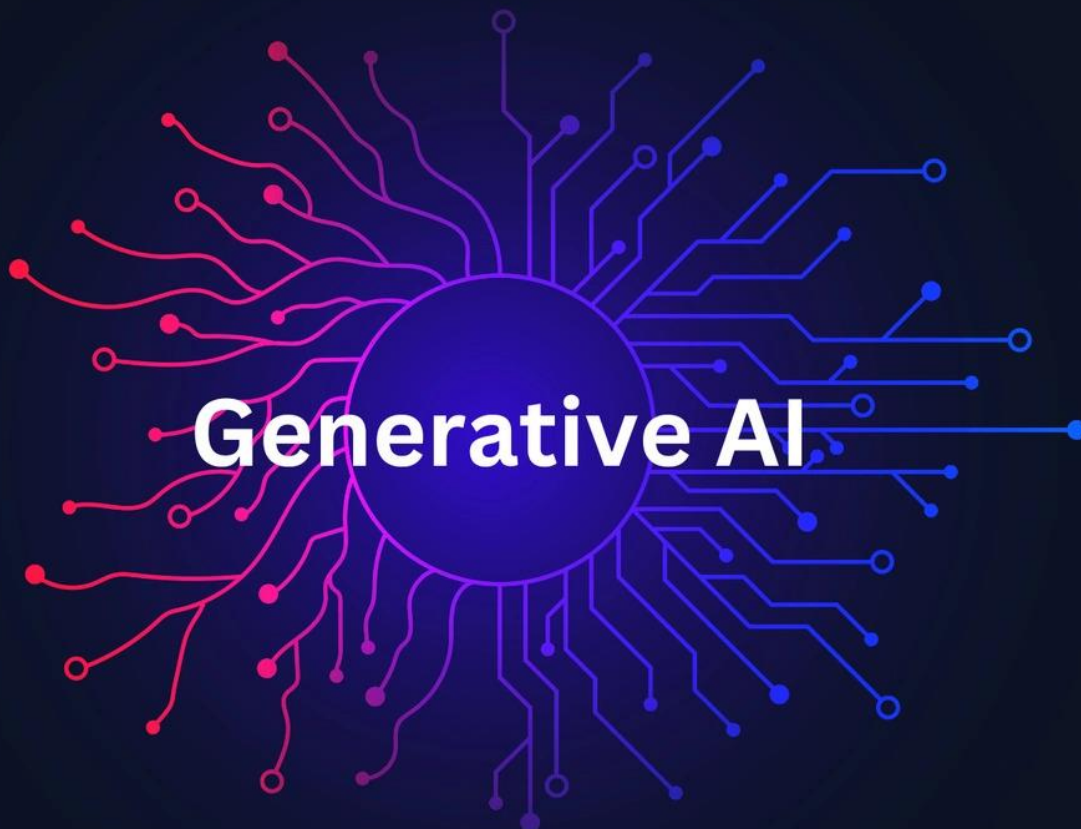


باسمه تعالی

ابزارهای هوش مصنوعی مولد



نسخه: ۱,۱,۰

دی ماه ۱۴۰۳

ارائه شده توسط: mmhosseini.ir

در کلیه مراحل تهیه، ترجمه و بازنویسی این جزوه از هوش مصنوعی بهره گرفته شده است.

فهرست مطالب

مقدمه	۶
تعریف هوش مصنوعی مولد	۷
آشنایی با ابزارهای متنوع هوش مصنوعی مولد	۹
دسته بندی ابزارهای هوش مصنوعی مولد	۹
۱. دسته بندی بر اساس نوع خروجی	۹
۲. دسته بندی بر اساس کاربرد	۱۰
معرفی ابزارهای کلیدی هوش مصنوعی مولد	۱۲
ChatGPT، تولید متن، پاسخ دهی هوشمند	۱۲
موارد استفاده از ChatGPT	۱۲
مزایا، چالش ها و محدودیتهای کار با ChatGPT	۱۳
معرفی چت جی پی تی	۱۶
راه اندازی حساب کاربری: ورود به دنیای چت جی پی تی	۱۶
مقایسه چت جی پی تی با موتورهای جستجو و ابزارهای تحلیل	۱۶
رقابت با چت جی پی تی در اشتغال: تهدید یا فرصت؟	۱۷
تعریف مجدد چت بات ها با ظهور چت جی پی تی و چت جی پی تی پلاس	۱۷
قرار گرفتن در صدر اخبار و مختل کردن کسب و کارها	۱۸
تفاوت چت جی پی تی در چیست؟ تمایز از سایر ابزارها	۱۹
Copilot، کمک به برنامه نویسی و کدنویسی	۲۱
موارد استفاده از copilot	۲۱
مزایا و چالش ها و محدودیتهای کار با کوپایلوت	۲۲
بررسی علمی و ساده ی هوش مصنوعی کوپایلوت	۲۳
مزایای کوپایلوت نسبت به سایر چت بات های هوش مصنوعی	۲۴
راه های دسترسی به کوپایلوت	۲۷
مقایسه طرح های رایگان و اشتراکی	۲۷
Gemini، جستجوی پیشرفته و ترجمه	۲۸
موارد استفاده از Gemini	۲۸
مزایا و محدودیتهای کار با جمینای	۲۹
گوگل جمینای چیست؟	۲۹
ویژگی های هوش مصنوعی گوگل جمینای	۳۰
قابلیت های گوگل جمینای	۳۱
نحوه ی عملکرد گوگل جمینای	۳۲
تفاوت گوگل جمینای با موتور جستجوی گوگل	۳۲

۳۳	ساخت تصویر با گوگل جمینای
۳۳	مقایسه‌ی جمینای با GPT-۴ و Claude
۳۵	Claude، دستیار نوشتاری و خلاقیت
۳۶	موارد استفاده از کلود
۳۷	مزایای کار با کلود
۳۷	محدودیت‌های کار با کلود
۳۸	معرفی جامع هوش مصنوعی Claude
۳۸	اهداف و ویژگی‌های کلیدی کلود
۳۸	قابلیت‌های کلود
۳۹	تفاوت‌های کلود با سایر مدل‌های زبانی بزرگ
۳۹	نسخه‌های کلود
۴۰	دسترسی به کلود
۴۱	DeepSeek، پدیده چینی تازه وارد
۴۱	ویژگی‌های برجسته دیپ‌سیک
۴۱	مزایای دیپ‌سیک
۴۲	محدودیت‌های دیپ‌سیک
۴۲	کاربردهای اصلی دیپ‌سیک
۴۳	دسترسی به هوش مصنوعی DeepSeek
۴۳	بررسی مقایسه‌ای ابزارها
۴۶	معرفی ابزارهای تولید تصویر هوش مصنوعی
۴۶	DALL-E، تولید تصویر از متن
۴۶	معرفی جامع هوش مصنوعی مولد DALL-E
۴۶	نحوه عملکرد DALL-E
۴۷	قابلیت‌های کلیدی DALL-E
۴۷	کاربردهای DALL-E
۴۸	مزایای استفاده از DALL-E
۴۸	محدودیت‌ها و چالش‌ها
۴۸	دسترسی به DALL-E
۴۹	MidJourney، طراحی تصاویر هنری
۴۹	موارد استفاده از MidJourney
۴۹	مزایا و محدودیت‌های کار با میدجرنی
۵۰	میدجرنی: خلق تصاویر شگفت‌انگیز با هوش مصنوعی
۵۰	فرایند ساخت تصویر در میدجرنی

۵۱	قدرت و تأثیر میدجربی.....
۵۱	نسخه‌های مختلف میدجربی و تغییرات آن‌ها.....
۵۲	Firefly، تولید تصویر از متن و فراتر از آن.....
۵۳	تاریخچه و نسخه‌های Firefly.....
۵۳	نحوه عملکرد Firefly.....
۵۴	قابلیت‌های کلیدی Firefly.....
۵۴	کاربردهای Firefly.....
۵۵	مزایای استفاده از Firefly.....
۵۵	محدودیت‌ها و چالش‌ها.....
۵۵	دسترسی به Firefly.....
۵۶	مقایسه بین ابزارهای هوش مصنوعی مولد تصویر.....
۵۷	Imagen۳، تولید کننده تصویر جمینای.....
۵۷	تاریخچه و نسخه‌های Imagen۳.....
۵۷	نحوه عملکرد Imagen۳.....
۵۸	قابلیت‌های کلیدی Imagen۳.....
۵۸	کاربردهای Imagen۳.....
۵۸	مزایای استفاده از Imagen۳.....
۵۹	محدودیت‌ها و چالش‌ها.....
۵۹	دسترسی به Imagen۳.....
۵۹	Stable Diffusion، تولید تصویر فراتر از تصور.....
۵۹	تاریخچه و نسخه‌های Stable Diffusion.....
۶۰	نحوه عملکرد Stable Diffusion.....
۶۰	قابلیت‌های کلیدی Stable Diffusion.....
۶۱	کاربردهای Stable Diffusion.....
۶۱	مزایای استفاده از Stable Diffusion.....
۶۱	محدودیت‌ها و چالش‌ها.....
۶۲	دسترسی به Stable Diffusion.....
۶۳	مقایسه چت‌بات‌ها، موتورهای جستجوی سنتی و موتورهای جستجوی هوش مصنوعی.....
۶۳	۱. موتورهای جستجوی سنتی (مانند گوگل و بینگ).....
۶۳	۲. چت‌بات‌ها مانند Gemini، ChatGPT.....
۶۴	۳. موتورهای جستجوی هوش مصنوعی مانند Perplexity AI.....
۶۵	تفاوت در تعامل.....
۶۵	کاربردهای تخصصی و عمومی.....

محدودیت‌ها و چالش‌های استفاده از ابزارهای هوش مصنوعی مولد.....	۶۶
۱. محدودیت‌های مربوط به کیفیت و صحت خروجی.....	۶۶
۲. محدودیت‌های مربوط به کنترل و شخصی‌سازی.....	۶۶
۳. چالش‌های مربوط به حریم خصوصی و امنیت.....	۶۷
۴. چالش‌های مربوط به دسترسی و هزینه.....	۶۷
۵. چالش‌های مربوط به اخلاق و مسئولیت‌پذیری.....	۶۸
نکاتی برای کاربران:.....	۶۸



مقدمه

یکی از شگفت‌انگیزترین دست‌یافته‌های بشر، هوش مصنوعی مولد است. فناوری که بسیار جالب و جذاب است و دانستن در مورد آن می‌تواند خیلی هیجان‌انگیز باشد. اما پیش از ورود به بحث به این مثال‌ها توجه کنید:

تصور کنید نویسنده‌ای سال‌ها درگیر نوشتن رمانی بوده، اما به بن‌بست رسیده. ناگهان با ابزاری آشنا می‌شود که می‌تواند ایده‌ها را به متن تبدیل کند، شخصیت‌ها را پرورش دهد و حتی پایان‌های مختلف را پیشنهاد دهد. این ابزار نه یک همکار نویسنده، بلکه یک هوش مصنوعی مولد است.

آیا این یک تقلب است یا انقلابی در دنیای نویسندگی؟

یا دانشجویی را تصور کنید که برای انجام یک تحقیق سنگین، با انبوهی از مقالات و داده‌ها روبروست. او به جای صرف ساعت‌ها برای جستجو و خلاصه‌برداری، از یک ابزار هوش مصنوعی مولد استفاده می‌کند که می‌تواند اطلاعات مرتبط را جمع‌آوری، دسته‌بندی و حتی خلاصه‌ای از آن‌ها ارائه دهد.

آیا این به معنای کاهش تلاش (تنبل شدن) دانشجویست یا افزایش بهره‌وری او؟

حتماً برای شما هم پیش آمده که پس از جستجوی یک محصول در اینترنت، تبلیغات مرتبط با آن را در سایت‌های مختلف ببینید. این نمونه‌ای ساده از کاربرد هوش مصنوعی مولد در بازاریابی است. اما تصور کنید این تبلیغات نه فقط مرتبط با محصول، بلکه کاملاً شخصی‌سازی‌شده و متناسب با سلیقه و نیازهای شما باشد. هوش مصنوعی مولد این امکان را فراهم می‌کند.

آهنگ‌های محبوبی را که این روزها می‌شنوید، در نظر بگیرید. آیا می‌دانستید که در آینده‌ای نزدیک بسیاری از آن‌ها با کمک هوش مصنوعی مولد ساخته خواهد شد؟ این ابزارها می‌توانند ملودی، ریتم و حتی شعر آهنگ را تولید کنند.

آیا این به معنای پایان خلاقیت هنرمندان است یا آغاز دوره‌ای جدید در موسیقی؟

چند سؤال:

- آیا هوش مصنوعی مولد جایگزین انسان‌ها خواهد شد؟
- آیا هوش مصنوعی مولد می‌تواند واقعاً خلاق باشد؟
- آیا استفاده از هوش مصنوعی مولد در آموزش، تقلب محسوب می‌شود؟

تعریف هوش مصنوعی مولد

هوش مصنوعی مولد (Generative AI) نوعی از هوش مصنوعی است که می‌تواند محتوای جدید و بدیع تولید کند.

برای تشریح بیشتر و درک بهتر مؤلفه‌های این تعریف، می‌توان آن را به شکل زیر گسترش داد:

هوش مصنوعی مولد، با استفاده از الگوریتم‌های یادگیری ماشین، از داده‌های موجود الگوها و ساختارها را یاد می‌گیرد و سپس از این دانش برای خلق محتوای جدید استفاده می‌کند. این محتوا می‌تواند شامل موارد زیر باشد:

- **متن:** مانند نوشتن داستان، شعر، مقاله، کد برنامه‌نویسی، ایمیل و غیره.
- **تصویر:** مانند خلق تصاویر واقعی، نقاشی، طرح‌های گرافیکی و غیره.
- **صوت:** مانند ساخت موسیقی، تولید صداهای مختلف و غیره.
- **ویدئو:** مانند ساخت ویدئوهای کوتاه، انیمیشن و غیره.
- **کد:** تولید کد برای زبان‌های برنامه‌نویسی مختلف

برای درک بهتر این تعریف، به نکات زیر توجه کنید:

- **یادگیری از داده‌ها:** هوش مصنوعی مولد ابتدا با تجزیه و تحلیل حجم زیادی از داده‌ها (مانند متن‌ها، تصاویر، صداها و غیره) الگوها و روابط موجود در آن‌ها را یاد می‌گیرد. این فرایند شبیه به یادگیری انسان از طریق تجربه و مشاهده است.
- **خلق محتوای جدید:** پس از یادگیری، هوش مصنوعی مولد می‌تواند محتوایی تولید کند که قبلاً وجود نداشته است. این محتوا می‌تواند کاملاً جدید و بدیع بوده یا ترکیبی از عناصر موجود باشد.
- **شباهت به داده‌های آموزشی:** محتوای تولید شده توسط هوش مصنوعی مولد معمولاً شباهت‌هایی به داده‌های آموزشی دارد، اما دقیقاً کپی آن‌ها نیست. به عبارت دیگر، هوش مصنوعی مولد الگوها را یاد می‌گیرد و نه خود داده‌ها را.

- **کاربردهای متنوع:** هوش مصنوعی مولد در زمینه‌های مختلفی (تقریباً همه ابعاد زندگی انسان) مانند هنر، طراحی، نویسندگی، برنامه‌نویسی، تبلیغات، پزشکی و غیره کاربرد دارد.

