar, kendi özel veya hassas veri setlerini sisteme yükleyerek, bu veriler üzerinde eğitilmiş bir model aracılığıyla sentetik veri üretimi yapabilmektedir. Bu yöntemin temel avantajı, üretilen sentetik verinin, kullanıcının özgün veri setinin istatistiksel özelliklerini ve içsel ilişkilerini koruyabilmesidir. Bu da, özellikle finans alanlarda veya özel iş gereksinimleri için oldukça değerlidir. Örneğin, bir şirket, müşteri memnuniyeti anketlerinden elde ettiği verileri kullanarak, daha büyük ve çeşitli bir sentetik veri seti oluşturabilir ve bu veri setini pazarlama stratejilerini optimize etmek veya yeni ürün fikirlerini test etmek için kullanabilir . Bu süreçte, veri gizliliği korunur çünkü sentetik veri, gerçek bireylerin bilgilerini içermez. Aynı zamanda, bu yaklaşım, veri paylaşımı ve iş birliği için de güvenli bir zemin sunar. Şirketler, hassas bilgileri ifşa etmeden, sentetik verileri paylaşarak ortak projeler yürütebilir veya dış danışmanlarla çalışabilir.

4.1.3. Veri Seti Kalitesini Artırmak için Filtreleme Teknikleri

2025 yılı boyunca, AI modellerinin performansını ve güvenliğini artırmak amacıyla veri seti filtreleme tekniklerine yönelik ilgi ve yatırım önemli ölçüde arttı. Bu eğilim, özellikle büyük dil modellerinin (LLM) ön-eğitim aşamasında kullanılan internet ölçekli verilerin içerdiği gürültü, yanlışlık ve zararlı içeriklerle başa çıkmak için kritik hale geldi. Anthropic ve Oxford Üniversitesi gibi öncü kurumlar, veri filtrelemenin model güvenliğini sağlamada proaktif ve etkili bir yöntem olduğunu gösteren önemli araştırmalar yayımladı . Bu çalışmalar, veri filtrelemenin sadece modelin zararlı bilgileri öğrenmesini engellemekle kalmadığını, aynı zamanda modelin genel yeteneklerini koruyarak bu güvenlik kazanımını sağladığını ortaya koydu. Veri filtreleme süreçleri, genellikle çok aşamalı bir yaklaşım benimsiyor. Örneğin, Oxford Üniversitesi'nin çalışmasında, anahtar kelime engel listeleri ve yüksek riskli içerikleri tespit etmek için eğitilmiş bir makine öğrenimi sınıflandırıcısından oluşan bir filtreleme hattı kullanıldı. Bu yöntem, veri setinin yalnızca ilgili materyallerini (yaklaşık %8–9'u) çıkararak genel bilgi derinliğini ve genişliğini korudu. Ayrıca, derin öğrenme için büyük ölçekli veri setleri hazırlanırken da özgün filtreleme teknikleri uygulanıyor. Örneğin, bir küresel deprem verisi oluşturma çalışmasında, fazla eksik veya geçersiz piksel içeren interferogramlar (>%20) otomatik olarak elendi. Bu tür filtreleme adımları, model eğitimi için yalnızca yüksek kaliteli ve güvenilir verilerin kullanılmasını sağlayarak, modelin genelleme yeteneğini ve kararlılığını artırıyor .

