

# ÜRETKEN YAPAY ZEKÂ MODELLERİ: KARŞILAŞTIRMA, ÜCRETLENDİRME VE KULLANIM MODELLERİ

Melik Başaran

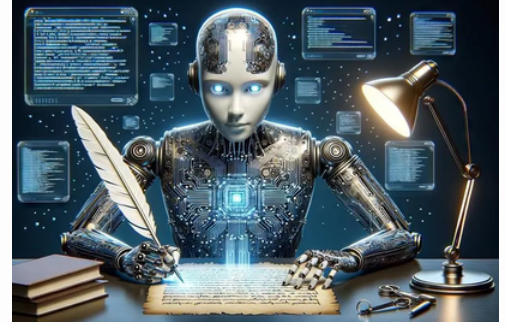
Gazi Üniversitesi Yapay Zeka Uygulama ve Araştırma Merkezi

basaranmelik2@gmail.com

DOI: 10.5281/zenodo.15367584

## 1. GİRİŞ

Üretken yapay zeka tarafından çok karmaşık yapıya sahip olan dilin kullanılması önemli bir dönüm noktası olmuştur. Bu teknolojiyi geleneksel makine öğrenmesi yaklaşımlarından ayıran temel fark, algoritmik tahminlerin ötesinde özgün içerikler üretebilmesidir (Bozkurt, 2023). Üretken yapay zekâ (Generative AI) modelleri, doğal dil işleme, görüntü üretimi, kod yazımı ve ses sentezi gibi birçok alanda çığır açan yenilikler sunmaktadır (Feuerriegel vd., 2024). Özellikle büyük dil modelleri (large language models - LLMs) metin üretimi ve bilgi erişimi sistemlerinde yaygın olarak kullanılmaktadır. Bu çalışmada, üretken yapay zekâ modellerinin temel kavramları, karşılaştırma kriterleri, ücretlendirme yapıları ve API/Chat kullanım modelleri ele alınacaktır. Üretken yapay zekâ modellerinin doğru değerlendirilmesi için şu temel kavramların anlaşılması gerekmektedir:



- LLM: Bu modeller, geniş bir veri kümesiyle eğitilen ve doğal dildeki karmaşıklıkları anlamak, bu karmaşayı çözümlmek ve metin üretmek için kullanılan derin öğrenme algoritmalarıdır (Akıllı & Şimşek, 2024).
- Parametre Sayısı: Bir modelin karmaşıklığını ve öğrenme kapasitesini belirleyen değişken sayısıdır.
- Token: Modelin metni işlemekte kullandığı en küçük birimdir ve genellikle kelime parçalarına bölünerek hesaplanır.
- Konteks Uzunluğu: Modelin tek seferde işleyebileceği maksimum token sayısını ifade eder.
- Application Programming Interface (API): Geliştiricilerin üretken yapay zekâ modellerine erişmesini ve kendi uygulamalarına entegre etmesini sağlayan ara yüzdür.
- Chat Kullanım Modeli: Web tabanlı veya uygulama içi sohbet arayüzleri üzerinden yapay zekâ ile etkileşim kurmayı sağlar.

Bu çalışmanın amacı, üretken yapay zekâ modellerini teknik ve ekonomik açıdan değerlendirmek, fiyatlandırma modellerini açıklamak ve farklı kullanım senaryolarına göre API ile chat tabanlı çözümleri karşılaştırmaktır.

## 2. ÜRETKEN YAPAY ZEKÂ MODELLERİNİN KARŞILAŞTIRILMASI

### 2.1. Model Özellikleri ve Performans Kriterleri

Üretken yapay zekâ modelleri, aşağıdaki kriterler esas alınarak karşılaştırılabilir:

- Model Adı ve Sürümü: Yapay zekâ modelinin (GPT-4, Claude 2) ve hangi sürüm numarasını kullandığını, böylece modelin teknik özelliklerini ve eğitim güncelliğini belirten ifadedir.
- Görsel ve Metin Tabanlı Dosya İnceleme: Modelin, PDF, DOCX veya benzeri biçimlerdeki görsel ve yazılı içerikleri analiz edip özetleme veya bilgi çıkarımı yapabilmesini ifade eder.
- Görsel Üretme: Yapay zekanın üretken çekışmeli ağlar (generative adversarial networks-GANs) ile daha önce var olmamış sayısız grafik tasarım ürünü üretebilme yeteneğini ifade eder (Ezgi & Atiker, 2020).
- Video Üretme: Metin veya diğer girdileri temel alarak kısa animasyonlar veya videolar oluşturma kapasitesini ifade eder.
- Sesli Kullanım: Modelin, konuşma tanıma veya metinden ses üretme (text to speech-TTS) gibi ses tabanlı giriş ve çıkışları desteklemesini belirtir.
- Akıl Yürütme Modu: Model tarafından üretilen bir dizi ara doğal dil akıl yürütme adımları sonucunda girdi isteminden nihai çıktıya ulaşması olarak tanımlanır (Wei vd., 2022).
- Web Arama Modu: Gerçek zamanlı veya çevrimiçi kaynaklara erişerek sorgu sonuçlarını inceleyip bunları cevaba entegre edebilme özelliğini ifade eder.

İlgili kriterler baz alınarak hazırlanan karşılaştırma tabloları Ek-1 ve Ek-2'de yer almaktadır.



**Melik Başaran**

Gazi Üniversitesi Teknoloji Fakültesi Bilgisayar Mühendisliği Bölümü 3. sınıf öğrencisiyim. Gazi Üniversitesi Yapay Zeka Uygulama ve Araştırma Merkezi'nde kısmi zamanlı olarak çalışmaktayım. Derin öğrenme, bilgisayarlı görü ve görüntü işleme alanlarıyla ilgilenmekteyim.

### 3. ÜCRETLENDİRME MODELLERİ

#### 3.1. Token Bazlı Fiyatlandırma

Token, modelin işlediği kelime, karakter veya sembol birimi olarak kabul edilebilir ve genellikle verinin boyutunu ölçen birimdir. Ücretlendirme, genellikle girdi tokenları ve çıktı tokenları üzerinden yapılır. Yani, hem gönderdiğiniz veri (input) hem de modelin ürettiği sonuç (output) toplam token sayısını etkiler. Üretken YZ modeli, kullanıcı istemi başına fiyatlandırılır. Bu belirlenen fiyat, birden çok farklı görev ve değişken sayıda istem turu için kullanıcı maliyetini belirler; örneğin, GPT-4'ü matematiksel akıl yürütme veya kod oluşturma için kullanmanın maliyeti yalnızca token başına fiyata, istemlerin uzunluğuna ve sayısına bağlıdır. Bu sistem, kullanıcıların ne kadar veri gönderdiğine ve ne kadar işlem yapıldığına göre ödeme yapılmasını sağlar (Mahmood, 2024).

- Girdi (Input): API'ye gönderilen ve modelin işleyip yanıt üreteceği ham veri veya sorgu içeriğidir.
- Önbelleğe Alınmış Girdi (Cached Input): Daha önce işlenmiş ve saklanmış olan, tekrar kullanmak istendiğinde API'ye yeniden gönderilmeksizin hızlıca erişilen girdi verisidir.
- Çıktı (Output): Modelin işlenmiş verilere dayanarak ürettiği sonuç veya yanıttır; bu, metin, sayısal değer veya yapılandırılmış bilgi olabilir.



#### 3.2. Abonelik ve Kotalı Kullanım

Bazı hizmet sağlayıcılar, sabit aylık ücretler karşılığında sınırsız veya kotalı kullanım sunabilmektedir. Örnek paket türleri:

- Ücretsiz Plan (Free): Temel özelliklere erişim sunar; genellikle günlük veya aylık kullanım kotası içerir.
- Pro/Plus Üyelikler: Gelişmiş model erişimi, daha yüksek kullanım kotaları ve ek avantajlar gibi Premium özelliklere sahip olup sabit aylık ücret karşılığında sunulur.
- Takım (Team) Planı: Birden fazla kullanıcının ortak kaynakları paylaştığı, iş birliğine yönelik orta düzey bir üyelik türüdür.
- Kurumsal Paketler (Enterprise): Büyük ölçekli kurumlar için özelleştirilmiş fiyatlandırma, gelişmiş güvenlik ve yönetim araçları gibi kurumsal düzeyde ek hizmetleri içerir.