مثالی ساده:

تصور کنید به یک کودک تعداد زیادی عکس از گربه‌ها نشان می‌دهید. کودک پس از مدتی می‌تواند تشخیص دهد که یک حیوان، گربه است یا خیر. هوش مصنوعی مولد نیز به همین شکل عمل می‌کند. با دیدن تعداد زیادی عکس از گربه‌ها، یاد می‌گیرد که گربه چه شکلی است و سپس می‌تواند تصاویری جدید از گربه‌ها خلق کند که قبلاً ندیده است.



آشنایی با ابزارهای متنوع هوش مصنوعی مولد

ابزارهای هوش مصنوعی مولد در سال‌های اخیر گستردگی بسیار زیادی پیدا نموده‌اند و روزی نیست که ابزاری جدید با امکانات و ویژگی‌های شگفتی آفرین عرضه نشود. با این حال آشنایی اصولی با انواع این ابزارها، دسته‌بندی‌ها و کارکردهای آن‌ها در استفاده مؤثر از این ابزارها بسیار مفید خواهد بود.

دسته‌بندی ابزارهای هوش مصنوعی مولد

هدف از این بخش، آشنا کردن شرکت‌کنندگان با انواع مختلف ابزارهای هوش مصنوعی مولد بر اساس نوع خروجی و کاربرد آن‌هاست. برای این منظور، ابزارها را به دو شیوه دسته‌بندی می‌کنیم.

۱. دسته‌بندی بر اساس نوع خروجی

در این دسته‌بندی، ابزارهای هوش مصنوعی مولد بر اساس نوع محتوایی که تولید می‌کنند، گروه‌بندی می‌شوند. این دسته‌بندی به درک بهتر کاربرد هر ابزار کمک می‌کند.

• ابزارهای تولید متن: این ابزارها قادر به تولید انواع مختلف متن مانند داستان،

مقاله، شعر، ایمیل، خلاصه متن، ترجمه و غیره هستند.

◦ **مثال‌ها:** ChatGPT، Gemini، Claude، Jasper.ai

◦ **کاربردها:** نوشتن محتوای وب، تولید گزارش، خلاصه نویسی، پاسخ به

سؤالات، ترجمه متون، تولید محتوای خلاقانه و ...

• ابزارهای تولید تصویر: این ابزارها می‌توانند تصاویر واقعی، نقاشی، طرح‌های

گرافیکی، تصاویر سه‌بعدی و غیره را تولید کنند.

◦ **مثال‌ها:** DALL-E ۲، Midjourney، Stable Diffusion، Bing Image Creator.

◦ **کاربردها:** طراحی لوگو، تولید تصاویر تبلیغاتی، خلق آثار هنری، طراحی

بازی‌های ویدئویی، ویرایش تصاویر و ...

- **ابزارهای تولید صوت:** این ابزارها قادر به تولید صدا، موسیقی، گفتار و غیره هستند.

- **مثال‌ها:** Jukebox (تولید موسیقی)، Descript (ویرایش صدا و تولید گفتار).
- **کاربردها:** ساخت موسیقی، تولید پادکست، دوبله فیلم، تولید جلوه‌های صوتی و ...

- **ابزارهای تولید ویدئو:** این ابزارها می‌توانند ویدئوهای کوتاه، انیمیشن و غیره را تولید کنند. این دسته هنوز در حال توسعه است اما پتانسیل بالایی دارد.

- **مثال‌ها:** Runway ML، Synthesia.
- **کاربردها:** تولید ویدئوهای تبلیغاتی، ساخت انیمیشن، تولید محتوای آموزشی و ...

- **ابزارهای تولید کد:** این ابزارها قادر به تولید کد برای زبان‌های برنامه‌نویسی مختلف هستند.

- **مثال‌ها:** GitHub Copilot، Tabnine.
- **کاربردها:** کمک به برنامه‌نویسان در نوشتن کد، تولید کد خودکار، تکمیل کد.

۲. دسته‌بندی بر اساس کاربرد

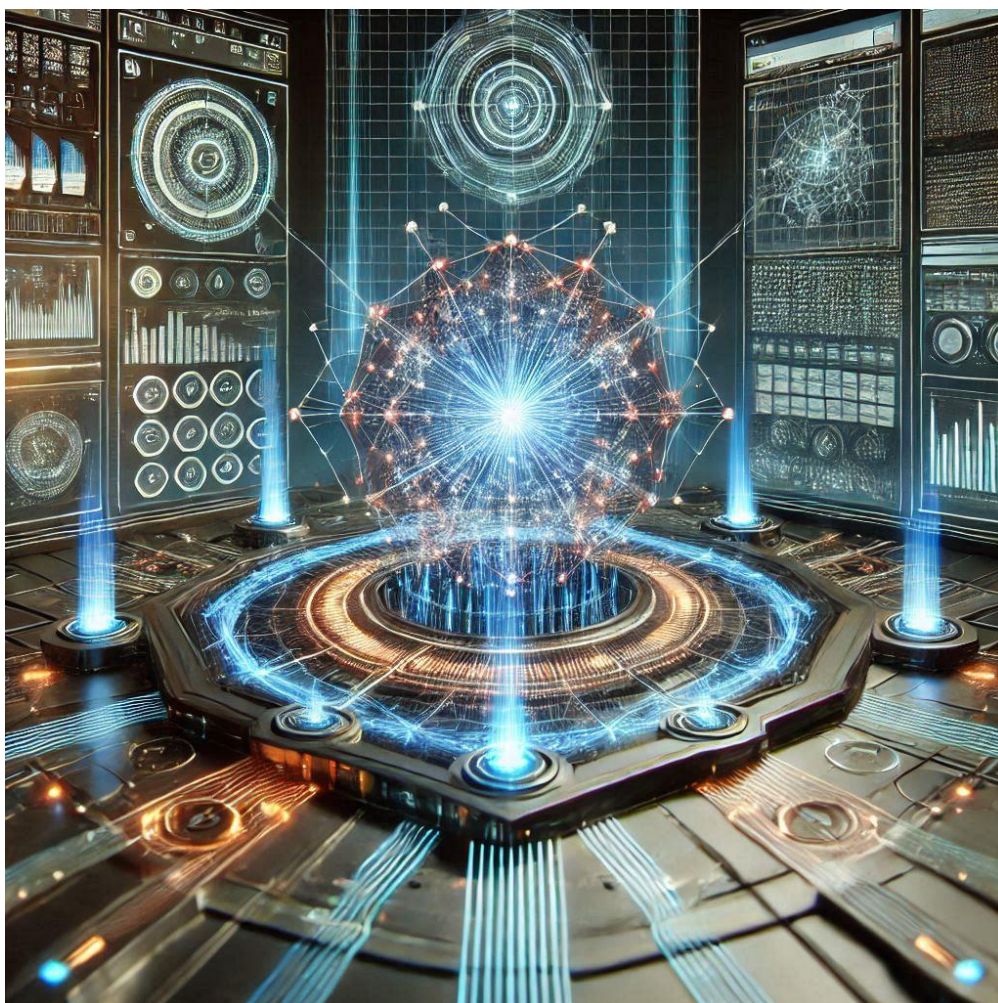
در این دسته‌بندی، ابزارها بر اساس نوع کاربردی که دارند، گروه‌بندی می‌شوند. این دسته‌بندی به کاربران کمک می‌کند تا ابزار مناسب برای نیاز خود را پیدا کنند. به عنوان مثال دسته بندی ابزارها به این صورت انجام می پذیرد.

- **ابزارهای خلاقیت و هنر:** این ابزارها برای تولید محتوای خلاقانه مانند داستان، شعر، موسیقی، نقاشی و غیره استفاده می‌شوند.
- **مثال‌ها:** DALL-E ۲، Midjourney، ChatGPT (برای تولید داستان و شعر).

- **ابزارهای بهره‌وری و کسب‌وکار:** این ابزارها برای افزایش بهره‌وری در کارها، تولید گزارش، پاسخ به سؤالات، خلاصه‌سازی متون و غیره استفاده می‌شوند.
 ○ **مثال‌ها:** Claude، Gemini، ChatGPT

- **ابزارهای آموزشی و پژوهشی:** این ابزارها برای کمک به یادگیری، انجام تحقیقات، جستجوی اطلاعات و غیره استفاده می‌شوند.
 ○ **مثال‌ها:** Wolfram Alpha (موتور دانش محاسباتی)، Consensus (موتور جستجوی مقالات علمی).

- **ابزارهای برنامه‌نویسی:** این ابزارها برای کمک به برنامه‌نویسان در نوشتن کد، تولید کد خودکار و غیره استفاده می‌شوند.
 ○ **مثال‌ها:** Tabnine، GitHub Copilot



معرفی ابزارهای کلیدی هوش مصنوعی مولد

ChatGPT، تولید متن، پاسخ‌دهی هوشمند

شاید بتوان چت جی‌پی‌تی را یکی از بهترین ابزارها برای تولید متن و پاسخ‌دهی هوشمند دانست. در ادامه با این ابزار هوش مصنوعی مولد بیشتر آشنا می‌شویم.

موارد استفاده از ChatGPT

۱. تولید متن و محتوا

- ✓ نوشتن خلاقانه: داستان، شعر، فیلمنامه، نمایشنامه، طنز، و غیره.
- ✓ تولید محتوای بازاریابی: متن تبلیغات، توضیحات محصول، پست‌های شبکه‌های اجتماعی، ایمیل‌های تبلیغاتی، شعار تبلیغاتی و غیره.
- ✓ تولید محتوای فنی: گزارش‌های فنی، مستندات، مقالات علمی، راهنماهای کاربری و غیره.
- ✓ نوشتن انواع متن اداری: نامه، ایمیل، گزارش، مقاله، رزومه و غیره.
- ✓ خلاصه‌سازی متن: خلاصه‌کردن متون طولانی به متون کوتاه و مفید.
- ✓ بازنویسی متن: بازنویسی متون با لحن و سبک متفاوت (رسمی، علمی، خودمانی و ...).
- ✓ تولید ایده: تولید ایده‌های جدید برای پروژه‌ها، کسب‌وکارها، موضوعات پژوهشی جدید، تولید ایده محتوای خلاقانه و غیره.
- ✓ تولید محتوا برای وبسایت‌ها و وبلاگ‌ها: نوشتن پست‌های وبلاگ، مقالات، توضیحات محصول و غیره.

۲. ارتباط و تعامل

- ✓ پاسخ به سوالات: پاسخ به سوالات کاربران در زمینه‌های مختلف.
- ✓ گفتگوی دوستانه: گفتگوی غیررسمی و دوستانه (گپ زدن).
- ✓ ارائه پشتیبانی مشتری: پاسخ به سوالات مشتریان، حل مشکلات و ارائه راهنمایی.
- ✓ آموزش و یادگیری: توضیح مفاهیم پیچیده، ارائه مثال‌های کاربردی، پاسخ به سوالات دانشجویان، آماده کردن سوالات امتحانی، تدوین برنامه درسی، ارائه درس و غیره.

- ✓ بازی و سرگرمی: انجام بازی‌های متنی، تعریف معما و غیره.
- ✓ شبیه‌سازی شخصیت: تعریف شخصیت‌های فرضی (پرسونا)، شبیه‌سازی گفتگو با شخصیت‌های تاریخی، خیالی و غیره.

۳. برنامه‌نویسی و توسعه نرم‌افزار

- ✓ تولید کد: تولید کد در زبان‌های برنامه‌نویسی مختلف.
- ✓ تکمیل کد: پیشنهاد تکمیل خودکار کد در حین نوشتن.
- ✓ رفع اشکال کد: پیدا کردن و رفع خطاهای کد.
- ✓ توضیح کد: تولید توضیحات برای کد.
- ✓ تبدیل زبان‌های برنامه‌نویسی: تبدیل کد از یک زبان به زبان دیگر.
- ✓ تولید مستندات کد: تولید مستندات برای پروژه‌های نرم‌افزاری.

۴. ترجمه و زبان

- ✓ ترجمه ماشینی: ترجمه متون از یک زبان به زبان دیگر.
- ✓ تصحیح گرامری و املائی: بررسی و تصحیح خطاهای گرامری و املائی.
- ✓ یادگیری زبان: تمرین مکالمه، یادگیری لغات جدید و غیره.

۵. سایر کاربردها

- ✓ تحقیق و پژوهش: جمع‌آوری اطلاعات، خلاصه‌سازی مقالات و غیره.
- ✓ سازماندهی اطلاعات: دسته‌بندی و سازماندهی اطلاعات.
- ✓ تولید داده‌های مصنوعی: تولید داده‌های مصنوعی برای آموزش مدل‌ها.
- ✓ تحلیل احساسات: تحلیل احساسات موجود در متن.
- ✓ تولید نام و شعار: تولید نام برای شرکت‌ها و محصولات، تولید شعارهای تبلیغاتی.

مزایا، چالش‌ها و محدودیتهای کار با ChatGPT

مزایای چت جی‌پی‌تی:

- ✓ دسترسی آسان و کاربرپسند: استفاده از چت جی‌پی‌تی معمولاً بسیار آسان است و نیاز به دانش فنی خاصی ندارد. رابط کاربری آن معمولاً ساده و قابل فهم است.

- ✓ سرعت بالا در پاسخگویی: چت جی‌پی‌تی می‌تواند در عرض چند ثانیه به سوالات و درخواست‌های کاربران پاسخ دهد.
- ✓ قابلیت تولید متون متنوع: این ابزار می‌تواند انواع مختلف متن را تولید کند، از داستان و شعر گرفته تا گزارش‌های فنی و ایمیل‌های رسمی.
- ✓ دسترسی شبانه‌روزی: چت جی‌پی‌تی به صورت ۲۴ ساعته در دسترس است و کاربران می‌توانند در هر زمان از آن استفاده کنند.
- ✓ قابلیت یادگیری و بهبود مستمر: مدل‌های زبانی مانند چت جی‌پی‌تی به طور مداوم در حال یادگیری و بهبود هستند و با گذشت زمان دقت و کیفیت پاسخ‌های آنها افزایش می‌یابد.
- ✓ صرفه‌جویی در زمان و هزینه: استفاده از چت جی‌پی‌تی می‌تواند در زمان و هزینه‌های مربوط به تولید محتوا، ترجمه، پشتیبانی مشتری و سایر موارد صرفه‌جویی کند.
- ✓ کمک به ایده‌پردازی و خلاقیت: چت جی‌پی‌تی می‌تواند به کاربران در ایده‌پردازی، تولید محتوای خلاقانه و حل مسائل کمک کند.
- ✓ قابلیت شخصی‌سازی (تا حدی): با استفاده از دستورات و ورودی‌های مناسب، می‌توان خروجی چت جی‌پی‌تی را تا حدی شخصی‌سازی کرد.

محدودیت‌های چت جی‌پی‌تی:

- احتمال تولید اطلاعات نادرست یا گمراه‌کننده: چت جی‌پی‌تی ممکن است گاهی اوقات اطلاعات نادرست، قدیمی یا بی‌اساس را به عنوان پاسخ ارائه دهد. به همین دلیل، بررسی صحت اطلاعات ارائه شده توسط این ابزار، بسیار مهم است.
- عدم درک کامل مفاهیم و ظرافت‌های زبانی: چت جی‌پی‌تی ممکن است در درک مفاهیم پیچیده، اصطلاحات، کنایه‌ها و مفاهیم فرهنگی دچار مشکل شود.
- عدم خلاقیت واقعی: چت جی‌پی‌تی بر اساس الگوهایی که از داده‌های آموزشی خود آموخته است، محتوا تولید می‌کند و خلاقیت واقعی به معنای خلق چیزی کاملاً جدید و بدیع را ندارد.
- وابستگی به داده‌های آموزشی: خروجی چت جی‌پی‌تی به شدت به داده‌های آموزشی آن وابسته است و هرگونه سوگیری یا نقص در داده‌ها می‌تواند در خروجی آن منعکس شود.
- مشکل در حفظ انسجام در متون طولانی: در تولید متون طولانی، ممکن است چت جی‌پی‌تی نتواند انسجام و پیوستگی لازم را حفظ کند.