4.2. Akıl Yürütme (Reasoning) Tabanlı Yaklaşımlar

4.2.1. Zincirleme Düşünce (Chain-of-Thought) ile Veri Seti Geliştirme

Zincirleme düşünce (Chain-of-Thought, CoT) tekniđi, 2025 yılında sentetik veri seti oluşturma süreçlerinde yaygın olarak kullanılan bir yöntem haline gelmiştir. CoT, bir problemin çözümünü adım adım açıklayan bir dizi ara adım (intermediate reasoning steps) oluşturmayı teşvik eder. Bu teknik, özellikle karmaşık akıl yürütme gerektiren görevlerde, büyük dil modellerinin (LLM) performansını artırmak için etkili olmuştur. Sentetik veri üretiminde CoT'nin kullanımı, modelin sadece sonuca ulaşmakla kalmayıp, aynı zamanda bu sonuca nasıl ulaştığını da açıklamasını sağlar. Bu da, üretilen verinin kalitesini ve güvenilirliğini artırır. Örneğin, bir matematik problemi için sentetik veri üretilirken, CoT kullanılarak modelin problemi çözerken izlediđi adımlar (örneğin, "Önce denklemleri düzenle, sonra değişkenleri yalnızlaştır, son olarak değeri hesapla") da veri setine dahil edilir. Bu yaklaşım, modelin sadece doğru cevabı üretmesini değil, aynı zamanda bu cevabı nasıl elde ettiğini de öğrenmesini sağlar. Bu da, modelin genelleme yeteneğini artırır ve daha karmaşık ve görülmemiş problemleri çözme kapasitesini geliştirir. Ayrıca, CoT tabanlı sentetik veri setleri, modelin açıklanabilirliğini (explainability) de artırır, çünkü modelin karar verme süreci daha şeffaf hale gelir.

4.2.2. Causal Reasoning (Nedensel Akıl Yürütme) ve LLM Entegrasyonu

2025 yılı boyunca, **nedensel akıl yürütme (causal reasoning)** AI alanında giderek daha önemli bir hale geldi. Geleneksel AI sistemlerinin, verilerdeki korelasyonları bulmada başarılı olsa da, neden-sonuç ilişkilerini anlamada yetersiz kaldığı görüldü. Bu nedenle, AI'nın "neden" sorusunu yanıtlayabilmesi ve insan benzeri bir şekilde neden-sonuç ilişkileri hakkında akıl yürütebilmesi gerektiđi vurgulandı. Bu yaklaşım, AI sistemlerinin şeffaflığını, açıklanabilirliğini ve değişen koşullara uyum sağlama yeteneğini artırarak, özellikle kurumsal AI uygulamalarında güven sorununu çözmeyi hedefliyor. Judea Pearl gibi öncülerin çalışmaları, nedensel akıl yürütmenin AI'nın bir sonraki sınırı olduđu fikrini güçlendirdi. Teknoloji devleri, nedensel AI'ya önemli yatırımlar yaptı. Microsoft Research, DoWhy ve EconML gibi açık kaynak kütüphaneleriyle nedensel çıkarım ve keşif alanındaki çabalarını birleştiren Project PyWhy'ı başlattı. IBM Research, biyomedikal araştırmalarda nedenselliđi kullanarak, kofaktörleri ayarlayarak daha doğru ilaç hedefi belirleme imkânı sağladı. Amazon Research, DoWhy kütüphanesine yeni nedensel algoritmalar ekleyerek, kök neden analizi ve kişiselleştirilmiş politika değerlendirmesi gibi yetenekleri geliştirdi. Google DeepMind ve Meta AI da, AI güvenlik sistemleri, öneri algoritmaları ve reklam optimizasyonu gibi alanlara nedensel akıl yürütmeyi entegre etti. Gartner'ın raporları, nedensel AI'nın 2-5 yıl içinde "yüksek etkili" bir teknoloji olacağını öngörüyor. 2023 yılında yapılan bir ankete göre, AI lideri şirketlerin %56'sı zaten nedensel AI tekniklerini kullanıyor veya deniyor, %25'i ise 2026 yılına kadar bu teknolojiyi benimsemeyi

planlıyor. Bu eğilim, AI'nın sadece tahmin etmekten ziyade kararları açıklamak ve optimize etmek için nasıl çalıştığını anlamaya yönelik bir kayısı gösteriyor .