### 4. API VE CHAT KULLANIM MODELLERİ

#### 4.1. API Kullanım Modeli

API, ana yazılımın kaynak koduna entegre edildikten sonra ana yazılıma hizmet eden farklı uygulama parçaları arasındaki bir sözleşmedir. Bu uygulama parçaları, her ikisinin de anladığı bir dilde ve gerekirse bir ağ üzerinden birbirleriyle iletişim kurar (Jacobson, Brail & Woods, 2012). ChatGPT, Claude, Grok, Gemini, DeepSeek gibi modeller API üzerinden erişilebilir ve bu API'ler, etkileşimli ya da otomatik süreçler oluşturulmasına olanak tanır. Ek-2'de de görüldüğü gibi her model API'sinin yetkin olduğu alanlar ve token ücretlendirmeleri farklıdır. Bu yüzden kullanılacak API seçilirken bütçeye ve kullanılmak istenen en uygun alana göre seçim yapılmalıdır. Yaygın kullanım senaryoları şunlardır:

- Yazılım Entegrasyonları: Mobil ve web uygulamalarına AI destekli özellikler ekler. Örneğin, akıllı öneri sistemleri veya arama motorları gibi uygulamalar API ile entegre edilebilir.
- Otomatik Raporlama ve Veri Analitiği: AI modelleri, veri analizi ve raporlama süreçlerini otomatikleştirir, içgörü sağlar. Örneğin, pazar tahminleri veya finansal raporlar oluşturulabilir.
- Sohbet Botları ve Asistanlar: Müşteri destek hizmetleri için API tabanlı botlar kullanılır. Bu, hızlı yanıt sistemleri oluşturmak için yaygın olarak kullanılır.
- İçerik Üretimi ve Yaratıcı Yazılım Çözümleri: Yaratıcı içerik üretimi (metin, görsel, video) için AI API'leri kullanılır. Örneğin, otomatik blog yazıları veya görüntü oluşturma yapılabilir.
- Eğitim ve Öğrenme Yönetim Sistemleri: Kişiselleştirilmiş eğitim önerileri ve eğitim sohbet botları gibi AI çözümleri, eğitim alanında API'lerle entegre edilebilir.
- Sağlık Hizmetleri ve Tıbbi Veriler: Hasta verilerini analiz etmek, tanı koyma ve tedavi önerileri için AI API'leri kullanılabilir. Sağlık hizmetlerinde, AI verileri işleyebilir ve hızlı kararlar alabilir (Tarcan, Balçık & Sebik, 2024).
- Doğal Dil İşleme ve Dil Çevirisi: Dil çevirisi, duygu analizi, metin özetleme gibi doğal dil işleme görevleri API'ler aracılığıyla yapılabilir.

#### 4.2. Chat Kullanım Modeli

ChatGPT, Claude, Grok, Gemini, DeepSeek gibi modellerin chat tabanlı sürümleri, kullanıcıların herhangi bir kod yazmadan web üzerinden veya özel entegre ortamlarda hızlıca etkileşimde bulunmasına olanak tanır. Ek-1'de görüldüğü üzere her model farklı alanlarda yetkinleşmiştir ve üyelik ücretleri değişiklik göstermektedir. Yaygın kullanım senaryoları şunlardır:

- Web Tabanlı Chatbotlar: ChatGPT ve Claude gibi platformlar, tarayıcı üzerinden kolayca erişilebilen web tabanlı chatbotlar sunmaktadır. Kullanıcılar, herhangi bir kurulum yapmadan hızlıca etkileşime geçebilmektedir.
- Özel Sohbet Asistanları: Kurumsal ve sektörel ihtiyaçlara göre özelleştirilmiş chat asistanları; müşteri destek, satış ve eğitim gibi süreçlere uyum sağlar. Fine-tuning ile sektöre özel çözümler sunmaktadır.
- Sektörel Chatbotlar ve Entegrasyonlar: Chatbotlar, finans, sağlık, eğitim gibi sektörlerde kullanılır. Her sektörün özel ihtiyaçlarına göre terminoloji ve iş akışlarına uyum sağlamaktadır.
- E-ticaret ve Alışveriş Yardımcıları: E-ticaret satış sohbet botları, müşteri desteği sağlayarak, satışları artırarak ve müşteri hizmetlerini daha uygun maliyetli ve tatmin edici hale getirerek e-ticaret üzerinde önemli bir etkiye sahip olma potansiyeli taşımaktadır (Khan, 2020).