- عدم آگاهی از رویدادهای بعد از آخرین به‌روزرسانی: چت جی‌پی‌تی از رویدادهایی که بعد از آخرین به‌روزرسانی مدل رخ داده‌اند، آگاهی ندارد. البته این ابزار با افزودن امکان اتصال به اینترنت، سعی کرده است تا حدودی این نقیصه را برطرف کند.
- مشکل در درک زمینه (Context): گاهی اوقات، چت جی‌پی‌تی در درک زمینه گفتگو و ارتباط بین سوالات و پاسخ‌ها دچار مشکل می‌شود. به همین دلیل تلاش در حفظ زمینه گفتگو و کنترل آن جهت خارج نشدن از زمینه بحث امری ضروری است.

چالش‌های استفاده از چت جی‌پی‌تی:

- مسائل اخلاقی و حقوقی: استفاده از چت جی‌پی‌تی برای تولید محتوای جعلی، همراه‌کننده یا مضر، نقض حقوق مالکیت معنوی و سایر مسائل اخلاقی و حقوقی، چالش‌های جدی را به وجود می‌آورد.
- مشکل در تشخیص محتوای تولید شده توسط هوش مصنوعی: تشخیص اینکه یک محتوا توسط انسان تولید شده یا توسط هوش مصنوعی، روز به روز دشوارتر می‌شود که می‌تواند مشکلاتی را در زمینه‌های مختلف مانند پژوهش و حقوق ایجاد کند.
- وابستگی بیش از حد به ابزار و کاهش مهارت‌های انسانی: استفاده زیاد از چت جی‌پی‌تی می‌تواند باعث کاهش مهارت‌های نوشتاری، تفکر انتقادی و حل مسئله در انسان‌ها شود.
- مسائل مربوط به حریم خصوصی و امنیت داده‌ها: استفاده از چت جی‌پی‌تی ممکن است نگرانی‌هایی در مورد حریم خصوصی و امنیت داده‌های کاربران ایجاد کند.
- سوگیری در پاسخ‌ها: به دلیل سوگیری‌های احتمالی در داده‌های آموزشی، ممکن است چت جی‌پی‌تی پاسخ‌هایی ارائه دهد که دارای سوگیری‌های نژادی، جنسیتی، مذهبی و غیره باشند.
- نیاز به نظارت انسانی: خروجی‌های چت جی‌پی‌تی نیازمند بررسی و نظارت انسانی هستند تا از صحت، دقت و مناسب بودن آنها اطمینان حاصل شود.

با در نظر گرفتن این مزایا، محدودیت‌ها و چالش‌ها، می‌توان از چت جی‌پی‌تی به صورت مؤثرتر و مسئولانه‌تر استفاده کرد. مهم است که به یاد داشته باشیم که چت

جی‌پی‌تی ابزاری قدرتمند است، اما نباید به آن به عنوان یک منبع قطعی و بی‌نقص اطلاعات اعتماد کرد.

معرفی چت جی‌پی‌تی

چت جی‌پی‌تی چیزی فراتر از یک ربات چت ساده است؛ این یک انقلاب در نحوه تعامل ما با دنیای رایانه‌ها و اطلاعات است. این ابزار قدرتمند مبتنی بر هوش مصنوعی، توسط شرکت OpenAI توسعه یافته و قادر است متن را به زبان طبیعی انسانی درک و تولید کند. این قابلیت، درهای جدیدی را به سوی کاربردهای مختلف باز کرده است، از پاسخگویی به سوالات ساده گرفته تا تولید محتوای خلاقانه و حتی برنامه نویسی رایانه‌ای.

راه‌اندازی حساب کاربری: ورود به دنیای چت جی‌پی‌تی

شروع کار با چت جی‌پی‌تی بسیار ساده است. برای دسترسی به این ابزار، کافی است به وبسایت [Chatgpt.com](https://chatgpt.com) مراجعه کرده و یک حساب کاربری ایجاد کنید. این فرآیند شامل وارد کردن آدرس ایمیل، ایجاد رمز عبور و تأیید حساب کاربری از طریق ایمیل است. پس از ایجاد حساب، می‌توانید به صفحه گفتگو چت جی‌پی‌تی دسترسی پیدا کرده و شروع به تعامل با آن کنید.

مقایسه چت جی‌پی‌تی با موتورهای جستجو و ابزارهای تحلیل

چت جی‌پی‌تی با ابزارهایی مانند موتورهای جستجو و نرم‌افزارهای تحلیل داده تفاوت‌های اساسی دارد.

- **موتورهای جستجو:** موتورهای جستجو مانند گوگل، لیستی از لینک‌ها به صفحات وب مرتبط با عبارت جستجوی شما را ارائه می‌دهند. شما باید خودتان در این صفحات به دنبال پاسخ سوال خود بگردید. در مقابل، چت جی‌پی‌تی مستقیماً به سوال شما پاسخ می‌دهد و می‌تواند اطلاعات را از منابع مختلف جمع‌آوری و به صورت منسجم ارائه کند.

• **ابزارهای تحلیل داده:** ابزارهای تحلیل داده مانند اکسل یا نرم‌افزارهای آماری، برای تحلیل داده‌های ساختاریافته (مانند جداول و پایگاه‌های داده) استفاده می‌شوند. آنها به شما کمک می‌کنند تا الگوها و روندهای موجود در داده‌ها را شناسایی کنید. در مقابل، چت جی‌پی‌تی با داده‌های متنی بدون غیرساختاریافته (مانند متن کتاب‌ها، مقالات و وب‌سایت‌ها) کار می‌کند و می‌تواند مفاهیم و روابط بین آنها را درک کند.

به طور خلاصه، موتورهای جستجو شما را به منابع اطلاعاتی هدایت می‌کنند، ابزارهای تحلیل داده به شما در تحلیل داده‌ها کمک می‌کنند، در حالی که چت جی‌پی‌تی اطلاعات را پردازش و به صورت مستقیم به شما ارائه می‌دهد.

رقابت با چت جی‌پی‌تی در اشتغال: تهدید یا فرصت؟

ظهور چت جی‌پی‌تی و سایر ابزارهای هوش مصنوعی مولد، نگرانی‌هایی را در مورد آینده مشاغل ایجاد کرده است. برخی از افراد نگرانند که این ابزارها جایگزین آنها در محیط کار شوند. با این حال، دیدگاه واقع‌بینانه‌تر این است که چت جی‌پی‌تی را به عنوان یک ابزار کمکی در نظر بگیریم که می‌تواند به ما در انجام وظایفمان کمک کند.

چت جی‌پی‌تی می‌تواند وظایف تکراری و زمان‌بر را خودکار کند، به ما در تولید محتوا، ایده‌پردازی و حل مسائل کمک کند و به این ترتیب، بهره‌وری ما را افزایش دهد. به جای دیدن آن به عنوان یک تهدید، باید یاد بگیریم که چگونه از آن به عنوان یک ابزار قدرتمند در جهت بهبود عملکرد خود استفاده کنیم؛ در چنین شرایطی احتمال حذف شدن از بازار کار بسیار پایین خواهد آمد.

تعریف مجدد چت‌بات‌ها با ظهور چت جی‌پی‌تی و چت جی‌پی‌تی پلاس

چت جی‌پی‌تی با توانایی‌های پیشرفته خود، تعریف جدیدی از چت‌بات‌ها ارائه داده است. چت‌بات‌های سنتی معمولاً بر اساس قوانین و الگوهای از پیش تعیین شده عمل می‌کردند و توانایی محدودی در درک زبان طبیعی داشتند. در مقابل، چت جی‌پی‌تی با استفاده از مدل‌های زبان بزرگ، قادر به درک و تولید متن به زبان طبیعی انسان است و می‌تواند مکالمات پیچیده‌تری را انجام دهد.

نسخه پلاس چت جی‌پی‌تی (ChatGPT Plus) که از مدل GPT-۴ بهره می‌برد، قابلیت‌های پیشرفته‌تری نسبت به نسخه رایگان دارد. این قابلیت‌ها شامل موارد زیر است:

- **درک عمیق‌تر از زبان:**

GPT-۴ در درک مفاهیم پیچیده و ظرافت‌های زبانی عملکرد بهتری دارد.

- **پاسخ‌های دقیق‌تر و منسجم‌تر:**

پاسخ‌های GPT-۴ معمولاً دقیق‌تر، منسجم‌تر و مرتبط‌تر با زمینه گفتگو هستند.

- **خلاقیت بیشتر:**

GPT-۴ در تولید محتوای خلاقانه عملکرد بهتری دارد.

- **دسترسی اولویت‌دار:**

مشترکین ChatGPT Plus در زمان‌های اوج ترافیک، دسترسی اولویت‌دار به سرویس دارند.

قرار گرفتن در صدر اخبار و مختل کردن کسب‌وکارها

ظهور چت جی‌پی‌تی از زمان عمومی شدن نسخه ۳ آن در سال ۲۰۲۲، توجه زیادی را به خود جلب کرده و تأثیر قابل توجهی بر کسب‌وکارها و صنایع مختلف داشته است. این فناوری، پتانسیل ایجاد تغییرات اساسی در نحوه انجام کارها و ارائه خدمات دارد.

چت جی‌پی‌تی نمونه‌ای از پیشرفت‌های تصاعدی با بهره‌گیری از هوش مصنوعی است. این فناوری به سرعت در حال توسعه است و پتانسیل ایجاد تغییرات بزرگی را در آینده دارد.

چت جی‌پی‌تی در حال حاضر تأثیر خود را بر صنایعی مانند خدمات مشتری، تولید محتوا، آموزش و برنامه‌نویسی به صورت محسوس نشان داده است.

از این رو، برای موفقیت در دنیای پر از تغییرات ناشی از هوش مصنوعی، باید خود را با این فناوری‌ها سازگار کنیم و از آنها به عنوان ابزاری برای بهبود عملکرد خود استفاده

کنیم. لذا هوش مصنوعی نباید به عنوان یک تهدید دیده شود، بلکه باید به عنوان ابزاری برای کمک به انسان‌ها و بهبود زندگی آنها در نظر گرفته شود.

باید توجه داشت که چت جی‌پی‌تی نقطه اوج این فناوری نیست، بلکه تنها شروعی است برای محصولات بزرگتر و قدرتمندتر هوش مصنوعی که در آینده شاهد آنها خواهیم بود.

تفاوت چت جی‌پی‌تی در چیست؟ تمایز از سایر ابزارها

همانطور که قبلاً اشاره شد، چت جی‌پی‌تی با چت‌بات‌های سنتی و سایر ابزارهای پردازش زبان طبیعی تفاوت‌های اساسی دارد. این تفاوت‌ها در نحوه درک و تولید زبان نهفته است. چت‌بات‌های سنتی اغلب بر اساس مجموعه‌ای از قوانین و الگوهای از پیش تعریف شده عمل می‌کنند و توانایی محدودی در درک پیچیدگی‌های زبان انسانی دارند. آنها معمولاً به سوالات ساده با پاسخ‌های از پیش تعیین شده پاسخ می‌دهند و قادر به انجام مکالمات پیچیده‌تر نیستند.

در مقابل، چت جی‌پی‌تی از مدل‌های زبان بزرگ (LLMs) استفاده می‌کند که با تحلیل حجم عظیمی از داده‌های متنی، الگوهای زبانی را یاد می‌گیرند و قادر به تولید متن‌های منسجم و مرتبط با زمینه گفتگو می‌شوند. این قابلیت، به چت جی‌پی‌تی اجازه می‌دهد تا مکالمات پیچیده‌تری را انجام دهد، به سوالات متنوع پاسخ دهد و حتی محتوای خلاقانه تولید کند.

معماری چت جی‌پی‌تی: زیرساخت‌های نرم‌افزاری

معماری چت جی‌پی‌تی بر پایه مدل‌های ترانسفورمر (Transformer) است. این معماری، انقلابی در زمینه پردازش زبان طبیعی ایجاد کرده است. مدل‌های ترانسفورمر با استفاده از مکانیسم توجه (Attention Mechanism) قادرند روابط بین کلمات در یک متن را به خوبی درک کنند. این مکانیسم به مدل اجازه می‌دهد تا به جای تمرکز صرف بر کلمات کنار هم قرار گرفته، به کلمات مهم‌تر در کل متن توجه کند و به این ترتیب، درک بهتری از معنای متن داشته باشد.

به عبارت ساده‌تر، مکانیسم توجه به مدل کمک می‌کند تا بفهمد کدام کلمات در یک جمله یا پاراگراف، بیشترین اهمیت را در انتقال معنا دارند. این قابلیت به ویژه در درک جملات طولانی و پیچیده و همچنین در ترجمه زبان بسیار مفید است.

ابرایانه‌ها و پردازنده‌های گرافیکی تحت آن: سخت‌افزار قدرتمند

آموزش و اجرای مدل‌های زبان بزرگ مانند چت جی‌پی‌تی نیازمند منابع محاسباتی بسیار زیادی است. این مدل‌ها بر روی ابررایانه‌های قدرتمند با هزاران پردازنده گرافیکی (GPU) آموزش داده می‌شوند. پردازنده‌های گرافیکی به دلیل توانایی بالای خود در پردازش موازی، برای آموزش مدل‌های یادگیری عمیق بسیار مناسب هستند.

آموزش این مدل‌ها فرآیندی زمان‌بر و پرهزینه است و نیازمند صرف انرژی زیادی است. به همین دلیل، تنها شرکت‌های بزرگ با منابع مالی و فنی کافی قادر به توسعه و آموزش چنین مدل‌هایی هستند.

در نظر گرفتن اهمیت ترانسفورمرها: کلید موفقیت

همانطور که اشاره شد، معماری ترانسفورمر نقش کلیدی در موفقیت چت جی‌پی‌تی ایفا می‌کند. این معماری مزایای متعددی نسبت به مدل‌های قبلی دارد، از جمله:

- **پردازش موازی:** ترانسفورمرها قادرند داده‌ها را به صورت موازی پردازش کنند که منجر به افزایش سرعت و کارایی می‌شود.
- **درک بهتر روابط بین کلمات:** مکانیسم توجه به مدل اجازه می‌دهد تا روابط بین کلمات در یک متن را به خوبی درک کند.
- **قابلیت پردازش متون طولانی:** ترانسفورمرها در پردازش متون طولانی نسبت به مدل‌های قبلی پردازش زبانی عملکرد بهتری دارند.

آموزش مدل - یادگیری از داده‌ها

مدل‌های زبانی بزرگ مانند چت جی‌پی‌تی با استفاده از حجم عظیمی از داده‌های متنی آموزش داده می‌شوند. این داده‌ها از منابع مختلفی مانند کتاب‌ها، مقالات، وبسایت‌ها، کدها و غیره جمع‌آوری می‌شوند. فرآیند آموزش به این صورت است که

مدل با تحلیل این داده‌ها، الگوهای زبانی را یاد می‌گیرد و قادر به پیش‌بینی کلمات بعدی در یک متن می‌شود.