4.2.3. Model Sağlamlığını Artırmak İçin Causal Learning

Nedensel öğrenme (causal learning) , büyük dil modellerinin (LLM) sağlamlığını ve güvenilirliğini artırmak için umut verici bir yaklaşım olarak ortaya çıktı. Geleneksel makine öğrenimi modelleri, eğitim verisindeki korelasyonlara dayalı olarak öğrendikleri için, bu korelasyonlar değiştiğinde performansları düşebilir. Nedensel öğrenme ise, modellerin veri içindeki neden–sonuç ilişkilerini öğrenmesini sağlayarak, bu tür değişikliklere karşı daha dirençli olmalarını sağlar. Örneğin, bir görüntü tanıma modeli, eğitim verisinde "açık hava" ortamı ile "arabalar" arasında bir korelasyon öğrenmiş, olabilir. Ancak, test verisinde kapalı bir garajda arabaların bulunduğu görüntüler varsa, modelin performansı düşebilir. Nedensel öğrenme, modelin "açık hava" ve "araba" arasındaki korelasyonu değil, "araba" nesnesinin asıl özelliklerini (örneğin, tekerlekler, kapıları, şekli) öğrenmesini sağlayarak, modelin farklı ortamlarda da arabaları tanıyabilmesini mümkün kılar. 2025 yılında, araştırmacılar, nedensel öğrenme tekniklerini kullanarak, modellerin spurious correlations (yanıltıcı korelasyonlar) üzerinde değil, gerçek ve kalıcı özellikler üzerinde öğrenmesini sağlayan yeni yöntemler geliştirmiştir. Bu da, modellerin dağılım kayması (distribution shift) gibi gerçek dünya senaryolarında daha sağlam ve güvenilir performans sergilemesini sağlar.

4.3. Eğitim Sonrası (Post–Training) ve İnce Ayar Teknikleri

4.3.1. LLM'lerin Eğitim Sonrası Optimizasyonu

2025 yılında, büyük dil modellerinin (LLM) eğitim sonrası optimizasyonu, modelin performansını ve güvenliğini artırmak için kritik bir aşama olarak öne çıktı. Bu süreç, genellikle üç ana adımdan oluşur: **süpervizörlü ince ayar (supervised fine–tuning – SFT)** , **ödül modeli eğitimi (reward model training)** ve **PPO (Proximal Policy Optimization)** gibi pekiştirmeli öğrenme teknikleriyle insan geri bildiriminden öğrenme (Reinforcement Learning from Human Feedback – RLHF). Bu adımlar, modelin belirli görevlerde daha iyi performans göstermesini, istenmeyen veya zararlı çıktılar üretmesini azaltmasını ve insan değerleriyle daha iyi hizalanmasını sağlar. 2025 yılında, bu tekniklerin yanı sıra, **DPO (Direct Preference Optimization)** gibi daha verimli ve kararlı alternatifler de geliştirildi. DPO, karmaşık bir ödül modeli ve PPO eğitimi gerektirmeden, doğrudan tercih verileri üzerinden modeli optimize ederek, ince ayar sürecini basitleştirir ve hızlandırır.

4.3.2. Öğretmen–Öğrenci Modeli ile Veri Seti Oluşturma

Öğretmen–öğrenci (teacher–student) modeli, 2025 yılında sentetik veri seti oluşturmak için kullanılan etkili bir teknik olarak öne çıktı. Bu yöntemde, büyük ve güçlü bir "öğretmen" model, geniş ve çeşitli bir veri seti üzerinde eğitilir. Ardından, bu öğretmen model, daha küçük ve verimli bir "öğrenci" modelin eğitimi için yüksek kaliteli sentetik veriler üretir. Öğretmen model, karmaşık görevlerde uzmanlaşmış ve derin akıl yürütme yeteneklerine sahip olabilir. Öğrenci model ise, bu sentetik verileri kullanarak, öğretmenin yeteneklerini daha hafif ve hızlı bir yapıda öğrenmeyi hedefler. Bu yaklaşım, özellikle kaynakları sınırlı olan cihazlarda (örneğin, mobil cihazlar, kenar cihazları) yüksek performanslı AI modelleri dağıtmak için kullanışlıdır. Ayrıca, öğretmen modelin ürettiği sentetik veriler, öğrenci modelin öğrenmesini yönlendirerek, daha iyi genelleme ve sağlamlık sağlar.