Chat ve API kullanım modelleri arasındaki temel farklar şunlardır:

	API Kullanımı	Chat Kullanımı
<b>Kurulum ve Entegrasyon</b>	Geliştirici, model API'sini kendi uygulamasına veya altyapısına entegre eder.	Kullanıcılar, web tabanlı veya özel sohbet arayüzlerinden doğrudan etkileşime geçer.
<b>Teknik Bilgi Gereksinimi</b>	Orta-yüksek düzeyde yazılım bilgisi (API çağrıları, anahtar yönetimi vb.) gerekir.	Son kullanıcı düzeyinde kullanım, ek bir yazılım bilgisi çoğunlukla gerekli değildir.
<b>Özelleştirme Düzeyi</b>	Yüksek özelleştirme imkânı (farklı veri kaynakları, kurumsal iş akışları).	Belirli sınırlar dahilinde modelin davranışı ayarlanabilir (prompt tasarımı vb.).
<b>Ölçekleme ve Hız</b>	Trafiğe ve kullanım senaryosuna göre API altyapısı ölçeklenebilir, yüksek hacimleri yönetebilir.	Kullanıcı bazlı bireysel oturumlar ağırlıklıdır, genellikle tek kullanıcı deneyimi temel alınır.
<b>Veri Gizliliği ve Güvenlik</b>	Model cevapları ve veriler kontrol edilebilir, özel sunucular veya ek güvenlik önlemleri uygulanabilir.	Hizmet sağlayıcının belirlediği veri işleme ve gizlilik politikalarına bağlı kalınır.
<b>Erişim Yönetimi</b>	API anahtarları veya OAuth benzeri kimlik doğrulama yöntemleriyle çoklu kullanıcı yönetimi.	Tipik olarak tekil kullanıcı hesapları, bazen kurumsal sohbet arayüzleri ve yönetim paneli.
<b>Geliştirme Süreçleri</b>	Sürekli entegrasyon, test, sürüm yönetimi; daha karmaşık projelerde büyük önem taşır.	Modelin arayüzünü sağlamak hizmet sağlayıcının sorumluluğundadır; kullanıcı tarafından ek geliştirme gerekmez.



## 5. SONUÇLAR

Ekonomi, politika, bilim ve eğitim dahil olmak üzere tüm sosyal sistemler, son zamanlarda yapay zekâ (YZ) teknolojisindeki hızlı gelişmelerden etkilenmiştir (Logg vd., 2019). Bu alanlarda üretken yapay zekâdan en verimli şekilde yararlanabilmek için ise doğru modelin seçilmesi, kullanım amacına uygun olarak eğitilmesi ve etik ilkeler doğrultusunda yönlendirilmesi büyük önem taşımaktadır.

Piyasadaki üretken yapay zekâ modellerinin seçimi; kullanım alanı, teknik gereksinimler ve bütçe gibi çeşitli faktörlere bağlı olarak değişiklik göstermektedir. API tabanlı çözümler, geliştiricilere esneklik, uyarlanabilirlik ve yüksek ölçeklenebilirlik sunarken; sohbet tabanlı çözümler (örneğin ChatGPT, Bard gibi) teknik bilgi gerektirmeyen kullanıcılar için kolay erişim, hızlı çıktı üretimi ve kullanıcı dostu bir deneyim sağlamaktadır.



Gelecekte üretken yapay zekânın, yaratıcı alanlarda (örneğin sanat, müzik, tasarım) ve araştırma-geliştirme süreçlerinde daha fazla rol oynaması; otonom sistemler, robotik ve IoT cihazlarının daha da yaygınlaşmasıyla, çok daha entegre bir biçimde günlük yaşamımıza girmesi beklenmektedir. Ancak bu gelişmelerin birlikte getireceği etik ve sosyal sorunlar, düzenleyici çerçevelerin oluşturulmasını gerektirecek (Corrêa, De Oliveira & Massmann, 2022; Fan, 2024) ve yapay zekâ teknolojilerinin geleceğinin büyük ölçüde bu denetimlerin etkinliğine bağlı olacağı düşünülmektedir.