به عبارت دیگر، مدل سعی می‌کند با دیدن متن‌های مختلف، یاد بگیرد که چگونه کلمات در کنار هم قرار می‌گیرند و چگونه جملات و پاراگراف‌ها ساخته می‌شوند. این فرآیند شبیه به یادگیری زبان توسط انسان است، با این تفاوت که مدل با سرعت و حجم بسیار بیشتری از داده‌ها آموزش می‌بیند.

درک اهمیت قالب گفتگو: تعامل انسان‌محور

چت جی‌پی‌تی به گونه‌ای طراحی شده است که بتواند در قالب گفتگو با کاربران تعامل داشته باشد. این ویژگی باعث می‌شود که استفاده از آن برای کاربران آسان‌تر و طبیعی‌تر به نظر برسد. کاربران می‌توانند سوالات خود را به زبان طبیعی مطرح کنند و پاسخ‌ها را نیز به همان شکل دریافت کنند.

این قالب گفتگو، تجربه کاربری بهتری را فراهم می‌کند و باعث می‌شود که کاربران احساس راحتی بیشتری در تعامل با این ابزار داشته باشند.

Copilot، کمک به برنامه‌نویسی و کدنویسی

موارد استفاده از copilot

۱. تولید و کمک به محتوا

- ✓ پرسش سؤال و پاسخ به آن
- ✓ تولید محتوای متنی (مقالات، ایمیل‌ها، گزارش‌ها)
- ✓ ساخت عکس با نوشتن متن یا بر اساس عکس
- ✓ ساخت موسیقی بر اساس متن
- ✓ خلاصه‌سازی متن

- ✓ ایجاد برنامه سفر و فرمول‌های اکسل
- ✓ نوشتن توضیحات و ادعاهای تبلیغاتی
- ✓ کمک به نگارش شعر و داستان (در زبان‌های مختلف)
- ✓ تحلیل محتوای صفحات وب و ویدئوهای متنی
- ✓ تولید متن در قالب فایل‌های PDF یا Word

۲. کدنویسی و توسعه نرم‌افزار

- ✓ کدنویسی و رفع اشکال
- ✓ تولید مستندات کد
- ✓ ادغام کد با دیگر نرم‌افزارها

۳. ترجمه و زبان‌شناسی

- ✓ ترجمه متون به زبان‌های مختلف
- ✓ اصلاح گرامری و املایی

۴. مدیریت داده و تحلیل

- ✓ ایجاد فرمول‌های اکسل
- ✓ مدیریت و پردازش داده‌ها

۵. سایر کاربردها

- ✓ حل تکالیف مدرسه و مسائل آموزشی
- ✓ تحلیل و تولید نام و شعار برای محصولات و کسب‌وکارها
- ✓ خلق تصاویر و گرافیک‌های متنی با کیفیت بالا
- ✓ ارائه سؤالات مرتبط با موضوعات مختلف و پیشنهادات تکمیلی

مزایا و چالش‌ها و محدودیتهای کار با کوپایلوت

مزایا:

- ✓ دسترسی به GPT-۴ و DALL-E ۳ برای ساخت تصاویر
- ✓ سرعت بالا در پاسخگویی به سؤالات
- ✓ قابلیت ادغام با مایکروسافت ۳۶۵ (ورد، اکسل، پاورپوینت)

- ✓ تولید ۷۲ تصویر رایگان در روز
- ✓ پشتیبانی از زبان‌های مختلف
- ✓ امکان ساخت موسیقی و تجزیه و تحلیل محتوای ویدئوها
- ✓ ارائه پاسخ‌های به‌روز با جستجو در منابع آنلاین
- ✓ سادگی استفاده و دسترسی از طریق وب و اپلیکیشن‌ها
- ✓ ایجاد تصاویری باکیفیت با جزئیات بالا
- ✓ عملکرد خوب در تشخیص محتوای تصویر
- ✓ ارائه منبع برای پاسخ‌ها
- ✓ قابل استفاده با IP ایران
- ✓ قابلیت ادغام در ورد، اکسل، پاورپوینت و اوت‌لوک
- ✓ قابلیت ارائه متن جواب در قالب فایل Word یا PDF

چالش‌ها

- دقت پایین برخی پاسخ‌ها
- ضعف در کدنویسی پیچیده و برخی کارهای فنی
- محدودیت در تحلیل محتوای پیچیده و نیاز به دستور حرفه‌ای
- عملکرد ضعیف در ایجاد پاورپوینت‌های پیشرفته
- عدم پشتیبانی از پرامپت صوتی فارسی

محدودیت‌ها

- محدودیت در نسخه رایگان (تعداد درخواست‌ها و تصاویر قابل تولید)
- عدم امکان ویرایش عکس‌های تولید شده
- محدودیت‌های زمانی در دسترسی به Turbo GPT-۴ در زمان اوج ترافیک

بررسی علمی و ساده‌ی هوش مصنوعی کوپایلوت

ابزار هوش مصنوعی کوپایلوت، به عنوان یک دستیار هوشمند مجازی، بهترین گزینه محسوب نمی شود اما می‌توان آن را در ابزاری برای جستجوی دقیق‌تر در اینترنت دانست. البته با ذکر این نکته که روز به روز بر امکانات و قابلیت‌های آن افزوده می شود.

این ابزار اغلب اطلاعات خود را از طریق جستجو در وب و خلاصه‌سازی محتوای وب‌سایت‌ها به دست می‌آورد. از این رو، با استفاده از دستورات (پرامپت) حرفه‌ای، می‌توان از آن در انجام برخی وظایف تحلیلی و پردازشی کمک گرفت.

هدف از توسعه‌ی کوپایلوت

کوپایلوت اساساً برای ادغام با نرم‌افزارهای آفیس، موتور جستجوی بینگ و سایر برنامه‌های تحت وب مایکروسافت طراحی شده است. با این حال، به دلیل استفاده از مدل زبانی GPT، می‌توان با ارائه‌ی توضیحات یا درخواست‌ها، سؤالات و نیازهای خود را به آن منتقل کرد. به‌عنوان مثال، می‌توان از آن درخواست آموزش آشپزی، ساخت رزومه، ترجمه‌ی متون تخصصی، نگارش محتوای کوتاه یا توضیح مفاهیم پیچیده‌ی فیزیک کوانتوم به زبان ساده را داشت. افزونه‌ها (پلاگین‌ها) نیز امکان انجام کارهای پیشرفته‌تر مانند ساخت موسیقی با هوش مصنوعی را فراهم می‌کنند. اما کوپایلوت در شرایط فعلی نمی‌تواند مانند ChatGPT یک تحلیل‌گر و پردازشگر حرفه‌ای باشد؛ از آن چنین توقعی هم نباید داشت چرا که هدف و مأموریت آن متفاوت است.

همکاری مایکروسافت و OpenAI

کوپایلوت حاصل همکاری مشترک مایکروسافت و OpenAI است. این همکاری با سرمایه‌گذاری مستقیم مایکروسافت در OpenAI ابتدا منجر به توسعه‌ی ChatGPT و سپس شکل‌گیری کوپایلوت بر مبنای مدل بزرگ زبانی GPT شد.

مزایای کوپایلوت نسبت به سایر چت‌بات‌های هوش مصنوعی

- **دسترسی گسترده:** کوپایلوت نه تنها از طریق وب و اپلیکیشن موبایل، بلکه در مرورگر Edge، ویندوز ۱۱، موتور جستجوی بینگ و حتی برنامه‌های مایکروسافت ۳۶۵ مانند ورد، اکسل و پاورپوینت نیز در دسترس است. همچنین، امکان ادغام آن با سایر نرم‌افزارها از طریق API وجود دارد.

- **اطلاعات به‌روز:** کوپایلوت با استفاده هم‌زمان از داده‌های آموزشی و جستجو در منابع معتبر از طریق بینگ، به درخواست‌های کاربران پاسخ می‌دهد و اطلاعات آن همواره

بهروز است. (در حال حاضر چت جی‌پی‌تی و جمینای نیز این قابلیت را به چت بات خود افزوده اند ولی تا رسیدن به نقطه مطلوب کمی فاصله دارند)

- **قابلیت‌های پیشرفته:** کوپایلویت توانایی خلق تصاویر، ساخت موسیقی (در نسخه‌ی رایگان) و ادغام با نرم‌افزارهای مایکروسافت ۳۶۵ را دارد.

- **انتخاب حالت چت:** کوپایلویت امکان انتخاب حالت چت (متعادل، دقیق و خلاقانه) توسط کاربران را در ابتدای مکالمه فراهم می‌کند.

کارایی کوپایلویت در پاسخ به سؤالات و جستجو در اینترنت

کوپایلویت قابلیت جستجو در اینترنت را دارد و می‌تواند در یافتن سریع پاسخ سؤالات کمک کند. برخلاف ChatGPT، نیازی به دستور صریح برای جستجو یا استفاده از افزونه‌ی خاص نیست. این قابلیت به صورت پیش‌فرض فعال است و به طور خودکار به منابع آنلاین مراجعه می‌کند.

تطبیق با زبان کاربر

یکی از ویژگی‌های برجسته‌ی کوپایلویت، تطبیق با زبان کاربر است. اگر به زبان فارسی پرسش شود، کوپایلویت در منابع فارسی جستجو می‌کند و بالعکس.

بهروز بودن اطلاعات

کوپایلویت در زمینه‌ی جستجو ابزاری قدرتمند و قابل اتکا به نظر می‌رسد و با ارجاع به منابع معتبر، امکان بررسی صحت اطلاعات را فراهم می‌کند. با این حال، مانند هر ابزار مبتنی بر هوش مصنوعی، بررسی دقیق اطلاعات، به ویژه در موضوعات تخصصی، ضروری است.

توانمندی‌ها در تولید محتوای متنی

کوپایلوت در تولید محتوای متنی قابلیت‌های گسترده‌ای دارد، از جمله پیشنهاد ساختار محتوا، تعیین موضوعات کلیدی، نگارش مقالات، ایمیل‌ها و حتی ادعای خلق شعر و داستان (که البته در زبان فارسی کیفیت پایینی دارد). همچنین، می‌تواند متون را خلاصه، گسترش یا بازنویسی کند و لحن متن را تغییر دهد. با این حال، عنصر خلاقیت هنوز به طور قابل توجهی در محتوای تولیدی آن دیده نمی‌شود و از این جهت نسبت به جمینای و چت جی‌پی‌تی عقب‌تر است.

آنالیز محتوای صفحات وب

کوپایلوت می‌تواند محتوای صفحات وب را خلاصه کند، نکات مهم را برجسته کند و به سؤالات مربوط به محتوا پاسخ دهد. این قابلیت برای دسترسی سریع به اطلاعات کلیدی بسیار مفید است.

درک محتوای ویدئوها

کوپایلوت می‌تواند ویدئوهایی که دارای متن (ترنسکریپت) هستند را تحلیل کند و به سؤالات مربوط به محتوا پاسخ دهد.

درک اسناد متنی

کوپایلوت می‌تواند محتوای اسناد متنی مانند PDF و Word را تحلیل کند. برای این کار، باید نسخه‌ی آنلاین اسناد در مرورگر Edge باز شود.

خلق تصاویر

ابزار Designer در کوپایلوت، با استفاده از DALL-E ۳، می‌تواند محتوای متنی یا صوتی را به تصاویر تبدیل کند. هرچند دقت آن گاهی پایین است و فرایند تولید عکس سریع نیست، اما برای نیازهای روزمره کافی است.

درک تصاویر

کوپایلویت توانایی درک و تحلیل تصاویر را دارد.

ترجمه

کوپایلویت قادر به ترجمه‌ی متون بین زبان‌های مختلف است و ابزاری قابل اعتماد برای ترجمه به شمار می‌رود. با این حال، در ترجمه‌ی اصطلاحات خاص و عبارات فرهنگی، بررسی دقیق‌تر در ترجمه تولیدی آن توصیه می‌شود.

ادغام با نرم‌افزارهای مایکروسافت ۳۶۵

کوپایلویت (در نسخه‌های پولی) با نرم‌افزارهای مایکروسافت ۳۶۵ ادغام می‌شود و بهره‌وری را افزایش می‌دهد. این ادغام شامل قابلیت‌هایی مانند ایجاد پیش‌نویس متن، ایجاد جداول و فرمول در اکسل و تهیه‌ی اسلاید در پاورپوینت است.

راه‌های دسترسی به کوپایلویت

- [وبگاه کوپایلویت](#)
- سایدبار کوپایلویت در ویندوز ۱۱
- موتور جستجوی بینگ
- سایدبار مرورگر مایکروسافت اج
- اپ موبایلی آیفون و اندروید
- برنامه‌های مایکروسافت ۳۶۵ (در اشتراک پولی)

مقایسه طرح‌های رایگان و اشتراکی

کوپایلویت در ۶ طرح مختلف ارائه می‌شود. طرح رایگان دارای محدودیت‌هایی در تعداد پیام در هر مکالمه، تعداد تصاویر قابل تولید و دسترسی به GPT-۴ Turbo در

زمان اوج ترافیک است. طرح Pro با پرداخت ماهانه ۲۰ دلار، این محدودیت‌ها را برطرف می‌کند و امکاناتی مانند دسترسی دائمی به GPT-۴ Turbo، سرعت بالاتر، تولید تصویر با کیفیت بالا و ادغام با برنامه‌های مایکروسافت ۳۶۵ را فراهم می‌کند.

Gemini، جستجوی پیشرفته و ترجمه

موارد استفاده از Gemini

۱. تولید و کمک به محتوا

- ✓ پرسش سؤال و پاسخ به آن
- ✓ تولید محتوای متنی (مقالات، ایمیل‌ها، گزارش‌ها)
- ✓ تولید تصاویر خلاقانه (در نسخه اولترا)
- ✓ خلاصه‌سازی محتوای متنی، ویدئویی یا صوتی
- ✓ برنامه‌ریزی سفر با Google Maps و Google Flights
- ✓ نوشتن ایمیل یا تولید و یافتن اطلاعات سریع در گوگل درایو و گوگل شیت

۲. کدنویسی و توسعه نرم‌افزار

- ✓ کدنویسی و رفع اشکال
- ✓ تولید مستندات کد
- ✓ ادغام کد با دیگر نرم‌افزارها

۳. ترجمه و زبان‌شناسی

- ✓ ترجمه متون به زبان‌های مختلف
- ✓ اصلاح گرامری و املائی

۴. مدیریت داده و تحلیل

- ✓ ایجاد فرمول‌های اکسل
- ✓ مدیریت و پردازش داده‌ها

۵. سایر کاربردها

- ✓ تحلیل و تولید نام و شعار برای محصولات و کسب‌وکارها
- ✓ خلق تصاویر و گرافیک‌های متنی با کیفیت بالا
- ✓ ارائه سؤالات مرتبط با موضوعات مختلف و پیشنهادات تکمیلی

مزایا و محدودیت‌های کار با جمینای

مزایا:

- ✓ رابط کاربری ساده و کاربرپسند
- ✓ قابل اجرا روی گوشی و کامپیوتر
- ✓ رایگان و پولی برای کاربری‌های خاص
- ✓ ادغام‌شده با سایر محصولات گوگل
- ✓ پاسخ‌گویی سریع به پرامپت‌های کاربران
- ✓ ارائه پاسخ‌ها و اطلاعات بلادرنگ و به‌روز
- ✓ توانمند در استدلال منطقی و حل مسئله
- ✓ ارائه قابلیت چندرسانه‌ای در یک پلتفرم واحد

محدودیت‌ها و چالش‌ها:

- نیاز به وی‌پی‌ان و نرم‌افزار تغییر IP
- تولیدنکردن تصویر در نسخه رایگان
- سوگیری در پاسخ‌گویی در برخی موارد
- ارائه اطلاعات نادرست یا گمراه‌کننده
- نداشتن خلاقیت و دقت کافی در پاسخ‌گویی

گوگل جمینای چیست؟

گوگل جمینای یک مدل زبانی بزرگ (LLM) و چت‌بات هوش مصنوعی است که ابتدا با نام بارد معرفی شد و سپس به جمینای تغییر نام داد. این ابزار قادر به درک و

پاسخگویی به زبان انسان است و مانند یک ربات گفتگوی متنی عمل می‌کند. جمینای با استفاده از الگوریتم‌های هوش مصنوعی مانند یادگیری ماشین، شبکه‌های عصبی و پردازش زبان طبیعی، محتوای ورودی را درک کرده و پاسخ مناسب تولید می‌کند.