5. Kasım 2025'teki Akademik Konferanslar ve Sempozyumlar

5.1. AAAI 2025 Fall Symposium Series

5.1.1. AI Güvenilirliği ve Risk Değerlendirmesi

Kasım 2025'te düzenlenen AAAI 2025 Fall Symposium Series kapsamında, AI güvenilirliği ve risk değerlendirmesi konularına odaklanan oturumlar düzenlendi. Bu oturumlarda, büyük dil modellerinin potansiyel riskleri, halüsinasyon eğilimleri ve istenmeyen davranışları ele alındı. Araştırmacılar, model güvenliğini artırmak için geliştirilen yeni teknikleri, örneğin ön–eğitim verisi filtreleme, adversarial training ve açıklanabilir AI yöntemlerini tartıştı. Bu çalışmalar, AI teknolojisinin daha güvenli ve sorumlu bir şekilde geliştirilmesi ve toplumun yararına kullanılması için önemli bir adım olarak değerlendirildi.

5.1.2. Causal Reasoning ve AI Uygulamaları

AAAI 2025 Fall Symposium Series'nde, nedensel akıl yürütme (causal reasoning) ve AI uygulamaları da önemli bir gündem maddesi oldu. Bu oturumlarda, nedensel ilişkileri anlama yeteneğinin AI sistemlerinin güvenilirliğini ve karar verme süreçlerinin şeffaflığını nasıl artırabileceği tartışıldı. Araştırmacılar, sağlık, finans ve sosyal bilimler gibi alanlarda nedensel AI'nın potansiyel uygulamalarını sundu. Bu çalışmalar, AI'nın sadece korelasyonları değil, aynı zamanda neden–sonuç ilişkilerini de anlayabilen bir sonraki seviyeye evrildiğini gösteriyor.

5.1.3. Agentic AI ve Otonom Sistemler

AAAI 2025 Fall Symposium Series kapsamında, agentic AI ve otonom sistemler konusunda da yoğun bir ilgi gözlemlendi. Bu oturumlarda, AI ajanlarının planlama, karar verme ve eyleme geçme yetenekleri ele alındı. Araştırmacılar, çoklu ajan sistemleri, insan-AI işbirliği ve otonom robotik gibi konulardaki en son gelişmeleri paylaştı. Bu çalışmalar, AI'nın sadece pasif bir araç değil, aynı zamanda aktif ve özerk bir ajan olarak kullanılabileceğini gösterdi.

5.2. PyTorch Conference 2025

5.2.1. Derin Öğrenme Framework'lerindeki Güncellemeler

PyTorch Conference 2025'te, derin öğrenme framework'lerindeki güncellemeler ve yeni özellikler detaylı bir şekilde ele alındı. PyTorch topluluğu, framework'ün performansını, kullanım kolaylığını ve ölçeklenebilirliğini artırmak için yaptığı çalışmaları paylaştı. Ayrıca, PyTorch'un diğer açık kaynak projelerle (örneğin, Ray) entegrasyonu ve endüstrideki kullanım durumları da tartışıldı. Bu oturumlar, PyTorch'un sadece bir araç değil, aynı zamanda sürekli olarak gelişen ve genişleyen bir ekosistem olduğunu vurguladı.

5.2.2. Endüstri ve Akademi İşbirliği Oturumları

PyTorch Conference 2025'te, endüstri ve akademi işbirliği oturumları da önemli bir yer tuttu. Bu oturumlarda, teknoloji şirketlerinin ve akademik kurumların PyTorch'u nasıl kullandığı, karşılaştıkları zorluklar ve çözüm yolları tartışıldı. Ayrıca, endüstrinin ihtiyaç duyduğu yeni özellikler ve akademik araştırmalardaki en son gelişmeler arasındaki köprülerin nasıl kurulabileceği ele alındı. Bu oturumlar, açık kaynak topluluğunun endüstri ve akademi arasındaki işbirliğini nasıl teşvik ettiğini ve bu işbirliğinin AI alanındaki ilerlemeyi nasıl hızlandırdığını gösterdi.