## 6. KAYNAKÇA

- Akıllı, E., & Şimşek, M. (2024). Dijital Diplomaside Büyük Dil Modelleri: Fırsatlar ve Riskler. İnsan ve Toplum, 14(4), 1-23.
- Jacobson, D., Brail, G., & Woods, D. (2012). APIs: A strategy guide. O'Reilly Media, Inc.
- Ezgi, Ş. E. N., & Atiker, B. (2020). Grafik Tasarım Uygulamalarında Yeni Bir Aktör: Yapay Zekâ. International Journal of Social and Humanities Sciences Research (JSHSR), 7(63), 3946-3957.
- Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2024). Generative AI. Business & Information Systems Engineering, 66 (1), 111-126.
- Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. Advances in neural information processing systems, 35, 24824-24837.
- Mahmood, R. (2024). Pricing and competition for generative ai. Advances in Neural Information Processing Systems, 37, 75727-75748.
- Tarcan, G. Y., Balçık, P. Y., & Sebik, N. B. (2024). Türkiye ve dünyada sağlık hizmetlerinde yapay zekâ. Mersin Üniversitesi Tıp Fakültesi Lokman Hekim Tıp Tarihi ve Folklorik Tıp Dergisi, 14(1), 50-60.
- Logg, J. M., Minson, J. A. ve Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. Organizational Behavior and Human Decision Processes, 151, 90-103.
- Bozkurt, A. (2023). ChatGPT üretken yapay zeka ve algoritmik paradigma değişikliği. Alanyazın, 4(1), 63-72.
- Khan, M. M. (2020, December). Development of an e-commerce sales chatbot. In 2020 IEEE 17th International Conference on Smart Communities: Improving Quality of Life Using ICT, IoT and AI (HONET) (pp. 173-176). IEEE.
- Fan, L. (2024). Artificial intelligence ethics: a dialogue between technological advances and human values. International Journal of Education and Humanities, 14(2), 260-265.
- Corrêa, N. K., De Oliveira, N., & Massmann, D. (2022). On the Efficiency of Ethics as a Governing Tool for Artificial Intelligence. arXiv preprint arXiv:2210.15289.



# Gazi Üniversitesi

## Yapay Zeka Araştırma ve Uygulama Merkezi

# AI TECH

Çizelge 1. Üretken Yapay Zekâ Modelleri Chat Kullanım Modellerinin Karşılaştırması

	Model Versiyonu	Görsel ve Metin Tabanlı Dosya İnceleme	Görsel Üretme	Vide Üretme	Sesli Kullanım	Akıllı Yürütme Modu	Web Arama Modu	Ücret (Aylık)	Yetkin Olduğu Konular
ChatGPT Free	GPT-4 (Limitli) GPT-4o Mini (Limitsiz)	Var (1 günde 2 adet)	Yok	Yok	Var (Limitli)	Var	Var	Free	Gelişmiş metin üretimi ve anlama yetenekleri, yaratıcı içerik oluşturma, custom GPT'ler sayesinde özelleştirme, Sora ve DALL-E sayesinde görsel ve video üretimi