جمینای در حال حاضر جدی‌ترین رقیب چت جی‌پی‌تی محسوب می‌شود. تفاوت اصلی جمینای با سایر مدل‌های مشابه مانند ChatGPT و Copilot، قابلیت چندرسانه‌ای (MultiModal) آن است. جمینای علاوه بر متن، توانایی درک عکس، صدا و فیلم را نیز دارد و می‌تواند با برنامه‌های گوگل مانند جی‌میل و گوگل درایو ادغام شود و از طریق موتور جستجوی گوگل، اطلاعات به‌روز ارائه دهد.

ویژگی‌های هوش مصنوعی گوگل جمینای

جمینای از مجموعه‌های داده‌ی متنی، صوتی، تصویری، ویدئویی و مدل‌های زبانی قدرتمند برای درک و تحلیل درخواست‌های کاربران استفاده می‌کند. طراحی این مدل بر اساس نحوه‌ی درک و تعامل انسان‌هاست و هدف آن، تبدیل شدن به یک دستیار کارآمد است.

این چت‌بات منعطف، قادر به درک و پاسخگویی به انواع محتواهاست، اما باید توجه داشت که مانند سایر مدل‌های هوش مصنوعی، در حال یادگیری است و نباید به عنوان منبعی کاملاً موثق در نظر گرفته شود.

جمینای در نسخه‌ها و مدل‌های مختلف برای سیستم‌های شخصی و تلفن‌های همراه در قالب وب اپلیکیشن، نرم‌افزار یا افزونه در دسترس است. جمینای پرو ۱.۵ ویژگی‌های قابل توجهی دارد:

۱. **توانایی درک و تولید محتوای چندرسانه‌ای:** این قابلیت به جمینای امکان می‌دهد تا کارهای متنوعی انجام دهد. به‌عنوان مثال، می‌تواند خلاصه‌ای از یک کتاب طولانی یا محتوای یک ویدئو را ارائه دهد و همچنین متن‌های داخل تصاویر را ترجمه، تایپ یا توصیف کند. در نسخه‌ی پریمیوم، امکان جداسازی بخش‌های مهم ویدئو نیز وجود دارد.

۲. **کدنویسی آسان‌تر:** جمینای می‌تواند در کدنویسی به کاربران کمک کند، خطاهای برنامه‌ها را شناسایی و رفع کند، کدهای ناقص را کامل کند، راه‌حل‌های مختلف

ارائه دهد و زبان‌های برنامه‌نویسی را به یکدیگر تبدیل کند. همچنین می‌تواند برنامه‌ای را نوشته و کدنویسی و اجرای آن را توضیح دهد. طبق آمار گوگل، جمینای پرو ۱.۵ می‌تواند در عرض چند ثانیه، هزاران خط کد را آنالیز کند. با این حال، باید توجه داشت که خروجی جمینای ممکن است همیشه دقیق نباشد و باید بررسی شود.

۳. **جستجوی آسان‌تر اطلاعات:** جمینای می‌تواند به سؤالات کاربران در موضوعات مختلف پاسخ دهد و با مراجعه آنلاین به منابع مختلف، پاسخ‌های جامع ارائه کند. همچنین، منابعی را برای کسب اطلاعات بیشتر پیشنهاد می‌دهد. با این حال، پاسخ‌های جمینای ممکن است همیشه قابل اعتماد و دقیق نباشند و باید با احتیاط مورد استفاده قرار گیرند.

۴. **ویژگی‌های خلاقانه و تعاملی:** جمینای ابزارها و امکاناتی را برای تولید محتوای خلاقانه ارائه می‌دهد. می‌تواند جوک بگوید، شعر بخواند یا داستان بنویسد. این ویژگی می‌تواند برای محتوانویسان در ایده‌پردازی مفید باشد.

توجه: جمینای با استفاده از یادگیری انتقالی به طور مداوم در حال یادگیری و تکامل است.

قابلیت‌های گوگل جمینای

نسخه ۱.۰ جمینای اولترا در ارزیابی‌های MMLU (سنجش درک زبان‌های چندوظیفه‌ای) از هوش انسانی پیشی گرفته است. این آزمون، توانایی مدل‌های زبانی در درک مفاهیم پیچیده و استدلال منطقی را می‌سنجد.

گوگل جمینای در سه مدل مختلف عرضه می‌شود:

۱. **جمینای نانو:** مدل کوچک و کم‌حجمی که در تلفن‌های همراه برای کارهایی مانند خلاصه‌سازی در برنامه‌ی ضبط صدا و پاسخ‌های هوشمند در Gboard استفاده می‌شود.
۲. **جمینای پرو:** مدل پیشرفته‌تر که در مراکز داده‌ی گوگل اجرا می‌شود و سرعت پردازش بالاتری دارد. این مدل در حل مسائل پیچیده و پردازش محتوای چندرسانه‌ای توانمند است و از طریق رابط برنامه‌نویسی Vertex AI قابل استفاده است.
۳. **جمینای اولترا:** پیشرفته‌ترین مدل جمینای که قابلیت چندوجهی دارد و می‌تواند مسائل پیچیده و فرمول‌های علمی را پردازش کند و تصاویر باکیفیت تولید کند. این مدل از طریق چت‌بات پیشرفته‌ی جمینای (قبلاً Bard Advanced) در دسترس است و با پرداخت هزینه‌ی اشتراک قابل استفاده است. در حال حاضر نسخه Flash آن به صورت رایگان در دسترس عموم قرار گرفته است.

نحوه‌ی عملکرد گوگل جمینای

جمینای با استفاده از الگوریتم‌های هوش مصنوعی مانند یادگیری عمیق، یادگیری ماشین و پردازش زبان طبیعی و با آموزش روی حجم عظیمی از داده‌های چندرسانه‌ای، محتوا را درک می‌کند، به سؤالات پاسخ می‌دهد و متن تولید می‌کند. از تکنیک یادگیری تقویتی با بازخورد انسان (RLHF) نیز برای بهبود پاسخ‌گویی استفاده می‌کند. به همین دلیل این مدل زبانی به صورت مداوم در حال بهبود کارایی و عملکرد خویش است.

تفاوت گوگل جمینای با موتور جستجوی گوگل

موتور جستجوی گوگل با خزیدن در صفحات وب و رتبه‌بندی آن‌ها بر اساس کلیدواژه‌ها، به جستجوها پاسخ می‌دهد. در حالی که جمینای با استفاده از هوش مصنوعی، رفتاری تعاملی و انسان‌گونه دارد و می‌تواند گفتگوی دوطرفه داشته باشد و کارهایی فراتر از جستجوی ساده انجام دهد.

ساخت تصویر با گوگل جمینای

قابلیت تولید تصویر در جمینای ابتدا به صورت موقت غیرفعال شده بود، اما اکنون در نسخه Advanced در دسترس است. برای تولید تصویر، کاربر باید یک پرامپت متنی (ترجیحاً به زبان انگلیسی) ارائه دهد. جمینای دو تصویر متناسب با پرامپت تولید می‌کند و کاربر می‌تواند تصاویر را دانلود کند. جمینای از مدل ۳ Imagen برای تولید تصاویر استفاده می‌کند.

مقایسه‌ی جمینای با GPT-۴ و Claude

در این بخش، به مقایسه جمینای، چت جی‌پی‌تی و کلود می‌پردازیم و نقاط قوت و ضعف هر کدام را مشخص می‌کنیم. مقایسه‌ی دقیق و جامعی بین این مدل‌ها به دلیل تغییرات مداوم و انتشار نسخه‌های جدید دشوار است، اما می‌توان به طور کلی به موارد زیر اشاره کرد:

۱. قابلیت چندوجهی: (MultiModal)

• گوگل جمینای

جمینای از ابتدا با قابلیت چندوجهی طراحی شده و به طور native می‌تواند متن، تصویر، صدا و ویدئو را پردازش کند. این یک مزیت بزرگ برای جمینای محسوب می‌شود.

• GPT-۴

چت جی‌پی‌تی نیز قابلیت چندوجهی دارد، اما این قابلیت به اندازه‌ی جمینای یکپارچه نیست و در برخی موارد نیاز به استفاده از API‌های جداگانه دارد.

• Claude

Claude در ابتدا بر روی پردازش متن تمرکز داشت، اما نسخه‌های جدیدتر آن قابلیت‌های چندوجهی محدودتری را نیز ارائه می‌دهند.

• درک زبان و تولید متن

هر سه مدل در درک زبان و تولید متن‌های روان و منسجم عملکرد بسیار خوبی دارند. با این حال، در برخی موارد، جمینای و GPT-۴ درک بهتری از مفاهیم پیچیده و ارائه‌ی پاسخ‌های دقیق‌تر نشان می‌دهند.

• کدنویسی

هر سه مدل می‌توانند در کدنویسی کمک کنند، اما GPT-۴ به دلیل آموزش بر روی داده‌های کد بیشتر، در این زمینه عملکرد قوی‌تری دارد. جمینای نیز در کدنویسی پیشرفت‌های زیادی داشته و می‌تواند در بسیاری از موارد به خوبی عمل کند.

• استدلال و منطق

جمینای اولترا در تست‌های استدلال و منطق، به خصوص MMLU، عملکرد بسیار خوبی از خود نشان داده و حتی از عملکرد انسانی نیز پیشی گرفته است. GPT-۴ نیز در این زمینه قوی است، اما Claude معمولاً در این بخش ضعیف‌تر عمل می‌کند.

• خلاقیت

هر سه مدل می‌توانند محتوای خلاقانه تولید کنند، اما کیفیت و نوع خلاقیت آن‌ها متفاوت است. GPT-۴ معمولاً در تولید داستان و شعر عملکرد بهتری دارد، در حالی که جمینای ممکن است در ارائه‌ی ایده‌های جدید و ترکیب مفاهیم مختلف قوی‌تر باشد.

• گوگل جمینای

جمینای پرو از طریق Gemini به صورت رایگان در دسترس است. جمینای اولترا نیز با اشتراک Google One AI Premium قابل دسترسی است.

• GPT-۴

دسترسی به GPT-۴ از طریق اشتراک ChatGPT Plus امکان‌پذیر است. همچنین، از طریق API نیز می‌توان به آن دسترسی داشت.

• Claude

Claude از طریق وبگاه Anthropic و همچنین از طریق API قابل دسترسی است.

۴. سایر ویژگی‌ها:

- **ادغام با سرویس‌ها**

جمینای به خوبی با سرویس‌های گوگل مانند جستجو، جی‌میل و درایو ادغام شده است که این یک مزیت بزرگ برای کاربران اکوسیستم گوگل محسوب می‌شود.

- **سرعت و مقیاس‌پذیری**

گوگل با زیرساخت قدرتمند خود، می‌تواند جمینای را به صورت گسترده و با سرعت بالا ارائه دهد.

به طور خلاصه:

- **جمینای:** در قابلیت چندوجهی، ادغام با سرویس‌های گوگل و استدلال و منطق قوی است.
- **GPT-۴:** در کدنویسی، تولید محتوای خلاقانه و دسترسی به افزونه‌های بیشتر قوی است.
- **Claude:** بر روی ایمنی و اخلاق در هوش مصنوعی تمرکز بیشتری دارد.

انتخاب بهترین مدل بستگی به نیازها و اولویت‌های کاربر دارد. اگر به قابلیت چندوجهی و ادغام با سرویس‌های گوگل نیاز دارید، جمینای گزینه مناسبی است. اگر به کدنویسی و تولید محتوای خلاقانه اهمیت می‌دهید، GPT-۴ ممکن است انتخاب بهتری باشد. و اگر به ایمنی و اخلاق در هوش مصنوعی اهمیت می‌دهید، Claude می‌تواند گزینه مناسبی باشد.

Claude، دستیار نوشتاری و خلاقیت

موارد استفاده از کلود

۱. تعامل طبیعی و اخلاقی:

- کلود از رویکرد "هوش مصنوعی مبتنی بر اصول ایمنی" استفاده می‌کند که آن را برای کاربران امن‌تر و پاسخ‌گوتر می‌کند.
- توانایی شبیه‌سازی گفت‌وگوهای انسانی طبیعی و ارائه پاسخ‌های شفاف و بدون تعصب.

۲. تمرکز بر شفافیت و امنیت داده:

- برخلاف بسیاری از ابزارهای دیگر، کلود طراحی شده تا با داده‌های کاربران به شکل ایمن برخورد کند و حریم خصوصی آنها را حفظ کند.

۳. تولید محتوای متنی پیشرفته:

- نوشتن مقالات و متون طولانی با لحن‌های مختلف (رسمی، غیررسمی و خلاقانه).
- خلاصه‌سازی متون با دقت بالا برای صرفه‌جویی در زمان کاربران.
- تولید ایده‌های نوآورانه برای پروژه‌های خلاقانه.

۴. کدنویسی پیشرفته:

- کلود با تمرکز بر خوانایی کد و مستندسازی، به توسعه‌دهندگان کمک می‌کند مشکلات پیچیده کدنویسی را حل کنند.
- توانایی ارائه راه‌حل‌های کدنویسی با بهره‌وری بالا، مشابه اما دقیق‌تر از بسیاری ابزارهای مشابه.

۵. استدلال منطقی و حل مسئله:

- کلود قابلیت درک پیچیدگی‌های مسئله و ارائه استدلال منطقی را دارد که در مقایسه با سایر ابزارها، پاسخ‌هایی عمیق‌تر و تحلیلی‌تر ارائه می‌کند.

۶. ترجمه و ویرایش زبانی:

- ترجمه دقیق‌تر با تمرکز بر مفهوم و ظرافت‌های زبانی، مناسب برای کاربردهای تخصصی.
- بهینه‌سازی گرامر و سبک نوشتاری متن‌ها به صورت حرفه‌ای.

۷. سرعت و کیفیت در پاسخ‌گویی:

- ارائه پاسخ‌های سریع‌تر در مقایسه با ابزارهای دیگر، حتی در زمینه درخواست‌های پیچیده و طولانی.

مزایای کار با کلود

- ✓ تمرکز بر ایمنی و اخلاق
- ✓ مفید بودن و پاسخگویی مناسب
- ✓ صداقت در پاسخ‌ها.
- ✓ مکالمه‌ی طبیعی و روان
- ✓ سرعت و کارایی
- ✓ قابلیت‌های متنوع تولید متن، خلاصه‌سازی، ترجمه، کدنویسی و تحلیل داده
- ✓ دسترسی از طریق API و وب
- ✓ رابط کاربری ساده و تطبیق‌پذیری
- ✓ عمق در ارائه پاسخ‌ها
- ✓ تعامل چندمرحله‌ای مؤثر
- ✓ امنیت داده
- ✓ قابلیت ادغام با نرم‌افزارها و سرویس‌های دیگر

محدودیت‌های کار با کلود

- محدودیت در قابلیت‌های چندوجهی و تمرکز بر متن
- احتمال تولید پاسخ‌های نادرست
- عدم درک کامل مفاهیم بسیار پیچیده
- محدودیت در دسترسی رایگان (بسته به نسخه)
- عدم پشتیبانی کامل از برخی زبان‌ها (از جمله فارسی در برخی موارد)
- محدودیت‌های مربوط به طول متن ورودی
- وابستگی به اینترنت و ارائه به صورت آنلاین
- محدودیت در شخصی‌سازی عمیق

معرفی جامع هوش مصنوعی Claude

کلود (Claude) یک چت بات هوش مصنوعی است که توسط شرکت آنتروپیک (Anthropic)، یک شرکت تحقیقاتی در زمینه‌ی هوش مصنوعی، در سال ۲۰۲۲ توسعه داده شده است. آنتروپیک بر روی توسعه‌ی هوش مصنوعی ایمن و مفید تمرکز دارد و کلود نیز با همین هدف طراحی شده است. هدف اصلی کلود، ایجاد یک دستیار هوشمند و مفید است که می‌تواند در زمینه‌های مختلف به انسان‌ها کمک کند.

اهداف و ویژگی‌های کلیدی کلود

کلود با تمرکز بر اصول زیر توسعه یافته است:

- مفید بودن (Helpfulness): کلود طوری طراحی شده که به سؤالات و درخواست‌های کاربران به بهترین شکل ممکن پاسخ دهد و در انجام وظایف مختلف به آن‌ها کمک کند.
 - بی‌ضرر بودن (Harmlessness): آنتروپیک تأکید زیادی بر ایمنی کلود دارد و تلاش می‌کند تا از تولید پاسخ‌های مضر، تبعیض‌آمیز یا خطرناک جلوگیری کند.
 - صداقت (Honesty): کلود طوری آموزش دیده که در صورت عدم اطمینان از صحت یک پاسخ، آن را به صراحت اعلام کند و از ارائه‌ی اطلاعات نادرست خودداری کند.
- این اصول باعث شده‌اند که کلود به عنوان یک چت بات هوش مصنوعی با تمرکز بر اخلاق و ایمنی شناخته شود.

قابلیت‌های کلود

- کلود قابلیت‌های متنوعی دارد که از جمله آن‌ها می‌توان به موارد زیر اشاره کرد:
- پاسخگویی به سؤالات: کلود می‌تواند به سؤالات مختلف کاربران در زمینه‌های گوناگون پاسخ دهد. این شامل سؤالات عمومی، اطلاعات تخصصی، سؤالات مربوط به کدنویسی و غیره می‌شود.
 - تولید متن: کلود می‌تواند انواع مختلف متن را تولید کند، از جمله مقالات، داستان‌ها، ایمیل‌ها، کد و غیره.

- خلاصه‌سازی متن: کلود می‌تواند متون طولانی را خلاصه کرده و نکات اصلی آن‌ها را ارائه دهد.
- ترجمه: کلود می‌تواند متون را بین زبان‌های مختلف ترجمه کند.
- کدنویسی: کلود می‌تواند در کدنویسی به برنامه‌نویسان کمک کند، کد بنویسد، کد را دیباگ کند و توضیحات مربوط به کد را ارائه دهد.
- تحلیل داده: کلود می‌تواند به تحلیل داده‌ها و ارائه‌ی بینش‌های مفید کمک کند.
- مکالمه‌ی طبیعی: کلود قادر به برقراری مکالمه‌ی طبیعی و تعاملی با کاربران است.

تفاوت‌های کلود با سایر مدل‌های زبانی بزرگ

کلود در برخی جنبه‌ها با سایر مدل‌های زبانی بزرگ مانند GPT-4 و جمینای تفاوت دارد:

- تمرکز بر ایمنی و اخلاق: آنتروپیک تمرکز ویژه‌ای بر ایمنی و اخلاق در توسعه‌ی کلود داشته است. این باعث شده که کلود در تولید پاسخ‌های مضر و خطرناک محافظه‌کارتر باشد.
- اندازه‌ی مدل: مدل‌های کلود معمولاً از نظر اندازه کوچک‌تر از مدل‌های GPT-4 هستند. این می‌تواند منجر به سرعت بیشتر و مصرف منابع کمتر شود، اما ممکن است در برخی موارد درک مفاهیم پیچیده را محدود کند.
- رویکرد آموزشی: آنتروپیک از روش‌های آموزشی خاصی برای کلود استفاده می‌کند که هدف آن افزایش ایمنی و مفید بودن آن است.

نسخه‌های کلود

آنتروپیک نسخه‌های مختلفی از کلود را ارائه کرده است که هر کدام ویژگی‌ها و قابلیت‌های خاص خود را دارند. از جمله نسخه‌های مهم می‌توان به موارد زیر اشاره کرد:

۱. **Claude 1**: اولین نسخه از مدل Claude در مارس ۲۰۲۳ عرضه شد. این نسخه نشان‌دهنده توانایی‌های بالای مدل در انجام وظایف مختلف بود، اما همچنان محدودیت‌هایی در درک و پاسخ‌دهی داشت.
۲. **Claude 2**: در ژوئیه ۲۰۲۳، نسخه دوم با بهبودهای قابل توجهی در دقت و سرعت پاسخ‌دهی معرفی شد. این نسخه توانست به عنوان رقیبی جدی برای مدل‌های مشابه مانند ChatGPT مطرح شود.
۳. **Claude 3**: این نسخه شامل سه مدل پیشرفته به نام‌های Haiku، Sonnet و Opus است که هر کدام برای کاربردهای خاصی طراحی شده‌اند:

- Claude ۳ Haiku: مدلی فشرده و کارآمد برای وظایف پایه‌ای زبانی با حجم کم و سرعت پردازش بالا.
 - Claude ۳ Sonnet: مدلی پیشرفته‌تر با قابلیت‌های پردازش زبان طبیعی برای تولید متن، خلاصه‌سازی و تحلیل احساسات.
 - Claude ۳ Opus: پیشرفته‌ترین مدل با توانایی‌های چندگانه در زمینه‌های مختلف مانند پردازش زبان، تحلیل داده و خلق محتوای خلاقانه.
۴. Claude ۳/۵: مدل Claude Sonnet ۳/۵ به تازگی معرفی شده و طبق ادعای شرکت آنتروپیک، قابلیت رقابت با هوش مصنوعی GPT-۴ یا هوش مصنوعی Gemini گوگل را دارد.

ویژگی‌های برجسته نسخه‌های جدید:

- **قدرت Vision قوی:** Claude ۳ دارای قابلیت پردازش و درک تصاویر است که آن را همگام با دیگر مدل‌های پیشرو می‌سازد و می‌تواند فرمت‌های بصری مختلف را بررسی و پردازش کند.
- **رفع محدودیت‌ها:** در نسخه‌های پیشین، کلود درک بالایی نداشت و به علت همین نقص از پاسخ به برخی موارد خودداری می‌کرد. در مدل‌های Claude ۳، این مسئله رفع شده و مدل می‌تواند آسیب‌ها را شناسایی و به سؤالات پاسخ دهد.
- **بهبود دقت پاسخ‌دهی:** نسخه‌های جدید با الگوریتم‌های پیشرفته‌تر، دقت بالاتری در پاسخ‌دهی به کاربران ارائه می‌دهند.

هوش مصنوعی کلود با این نسخه‌ها و بهبودهای مستمر، به عنوان یکی از پیشرفته‌ترین مدل‌های زبانی در حوزه هوش مصنوعی شناخته می‌شود و می‌تواند در کاربردهای متنوعی مورد استفاده قرار گیرد.

دسترسی به کلود

کلود از طریق [وبگاه کلاد](#) آن و همچنین از طریق API قابل دسترسی است. همچنین، برخی از پلتفرم‌ها و برنامه‌ها نیز از API کلود برای ارائه‌ی خدمات خود استفاده می‌کنند.

DeepSeek، پدیده چینی تازه وارد

دیپ‌سیک (DeepSeek) یک مدل زبان بزرگ هوش مصنوعی است که برای پاسخ‌گویی به سؤالات و انجام درخواست‌های کاربران طراحی شده است. این چت‌بات و دستیار مجازی با استفاده از داده‌های گسترده آموزش دیده و عملکرد قابل‌توجهی در تولید محتوای متنی، کدنویسی، ترجمه و تحلیل داده‌ها دارد.

ویژگی‌های برجسته دیپ‌سیک

- **پاسخ‌دهی هوشمند:** قابلیت ارائه پاسخ‌های دقیق و به‌روز به سؤالات پیچیده کاربران با استفاده از منابع معتبر.
- **تولید محتوای باکیفیت:** توانایی تولید متون متنوع از جمله مقالات وبلاگی و پست‌های شبکه‌های اجتماعی با کیفیت بالا و بهینه برای موتورهای جستجو.
- **پشتیبانی از چند زبان:** سازگاری با زبان‌های مختلف، امکان تعامل با کاربران در سراسر جهان را فراهم می‌کند.
- **سرعت پردازش بالا:** نسخه جدید دیپ‌سیک (DeepSeek-V3) با بهبود سرعت استنتاج، تجربه کاربری بهتری را ارائه می‌دهد.
- **قابلیت‌های کدنویسی:** امکان انجام وظایف مرتبط با کدنویسی، ترجمه و نوشتن مقالات و ایمیل‌ها بر اساس درخواست کاربران.

مزایای دیپ‌سیک

- **رابط کاربری ساده:** طراحی کاربرپسند که استفاده از آن را برای کاربران با سطوح مهارتی مختلف آسان می‌کند.
- **دسترسی رایگان:** امکان استفاده رایگان از دیپ‌سیک V3- برای تجربه مدل هوشمند و قدرتمند.
- **ادغام آسان:** قابلیت ادغام با سایر نرم‌افزارها و سرویس‌ها از طریق API، امکان توسعه و استفاده در پروژه‌های مختلف را فراهم می‌کند.

محدودیت‌های دیپ‌سیک

- **وابستگی به اینترنت:** برای استفاده از دیپ‌سیک، اتصال به اینترنت ضروری است.
- **نیاز به نسخه پولی برای امکانات پیشرفته:** برخی قابلیت‌های پیشرفته ممکن است در نسخه‌های پولی در دسترس باشند.
- **احتمال خطا در پاسخ‌گویی:** مانند سایر مدل‌های هوش مصنوعی، امکان ارائه اطلاعات نادرست یا گمراه‌کننده وجود دارد؛ بنابراین، توصیه می‌شود پاسخ‌ها با منابع معتبر مقایسه شوند.

کاربردهای اصلی دیپ‌سیک

کاربردهای اصلی DeepSeek عبارت‌اند از:

- ✓ **پاسخ به سؤالات:** ارائه اطلاعات دقیق و مفید در موضوعات گوناگون مانند علم، فناوری، تاریخ و فرهنگ.
- ✓ **یادگیری و آموزش:** کمک به درک مفاهیم پیچیده، حل مسائل درسی و ارائه توضیحات ساده‌شده.
- ✓ **ترجمه:** ترجمه متون بین زبان‌های مختلف با حفظ معنای اصلی.
- ✓ **تولید متن:** نوشتن مقالات، داستان‌ها، ایمیل‌ها، گزارش‌ها و سایر محتوای متنی.
- ✓ **برنامه‌نویسی:** کمک در نوشتن، اشکال‌زدایی و بهینه‌سازی کدهای برنامه‌نویسی.
- ✓ **خلاقیت:** ارائه ایده‌های جدید برای پروژه‌ها، نام‌گذاری، شعارها و محتوای خلاقانه.
- ✓ **پشتیبانی تصمیم‌گیری:** تحلیل اطلاعات و ارائه پیشنهادات بر اساس داده‌ها.
- ✓ **ویرایش و اصلاح متن:** بررسی و بهبود نوشته‌ها از نظر گرامر، ساختار و سبک.
- ✓ **سرگرمی:** گفت‌وگوهای جذاب، پیشنهاد فیلم، کتاب یا بازی و ایجاد محتوای سرگرم‌کننده.
- ✓ **کسب‌وکار:** کمک در تدوین استراتژی‌های بازاریابی، تحلیل بازار و بهبود ارتباط با مشتریان.
- ✓ **جستجو در وب:** امکان جستجوی آسان در وب و یافتن منابع مورد نیاز.

دسترسی به هوش مصنوعی DeepSeek

DeepSeek یک شرکت هوش مصنوعی چینی است که در ابتدا بر توسعه‌ی مدل‌های زبانی بزرگ (LLM) با تمرکز ویژه بر کدنویسی و تسلط بر زبان‌های برنامه‌نویسی مختلف فعالیت می‌کرد. با این حال، با توجه به دسترسی آنلاین ارائه شده در آدرس <https://chat.deepseek.com>، مشخص می‌شود که DeepSeek دامنه فعالیت خود را گسترش داده و اکنون حوزه‌های بیشتری را پوشش می‌دهد.

بررسی مقایسه‌ای ابزارها

بعد از معرفی ابزارهای هوش مصنوعی مولد متن، برای فهم بیشتر در قالب یک جدول ویژگی‌های این ابزارها را با هم مقایسه می‌کنیم.