ChatGPT Plus	GPT-4o (3 saatte 80 mesaj) GPT-4 (3 saatte 40 mesaj) GPT-4o mini	Var	Var (Dall-E)	Var (720p - 5sn - 50 adet) (Sora)	Var	Var	Var	\$20	
ChatGPT Pro	GPT-4 GPT-4o GPT-4.5 (Araştırma)	Var	Var (Dall-E)	Var (1080p - 20sn - 500 adet) (Sora)	Var	Var	Var	\$200	
Gemini Free	Gemini 2.0 Gemini 2.0 Flash Thinking Experimental	Var	Var (Imagen 3)	Yok	Sadece sesli girdi	Var (Gemini 2.0 Flash Thinking Experimental)	Var	Free	Google hizmetleriyle entegrasyon, kullanıcının Google verilerini kullanarak kişiselleştirilmiş deneyim; görsel, video ve doküman incelemesi, veri analizi ve veri işleme
Gemini Advanced	Gemini 2.0 Gemini 2.0 Flash Thinking Experimental Gemini Ultra	Var	Var (Imagen 3)	Yok	Sadece sesli girdi	Var (Gemini 2.0 Flash Thinking Experimental)	Var	\$20 (Google One Premium paketi içerisinde)	Ücretsiz pakete ek olarak Gemini Ultra modeline erişim sağlama, limitlerin artırılması
Claude Free	Claude 3.7 Sonnet (1 günde 40 mesaj)	Var	Yok	Yok	Yok	Var (Claude 3.7 Sonnet)	Yok	Free	Kodlama, hata ayıklama, doğal dilde sohbet, web geliştirme
Claude Pro	Claude 3.7 Sonnet Claude 3.5 Sonnet Claude 3.5 Haiku Claude 3 Opus	Var	Yok	Yok	Yok	Var (Claude 3.7 Sonnet)	Yok	\$15	
Microsoft Copilot Free	Microsoft Prometheus	Var	Var (1 günde 15 adet)	Yok	Sadece sesli girdi	Var	Var	Free	Kullanıcın Microsoft verilerini kullanarak kişiselleştirilmiş deneyim, DALL-E 3 ile görsel üretme
Microsoft Copilot Pro	Microsoft Prometheus	Var	Var (1 günde 100 adet)	Yok	Sadece sesli girdi	Var	Var	\$20	Ücretsiz pakete ek olarak Microsoft 365 uygulamaları ile entegre çalışma, limitlerin artırılması
DeepSeek	DeepSeek V3	Sadece metin tabanlı dokümanlar	Yok	Yok	Yok	Var (DeepThink R1)	Var	Free	Açık kaynak, geliştirilmeye açık
DeepAI Free	GPT-3.5	Yok	Var (Limitli)	Var (Limitli)	Yok	Var (Genius)	Yok	Free	Yaratıcılık gerektiren işlerde iyi; görsel, metin ve video oluşturmada başarılı
DeepAI Pro	GPT-3.5	Yok	Var (1 ayda 500 adet)	Var (1 ayda 30 adet)	Yok	Var (Genius)	Yok	\$5	
Grok Free	Grok 3 (sınırlı yoğunluğuna bağlı limitli kullanım) Grok 2 (sınırlı yoğunluğuna bağlı limitli kullanım)	Var	Var (Aurora) (Sınırlı)	Yok	Yok	Var (Grok 3)	Var	Free	X platformuna direkt bağlı olduğundan dolayı gündemi yakından takip edebilme
SuperGrok	Grok 3 (2 saatte 50 sorgu) Grok 3 mini Grok 2 (1 saatte 30 sorgu)	Var	Var (Aurora)	Yok	Yok	Var (Grok 3)	Var	\$30	