ویژگی	Microsoft Copilot	ChatGPT	Google Gemini	Claude
توسعه‌دهنده	Microsoft	Open AI	Google DeepMind	Anthropic
مدل پایه	GPT-۴	GPT-۴	Gemini Ultra	Claude ۲
نوع داده ورودی	متن، لینک، عکس، صوت	متن، لینک، عکس، فایل و صوت	متن، لینک، عکس، صوت	متن، عکس، فایل متنی
انواع مدل	نسخه کوپایلوت معمولی همراه با چت‌جی‌پی‌تی ۴ و مایکروسافت کوپایلوت پرو	GPT-۴ GPT-۴o	Ultra Pro Nano	بر اساس اندازه شامل Haiku و Opus و Sonnet
رابط کاربری	وبگاه اپ دسکتاپی اپ موبایل سایدبار ویندوز ۱۱ مرورگر مایکروسافت اج افزونه مرورگر کروم بات تلگرام	وبگاه اپ موبایل اپ دسکتاپ افزونه مرورگر API	وبگاه API	وبگاه اپ موبایل API

			API	
رایگان و ۲۰ دلار در ماه	رایگان و ۲۰ دلار در ماه	رایگان و ۲۰ دلار در ماه	رایگان و ۲۰ دلار در ماه	قیمت
سریع	متوسط به سریع	بسیار سریع	متوسط به سریع	سرعت پاسخگویی
Slack Zoom Notion پشتیبانی از API	Google Docs Google Sheets Google Drive Gmail Youtube پشتیبانی از API	اتصال به Google Drive پشتیبانی از API	Microsoft Office Microsoft Teams Visual Studio Code و پشتیبانی از API	امکان ادغام با دیگر نرم افزارها
متن، کد	متن، کد، عکس، لینک	متن، لینک، کد، عکس	متن، لینک، کد، عکس	فرمت پاسخها
ندارد	دارد (در ایران فعال نیست)	دارد	دارد	توانایی خلق تصویر
دارد	دارد	دارد	ندارد	نیاز به VPN
حافظه و مدیریت زمینه عالی، به ویژه در مکالمات طولانی	حافظه خوب، بهینه شده برای وظایف پیچیده با داده های متنی و بصری	خوب، اما ممکن است در رشته های بسیار طولانی دچار مشکل شوند	محدود؛ اغلب زمینه مکالمه را تنظیم مجدد می کند	امکان دریافت ورودی های چند مرحله ای

منابع پاسخگویی	استفاده از جستجوی Bing برای اطلاعات بهروز	پاسخهای با کیفیت بالا بر اساس یک مجموعه داده گسترده، اما گاهی اوقات با توهم. دسترسی به جستجو در اینترنت	دقت بالا با استفاده از جستجوی گوگل و منابع داده یکپارچه	نرخ پایین توهم؛ استفاده از مجموعه دادههای گزینش شده گسترده
حوزه تخصصی	یکپارچگی با برنامه‌های Microsoft ۳۶۵ خوب در خلاصه‌سازی و جستجوهای مبتنی بر وب	چندکاره در حوزه‌های مختلف، از جمله وظایف خلاقانه، تحلیلی و تخصصی	مناسب برای مستندسازی فنی و پژوهش	قوی در کدنویسی، استدلال و پرسش‌های پیچیده
دقت و صحت پاسخگویی	دقیق برای جستجوهای مبتنی بر وب؛ به شدت به داده‌های وب متکی است	به طور کلی بالا، به‌ویژه در وظایف مرتبط با زبان	دقت بالا، به‌ویژه در وظایفی که ترکیبی از اطلاعات متنی و بصری هستند	دقت بالا، به‌ویژه در استدلال و پرسش‌های دقیقی.

معرفی ابزارهای تولید تصویر هوش مصنوعی

DALL-E، تولید تصویر از متن

معرفی جامع هوش مصنوعی مولد DALL-E

DALL-E یک مدل هوش مصنوعی مولد است که توسط شرکت OpenAI توسعه داده شده و قادر به تولید تصاویر دیجیتال از توضیحات متنی است. به عبارت دیگر، شما با نوشتن یک متن (که به آن "پرامپت" می‌گویند) می‌توانید از DALL-E بخواهید تصویری مرتبط با آن متن را خلق کند. این توانایی، DALL-E را به ابزاری قدرتمند برای خلق آثار هنری، طراحی، ایده‌پردازی و بسیاری کاربردهای دیگر تبدیل کرده است.

تاریخچه و نسخه‌های DALL-E

- DALL-E ۱: اولین نسخه DALL-E در ژانویه ۲۰۲۱ معرفی شد. این نسخه نشان داد که چگونه یک مدل زبانی بزرگ می‌تواند برای تولید تصاویر از متن استفاده شود، اما کیفیت و دقت تصاویر آن محدود بود.
- DALL-E ۲: نسخه دوم در آوریل ۲۰۲۲ معرفی شد و بهبود چشمگیری در کیفیت، وضوح و دقت تصاویر نسبت به نسخه اول داشت. DALL-E ۲ همچنین قابلیت‌های جدیدی مانند ویرایش تصاویر موجود، ایجاد تغییرات جزئی در تصاویر و ترکیب چندین تصویر را ارائه داد.
- DALL-E ۳: جدیدترین نسخه DALL-E است که در سپتامبر ۲۰۲۳ معرفی شد. این نسخه با بهبود درک پرامپت‌ها و تولید تصاویر واقعی‌تر و با جزئیات بیشتر، جهش بزرگی نسبت به نسخه‌های قبلی داشته است. DALL-E ۳ به طور مستقیم با ChatGPT ادغام شده است، به این معنی که کاربران می‌توانند به راحتی از طریق چت با ChatGPT، تصاویر مورد نظر خود را تولید کنند.

نحوه عملکرد DALL-E

DALL-E بر پایه معماری ترانسفورمر (Transformer) ساخته شده است، همان معماری که در مدل‌های زبانی بزرگی مانند GPT استفاده می‌شود. این مدل با آموزش بر روی مجموعه داده‌های عظیمی از تصاویر و متون مرتبط، یاد می‌گیرد که چگونه بین متن و تصویر ارتباط برقرار کند. به طور خلاصه، DALL-E مراحل زیر را طی می‌کند:

۱. **دریافت پرامپت:** کاربر یک متن (پرامپت) را به DALL-E می‌دهد که تصویری را که می‌خواهد ببیند توصیف می‌کند.
۲. **درک پرامپت:** DALL-E با استفاده از مدل زبانی خود، پرامپت را درک می‌کند و مفاهیم و عناصر کلیدی آن را استخراج می‌کند.
۳. **تولید تصویر:** DALL-E با استفاده از دانش خود از ارتباط بین متن و تصویر، تصویری را تولید می‌کند که با پرامپت مطابقت داشته باشد.

قابلیت‌های کلیدی DALL-E

۱. **تولید تصویر از متن:** اصلی‌ترین قابلیت DALL-E، تولید تصاویر جدید بر اساس توضیحات متنی است.
۲. **ویرایش تصاویر موجود:** DALL-E ۲ و ۳ می‌توانند تصاویر موجود را ویرایش کنند، عناصری را به آن‌ها اضافه یا حذف کنند و تغییرات جزئی در آن‌ها ایجاد کنند.
۳. **تغییرات تصویر (Variations):** DALL-E می‌تواند تغییرات مختلفی از یک تصویر را با حفظ عناصر اصلی آن تولید کند.
۴. **تولید تصاویر با سبک‌های هنری مختلف:** DALL-E می‌تواند تصاویری را با سبک‌های هنری مختلف، مانند نقاشی رنگ روغن، طراحی با مداد، عکاسی و غیره تولید کند.
۵. **ترکیب تصاویر:** DALL-E می‌تواند چندین تصویر را با هم ترکیب کند و تصویر جدیدی ایجاد کند.

کاربردهای DALL-E

- ✓ **خلق آثار هنری:** هنرمندان می‌توانند از DALL-E برای خلق آثار هنری دیجیتال، ایده‌پردازی و آزمایش سبک‌های جدید استفاده کنند.
- ✓ **طراحی گرافیک و تبلیغات:** طراحان گرافیک و متخصصان تبلیغات می‌توانند از DALL-E برای تولید تصاویر تبلیغاتی، لوگو، بنر و سایر موارد بصری استفاده کنند.
- ✓ **طراحی محصول:** طراحان محصول می‌توانند از DALL-E برای تجسم ایده‌های خود و تولید نمونه‌های اولیه استفاده کنند.
- ✓ **آموزش و سرگرمی:** DALL-E می‌تواند در آموزش، برای تجسم مفاهیم پیچیده و همچنین برای سرگرمی و خلق تصاویر خلاقانه استفاده شود.

مزایای استفاده از DALL-E

- **سرعت و کارایی:** DALL-E می‌تواند تصاویر را در عرض چند ثانیه تولید کند، که بسیار سریع‌تر از روش‌های سنتی ایجاد تصویر است.
- **خلاقیت و نوآوری:** DALL-E می‌تواند تصاویری را خلق کند که قبلاً هرگز دیده نشده‌اند و ایده‌های جدیدی را برای کاربران ارائه دهد.
- **دسترسی آسان (به ویژه با ادغام با ChatGPT):** استفاده از DALL-E به لطف رابط کاربری ساده و ادغام آن با ChatGPT بسیار آسان شده است.

محدودیت‌ها و چالش‌ها

- **مسائل اخلاقی:** استفاده از DALL-E می‌تواند مسائل اخلاقی مربوط به مالکیت معنوی، کپی‌رایت و سوءاستفاده از تصاویر را به دنبال داشته باشد.
- **احتمال تولید تصاویر نامناسب یا تبعیض‌آمیز:** اگرچه OpenAI تلاش زیادی برای جلوگیری از تولید تصاویر نامناسب توسط DALL-E انجام داده است، اما همچنان احتمال وقوع این مشکل وجود دارد.
- **نیاز به پرامپت‌های دقیق:** برای دریافت بهترین نتیجه از DALL-E، لازم است پرامپت‌ها دقیق و واضح باشند. نوشتن پرامپت‌های مؤثر می‌تواند یک چالش باشد. حتی ممکن است که استفاده از پرامپت‌های نوشته شده به زبان انگلیسی، بهتر درک شده و خروجی بهتری را به دنبال داشته باشند.
- **محدودیت‌های دسترسی (بسته به نسخه و پلتفرم):** دسترسی به نسخه‌های مختلف DALL-E محدودیت داشته و نیاز به عضویت یا پرداخت هزینه داشته باشد.

دسترسی به DALL-E

DALL-E وبگاهی برای استفاده مستقیم نداشته و در حال حاضر، بهترین و ساده‌ترین راه برای دسترسی به DALL-E ۳ استفاده از ChatGPT Plus است. با خرید اشتراک ChatGPT Plus، شما به DALL-E ۳ نیز دسترسی خواهید داشت و می‌توانید به راحتی از طریق چت با ChatGPT، تصاویر مورد نظر خود را تولید کنید.

در نهایت، DALL-E یک ابزار قدرتمند و نوآورانه است که پتانسیل زیادی برای تغییر نحوه خلق و استفاده از تصاویر دارد. با پیشرفت روزافزون این فناوری، می‌توان انتظار داشت که کاربردهای آن در آینده گسترده‌تر شود.

MidJourney، طراحی تصاویر هنری

موارد استفاده از MidJourney

- ✓ ایجاد تصاویر فانتزی و خلاقانه
- ✓ ایجاد تصاویر آموزشی
- ✓ طراحی تصاویر تبلیغاتی
- ✓ ایجاد تصاویر برای کمپین‌های بازاریابی
- ✓ طراحی شخصیت و کاراکتر برای کتاب یا فیلم
- ✓ طراحی مد و لباس

مزایا و محدودیت‌های کار با میدج‌رنی

مزایا:

- ✓ امکان ایجاد تصاویر رایگان در نسخه غیرپولی
- ✓ استفاده از کارت گرافیک سیستم برای ایجاد تصاویر
- ✓ فهم زبان فارسی در پرامپت برای ایجاد عکس
- ✓ قابلیت ایجاد سرور اختصاصی برای ساخت عکس در نسخه پولی
- ✓ ساخت سریع تصاویر در نسخه پولی

محدودیت‌ها:

- نداشتن نسخه رایگان برای ایجاد تصویر
- دشوار بودن استفاده از ابزار
- قابل استفاده بودن نسخه رایگان تنها از طریق پلتفرم Discord
- سرعت پایین در ایجاد تصویر در نسخه رایگان (۳۰ تا ۶۰ ثانیه)
- تکمیل ظرفیت سرور ایجاد تصویر در نسخه رایگان در برخی از مواقع

میدجرنی: خلق تصاویر شگفت‌انگیز با هوش مصنوعی

میدجرنی (Midjourney) در ابتدا یک ابزار مبتنی بر دیسکورد بود. اما در حال حاضر دارای وبگاه اختصاصی است که با استفاده از دستورات متنی (پرامپت)، تصاویر خلاقانه و چشمگیری تولید می‌کند. کافیست توصیف تصویر مورد نظر خود را به صورت متن وارد کنید تا میدجرنی تصویر متناظر آن را خلق کند.

یکی از ویژگی‌های برجسته میدجرنی، ترکیب دسترسی آسان و کیفیت هنری بالاست. رابط کاربری آن ساده و شبیه به محیط چت است، در عین حال، خروجی تصاویر بسیار با کیفیت و با جزئیات دقیق هستند.

میدجرنی از طریق دیسکورد (و اخیراً وبگاه خود) از یک جامعه‌ی کاربری فعال بهره می‌برد که امکان به اشتراک‌گذاری آثار و دریافت بازخورد را فراهم می‌کند. این ویژگی‌ها میدجرنی را به ابزاری ایده‌آل برای خلق تصاویر بصری جذاب و هنری تبدیل کرده است.

دسترسی بدون دیسکورد

میدجرنی علاوه بر دیسکورد، یک [وبگاه اختصاصی](#) نیز راه‌اندازی کرده که به کاربران امکان تولید تصاویر را بدون نیاز به دیسکورد می‌دهد. البته در حال حاضر، این وبگاه فقط برای کاربرانی با اشتراک فعال و بیش از ۱۰۰ تصویر ساخته شده در حسابشان، در دسترس است.

فرایند ساخت تصویر در میدجرنی

فرایند تولید تصویر در میدجرنی شامل دو مرحله اصلی است:

مرحله اول: آموزش مدل (پشت صحنه):

در این مرحله، که برای کاربران قابل مشاهده نیست، میدجرنی با استفاده از شبکه‌های عصبی مولد مقابله‌ای (GAN) آموزش داده می‌شود. در این روش، دو شبکه عصبی به نام‌های "مولد" و "متمایز کننده" با یکدیگر همکاری می‌کنند. مولد تصاویر جدیدی ایجاد می‌کند و متمایز کننده آن‌ها را با تصاویر واقعی مقایسه می‌کند تا میزان شباهت

آن‌ها به واقعیت را بسنجد. این فرایند تکرار می‌شود تا تصاویر تولید شده با توصیفات ورودی مطابقت بالایی داشته باشند. این مرحله شامل فرایندهای زیر است:

- **جمع‌آوری و آموزش داده‌ها:** مجموعه‌ی عظیمی از تصاویر به سیستم داده می‌شود تا الگوها، رنگ‌ها، اشکال و ساختارها را یاد بگیرد.
- **یادگیری عمیق:** شبکه‌های عصبی مصنوعی از داده‌های ورودی یاد می‌گیرند.
- **تولید تصویر:** مولد تصاویر جدید تولید می‌کند و متمایز کننده آن‌ها را ارزیابی می‌کند.
- **بازخورد انسانی:** بازخورد کاربران به بهبود سیستم کمک می‌کند.

مرحله دوم: تولید تصویر توسط کاربر (در لحظه):

در این مرحله، کاربر توصیف متنی (پرامپت) خود را وارد می‌کند و میدجرنی تصویر متناظر را تولید می‌کند. این فرایند شامل مراحل زیر است:

۱. **توصیف عکس (پرامپت):** کاربر شرح متنی تصویر مورد نظر خود را وارد می‌کند. این شرح می‌تواند ساده یا بسیار دقیق و جزئی باشد.
۲. **پردازش متن:** میدجرنی متن ورودی را تجزیه و تحلیل می‌کند تا مفاهیم، اشیاء، مکان‌ها و حالات موجود در آن را درک کند.
۳. **ساخت و ارائه تصویر:** میدجرنی با استفاده از درک خود از متن و الگوریتم‌های پیچیده، تصویری جدید و منحصر به فرد را تولید می‌کند.

قدرت و تأثیر میدجرنی

میدجرنی به دلیل توانایی خود در تولید تصاویر واقعی و هنری، توجه زیادی را به خود جلب کرده است. نمونه‌ی مشهور آن، تصویر پاپ فرانسیس با ژاکت پف‌کرده بود که در سال ۲۰۲۳ بسیار وایرال شد و قدرت هوش مصنوعی در تولید تصاویر واقعی را به نمایش گذاشت.