# Gazi Üniversitesi

## Yapay Zeka Araştırma ve Uygulama Merkezi

# AI TECH

Çizelge 2. Üretken Yapay Zekâ Modelleri API Kullanım Modellerinin Karşılaştırması

	Price per 1M tokens			Yetkin Olduğu Konular	
	Model Name	Input	Cached input	Output	
ChatGPT	gpt-4.5-preview	\$75,00	\$37,50	\$150,00	OpenAI'nin en son geliştirdiği dil modelidir. Gelişmiş metin ve görsel görevlerde yüksek performans sunar.
	gpt-4o	\$2,50	\$1,25	\$10,00	Metin, görsel ve sesli verileri işleyebilen çok modlu bir yapay zeka modelidir.
	gpt-4o-audio-preview	\$2,50	-	\$10,00	Sesli etkileşimler ve sesli veri işleme konusunda uzmanlaşmış bir modeldir.
	gpt-4o-realtime-preview	\$5,00	\$2,50	\$20,00	Düşük gecikmeli, gerçek zamanlı sesli etkileşimler için optimize edilmiş bir modeldir.
	gpt-4o-mini	\$0,15	\$0,08	\$0,60	GPT-4o'nun daha küçük ve hızlı bir versiyonudur, metin ve sesli verileri işleme kapasitesine sahiptir.
	gpt-4o-mini-audio-preview	\$0,15	-	\$0,60	Küçük boyutlu ve hızlı bir sesli etkileşim modelidir.
	gpt-4o-mini-realtime-preview	\$0,60	\$0,30	\$2,40	Gerçek zamanlı sesli etkileşimler için optimize edilmiş küçük boyutlu bir modeldir.
	o1	\$15,00	\$7,50	\$60,00	Derin düşünme ve karmaşık problem çözme yeteneklerine sahip bir yapay zeka modelidir.
	o3-mini	\$1,10	\$0,55	\$4,40	o3 modelinin daha hızlı ve hafif bir versiyonudur, gelişmiş akıl yürütme yeteneklerine sahiptir.
	o1-mini	\$1,10	\$0,55	\$4,40	o1 modelinin daha küçük ve hızlı bir versiyonudur, temel görevlerde etkili performans sunar.
Gemini	DALL-E 3	Quality	1024x1024	1024x1792	Metin açıklamalarından yüksek kaliteli görseller oluşturabilen bir yapay zeka modelidir.
		Standart	\$0,04	\$0,08	
		HD	\$0,08	\$0,12	
Gemini	Gemini 2.0 Flash	\$0.10 (text / image / video) \$0.70 (audio)	\$0.025 / 1,000,000 tokens (text/image/video) \$0.175 / 1,000,000 tokens (audio)	\$0,40	Düşük gecikme süresi ve yüksek performansıyla, ajan tabanlı deneyimleri destekleyen güçlü bir modeldir.
	Gemini 2.0 Flash-Lite	\$0.075	Available March 31, 2025	\$0,30	Maliyet etkinliğiyle dikkat çeken bu model, daha hafif uygulamalar için tasarlanmıştır.
	Gemini 1.5 Flash	\$0.075, prompts <= 128k tokens \$0.15, prompts > 128k tokens	\$0.30, prompts <= 128k tokens \$0.60, prompts > 128k tokens	\$0.01875, prompts <= 128k tokens \$0.0375, prompts > 128k tokens	Geliştirilmiş hız ve performans sunan bu model, hızlı yanıt gerektiren uygulamalar için idealdir.
	Gemini 1.5 Flash-8B	\$0.0375, prompts <= 128k tokens \$0.075, prompts > 128k tokens	\$0.15, prompts <= 128k tokens \$0.30, prompts > 128k tokens	\$0.01, prompts <= 128k tokens \$0.02, prompts > 128k tokens	8 milyar parametreye sahip, daha kompakt ve verimli çalışması için tasarlanmış,Gemini 1.5 Flash modeli varyasyonudur.
	Gemini 1.5 Pro	\$1.25, prompts <= 128k tokens \$2.50, prompts > 128k tokens	\$5.00, prompts <= 128k tokens \$10.00, prompts > 128k tokens	\$0.3125, prompts <= 128k tokens \$0.625, prompts > 128k tokens	Çeşitli görevlerde yüksek performans sunan bu model, geniş bir uygulama yelpazesinde kullanılabilir.
Claude	Claude 3.7 Sonnet	\$3	Prompt caching write	\$15	Gelişmiş mantıksal akıl yürütme ve karmaşık problem çözme yetenekleriyle tanınır.
	Claude 3.5 Haiku	\$0,80	Prompt caching write	\$4	Hızlı yanıt süreleriyle genel amaçlı kullanım için tasarlanmıştır.
	Claude 3.5 Sonnet	\$3	Prompt caching write	\$15	Kodlama, çok adımlı iş akışları ve görsel veri işleme konularında üstün performans sergiler.
	Claude 3 Opus	\$15	Prompt caching write	\$75	En gelişmiş model olup, geniş bilgi tabanı ve yüksek işlem kapasitesiyle karmaşık görevleri yerine getirebilir.
	Claude 3 Haiku	\$0,08	Prompt caching write	\$1,25	Daha küçük ve hızlı bir model olup, temel görevlerde verimli performans sunar.
DeepSeek	deepseek-chat	\$0,07	\$0,27	\$1,10	Genel sohbet uygulamaları için tasarlanmış olup, kullanıcılarla etkileşimli ve anlamlı diyaloglar kurma yeteneğine sahiptir.
	deepseek-reasoner	\$0,14	\$0,55	\$2,19	Özellikle matematiksel ve programlama görevlerinde adım adım düşünme yaklaşımını kullanarak yüksek doğruluklu sonuçlar elde eder.
DeepAI	DeepAI Free	100 çağrı başına \$5			Yapay zekâ destekli araçlar sunarak metin, görsel ve video oluşturma, düzenleme ve analiz etme konularında hizmet vermektedir.
	DeepAI Pro	500 çağrı başına \$5			
Grok	grok-2-vision-1212	\$2,00	-	\$10,00	Geliştirilmiş doğruluk, talimat takibi ve çok dilli destek sunar.
	grok-2-1212	\$2,00	-	\$10,00	Görsel içerik analizi ve üretimi konusunda yetenekli olup, çok modlu uygulamalar için idealdir.
	grok-vision-beta	\$5,00	-	\$15,00	Görsel anlayış ve üretim kapasitesini test etmek amacıyla kullanılan beta sürümüdür.
	grok-beta	\$5,00	-	\$15,00	Grok beta sürümüdür.