نسخه‌های مختلف میدجرنی و تغییرات آن‌ها

میدجرنی از زمان راه‌اندازی خود در سال ۲۰۲۲، نسخه‌های متعددی را منتشر کرده که هر کدام بهبودها و ویژگی‌های جدیدی داشته‌اند:

- **نسخه ۱ (فوریه ۲۰۲۲):** اولین نسخه با قابلیت‌های پایه برای تولید تصاویر ساده.

- **نسخه ۲ (آوریل ۲۰۲۲):** بهبود قابل توجه در کیفیت تصاویر و درک بهتر توصیفات متنی.
- **نسخه ۳ (ژانویه ۲۰۲۲):** تمرکز بیشتر بر تعامل کاربر و امکان دریافت بازخورد. (تاریخ انتشار ظاهراً اشتباه است و باید بعد از نسخه ۲ باشد)
- **نسخه ۴ (نوامبر ۲۰۲۲):** جهش بزرگ در کیفیت تصاویر و واقع‌گرایی آن‌ها.
- **نسخه ۵ (مارس ۲۰۲۳) و زیر نسخه‌های ۵.۱ و ۵.۲ (می ۲۰۲۳ و ژانویه ۲۰۲۳):** تمرکز بیشتر بر واقع‌گرایی و بهبود جزئیات و پاسخ‌دهی به توصیفات. (تاریخ انتشار ژانویه ۲۰۲۳ برای ۵.۲ ظاهراً اشتباه است و باید بعد از می ۲۰۲۳ باشد)
- **نسخه ۶ (دسامبر ۲۰۲۳):** بهبود چشمگیر کیفیت تصاویر، درک بهتر توصیفات ساده، جزئیات بالاتر، تنوع بیشتر در سبک‌ها، افزایش سرعت و تعامل بهتر با کاربر.

در نسخه ۶، تشخیص و درک متن به طور قابل توجهی بهبود یافته و امکان ایجاد تصاویر با جزئیات بسیار بالا فراهم شده است. همچنین، تنوع سبک‌ها و زمینه‌ها افزایش یافته و سرعت پردازش و تعامل با کاربر بهبود یافته است.

نکات مهم:

- میدجرنی در ابتدا یک دوره آزمایشی رایگان محدود ارائه می‌داد، اما در حال حاضر برای استفاده کامل از آن نیاز به خرید اشتراک دارید. می‌توانید پلن‌های اشتراک مختلف را در وبسایت میدجرنی مشاهده کنید.
- برای استفاده از میدجرنی، باید به اینترنت متصل باشید.
- نوشتن پرامپت‌های دقیق و واضح می‌تواند به تولید تصاویر بهتر و دقیق‌تر کمک کند.
- میدجرنی به طور مداوم در حال به‌روزرسانی و بهبود است و ممکن است رابط کاربری و نحوه استفاده از آن کمی تغییر کند.

Firefly، تولید تصویر از متن و فراتر از آن

Firefly خانواده‌ای از مدل‌های هوش مصنوعی مولد است که توسط شرکت ادوبی (Adobe) توسعه داده شده و قادر به تولید تصاویر دیجیتال، جلوه‌های متنی، وکتورها، مدل‌های سه‌بعدی و موارد دیگر از توضیحات متنی (پرامپت) است.

به عبارت دیگر، شما با نوشتن یک متن می‌توانید از Firefly بخواهید تصویری مرتبط با آن متن را خلق کند، متن موجود را به سبک خاصی درآورد، یک تصویر وکتور یا حتی

یک مدل سه‌بعدی ایجاد کند. این توانایی، Firefly را به ابزاری قدرتمند برای خلق آثار هنری، طراحی، ایده‌پردازی، ویرایش تصاویر، تولید محتوای سه‌بعدی و بسیاری کاربردهای دیگر تبدیل کرده است. مزیت اصلی Firefly، ادغام عمیق آن با نرم‌افزارهای محبوب ادوبی مانند فتوشاپ، ایلاستریتور، ادوبی اکسپرس و Substance 3D است که گردش کار طراحان را بسیار تسهیل می‌کند.

تاریخچه و نسخه‌های Firefly

Firefly به صورت یکپارچه در محصولات ادوبی عرضه می‌شود و به طور مداوم در حال به‌روزرسانی است. به جای نسخه‌های مجزا مانند DALL-E یا Stable Diffusion، Firefly به عنوان یک پلتفرم در حال تکامل در نظر گرفته می‌شود که قابلیت‌های جدید به طور مرتب به آن اضافه می‌شوند. ادوبی تمرکز زیادی بر آموزش مدل‌های Firefly بر روی داده‌های دارای مجوز و محتوای عمومی دارد تا از مسائل مربوط به کپی‌رایت جلوگیری کند.

نحوه عملکرد Firefly

Firefly بر پایه مدل‌های یادگیری عمیق و با استفاده از مجموعه داده‌های عظیمی از تصاویر، متون مرتبط و مدل‌های سه‌بعدی آموزش داده شده است. این مدل یاد می‌گیرد که چگونه بین متن و تصویر، متن و وکتور، متن و مدل سه‌بعدی و سایر انواع محتوا ارتباط برقرار کند.

به طور خلاصه، Firefly مراحل زیر را طی می‌کند:

۱. **دریافت پرامپت:** کاربر یک متن (پرامپت) را به Firefly می‌دهد که تصویری، جلوه متنی، عنصر طراحی، وکتور یا مدل سه‌بعدی مورد نظر را توصیف می‌کند.
۲. **درک پرامپت:** Firefly با استفاده از مدل زبانی خود، پرامپت را درک می‌کند و مفاهیم و عناصر کلیدی آن را استخراج می‌کند.
۳. **تولید محتوا:** Firefly با استفاده از دانش خود از ارتباط بین متن و انواع محتوا، خروجی مورد نظر را تولید می‌کند.

قابلیت‌های کلیدی Firefly

- **تولید تصویر از متن (Text to Image):** اصلی‌ترین قابلیت Firefly، تولید تصاویر جدید بر اساس توضیحات متنی است.
- **جلوه‌های متنی (Text Effects):** Firefly می‌تواند متن‌های موجود را با سبک‌ها و جلوه‌های بصری مختلف، مانند بافت‌های سه‌بعدی، سایه‌ها و افکت‌های هنری، تغییر دهد.
- **رنگ‌آمیزی مجدد مولد (Generative Recolor):** در ایلاستریاتور، این قابلیت به شما امکان می‌دهد رنگ آثار هنری وکتور خود را به سرعت با استفاده از پرامپت‌های متنی تغییر دهید.
- **پر کردن مولد (Generative Fill):** در فتوشاپ، این قابلیت به شما امکان می‌دهد بخش‌هایی از تصویر را با محتوای جدید، بر اساس پرامپت‌های متنی، پر کنید یا حذف کنید.
- **گسترش مولد (Generative Expand):** در فتوشاپ، این قابلیت به شما امکان می‌دهد ابعاد تصویر خود را گسترش دهید و Firefly به طور خودکار محتوای جدید را برای پر کردن فضاهای خالی ایجاد می‌کند.
- **تولید وکتور از متن (Text to Vector Graphic):** در ایلاستریاتور، این قابلیت به شما امکان می‌دهد تصاویر وکتور را مستقیماً از پرامپت‌های متنی ایجاد کنید.
- **تولید مدل سه‌بعدی از متن (Text to 3D Model & Texture):** در Substance 3D، این قابلیت‌ها به شما امکان می‌دهند بافت‌ها و مدل‌های سه‌بعدی را با استفاده از پرامپت‌های متنی تولید کنید.

کاربردهای Firefly

- **خلق آثار هنری:** هنرمندان می‌توانند از Firefly برای خلق آثار هنری دیجیتال، ایده‌پردازی و آزمایش سبک‌های جدید استفاده کنند.
- **طراحی گرافیک و تبلیغات:** طراحان گرافیک و متخصصان تبلیغات می‌توانند از Firefly برای تولید تصاویر تبلیغاتی، لوگو، بزر و سایر موارد بصری استفاده کنند.
- **طراحی محصول:** طراحان محصول می‌توانند از Firefly برای تجسم ایده‌های خود و تولید نمونه‌های اولیه استفاده کنند.
- **ویرایش و روتوش عکس:** با قابلیت‌هایی مانند پر کردن و گسترش مولد در فتوشاپ، ویرایش و روتوش عکس‌ها بسیار سریع‌تر و آسان‌تر می‌شود.

- **تولید محتوای سه‌بعدی:** طراحان سه‌بعدی می‌توانند از Firefly برای تولید بافت‌ها، مدل‌ها و صحنه‌های سه‌بعدی استفاده کنند.
- **آموزش و سرگرمی:** Firefly می‌تواند در آموزش، برای تجسم مفاهیم پیچیده و همچنین برای سرگرمی و خلق تصاویر خلاقانه استفاده شود.

مزایای استفاده از Firefly

- **ادغام با نرم‌افزارهای ادوبی:** ادغام عمیق Firefly با نرم‌افزارهای محبوب ادوبی، گردش کار طراحان را بسیار بهبود می‌بخشد.
- **سهولت استفاده:** رابط کاربری Firefly بسیار ساده و کاربرپسند است.
- **سرعت و کارایی:** Firefly می‌تواند تصاویر و جلوه‌ها را در عرض چند ثانیه تولید کند.
- **محتوای قانونی:** ادوبی اعلام کرده است که Firefly بر روی مجموعه داده‌هایی آموزش داده شده است که فاقد مسائل کپی‌رایت هستند و بنابراین محتوای تولید شده توسط آن قانونی است.
- **تنوع در خروجی:** Firefly می‌تواند انواع مختلفی از محتوا از جمله تصویر، متن، وکتور و مدل سه‌بعدی تولید کند.

محدودیت‌ها و چالش‌ها

- **وابستگی به اکوسیستم ادوبی:** برای استفاده از Firefly، معمولاً نیاز به اشتراک در سرویس‌های ادوبی دارید.
- **نیاز به پرامپت‌های دقیق:** برای دریافت بهترین نتیجه از Firefly، لازم است پرامپت‌ها دقیق و واضح باشند.
- **نیاز به سخت‌افزار مناسب (در برخی موارد):** برای استفاده از برخی قابلیت‌های سنگین‌تر Firefly، ممکن است به سخت‌افزار قدرتمندتری نیاز باشد.

دسترسی به Firefly

Firefly به صورت یکپارچه در محصولات ادوبی مانند فتوشاپ، ایلاستریاتور، ادوبی اکسپرس و Substance 3D در دسترس است. برای استفاده از آن، باید یکی از این نسخه‌های به روز این نرم‌افزارها را داشته باشید. همچنین، برخی از قابلیت‌های Firefly به صورت [تحت وب](#) نیز در دسترس هستند.

در نهایت، Firefly یک ابزار قدرتمند و نوآورانه است که پتانسیل زیادی برای تغییر نحوه خلق و استفاده از تصاویر و سایر محتوای بصری و سه‌بعدی دارد. با پیشرفت روزافزون این فناوری و ادغام آن با سایر محصولات ادوبی، می‌توان انتظار داشت که کاربردهای آن در آینده گسترده‌تر شود.

مقایسه بین ابزارهای هوش مصنوعی مولد تصویر

در حال حاضر بهترین رقبای میدجرنی برای ساخت عکس با هوش مصنوعی ابزارهای Firefly و Dall-E هستند. هر دو این ابزارها مدل اختصاصی خود را دارند و در خروجی‌هایی که ارائه می‌دهند می‌توانند با میدجرنی رقابت کنند. با کمک جدولی که در ادامه می‌آید، می‌توانید ویژگی‌های این ابزارها را با یکدیگر مقایسه کنید. این نمودار مقایسه‌ای از عملکرد سه ابزار ساخت عکس با هوش مصنوعی Firefly، Dall-E و Midjourney را نشان می‌دهد. معیارهای مقایسه شامل کیفیت تصویر، سهولت استفاده، قابلیت سفارشی‌سازی و وضوح تصویر هستند. بر اساس این نمودار می‌توان نتایج زیر را کسب کرد:

ویژگی	Midjourney	Dall-E (OpenAI)	Firefly
کیفیت تصاویر	بسیار بالا و واقع‌گرایانه	بسیار بالا و واقع‌گرایانه	متوسط تا بالا
مدل‌های استفاده شده	یادگیری ماشین	ترنسفورمر و GPT-4	مشابه با Dall-E
سرعت تولید	بالا	متوسط روبه‌بالا	متوسط
تعامل با کاربر	پلتفرم Discord، بازخورد کاربران	رابط کاربری مستقل، بازخورد کمتر	رابط کاربری مستقل، بازخورد متوسط
تنوع سبک‌ها	بسیار متنوع	متنوع	کمتر متنوع
سهولت استفاده	آسان و کاربرپسند	نسبتاً آسان	آسان

در ادامه برخی دیگر از ابزارهای هوش مصنوعی مولد تصویر را نیز معرفی می کنیم:

Imagen^۳، تولید کننده تصویر جمینای

Imagen^۳ یک مدل هوش مصنوعی مولد است که توسط شرکت گوگل توسعه داده شده و قادر به تولید تصاویر دیجیتال از توضیحات متنی است. این توانایی، Imagen^۳ را به ابزاری قدرتمند برای خلق آثار هنری، طراحی، ایده پردازی و بسیاری کاربردهای دیگر تبدیل کرده است.

تاریخچه و نسخه های Imagen

- **Imagen:** اولین نسخه Imagen در سال ۲۰۲۲ معرفی شد. این نسخه نشان داد که چگونه یک مدل زبانی بزرگ می تواند برای تولید تصاویر از متن استفاده شود، اما کیفیت و دقت تصاویر آن محدود بود.
- **Imagen^۲:** نسخه دوم با بهبود چشمگیری در کیفیت، وضوح و دقت تصاویر نسبت به نسخه اول ارائه شد.
- **Imagen^۳:** جدیدترین نسخه Imagen است که با بهبود درک پرامپت ها و تولید تصاویر واقعی تر و با جزئیات بیشتر، جهش بزرگی نسبت به نسخه های قبلی داشته است. این نسخه با Gemini ادغام شده است، به این معنی که کاربران می توانند به راحتی از طریق تعامل با Gemini، تصاویر مورد نظر خود را تولید کنند.

نحوه عملکرد Imagen^۳

Imagen^۳ بر پایه معماری ترانسفورمر (Transformer) ساخته شده است، همان معماری که در مدل های زبانی بزرگی مانند مدل های زبانی گوگل استفاده می شود. این مدل با آموزش بر روی مجموعه داده های عظیمی از تصاویر و متون مرتبط، یاد می گیرد که چگونه بین متن و تصویر ارتباط برقرار کند. به طور خلاصه، Imagen^۳ مراحل زیر را طی می کند:

۱. **دریافت پرامپت:** کاربر یک متن (پرامپت) را به Imagen^۳ می دهد که تصویری را که می خواهد ببیند توصیف می کند.

۲. **درک پرامپت:** Imagen۳ با استفاده از مدل زبانی خود، پرامپت را درک می‌کند و مفاهیم و عناصر کلیدی آن را استخراج می‌کند.
۳. **تولید تصویر:** Imagen۳ با استفاده از دانش خود از ارتباط بین متن و تصویر، تصویری را تولید می‌کند که با پرامپت مطابقت داشته باشد.

قابلیت‌های کلیدی Imagen۳

- **تولید تصویر از متن:** اصلی‌ترین قابلیت Imagen۳، تولید تصاویر جدید بر اساس توضیحات متنی است.
- **تولید تصاویر با سبک‌های هنری مختلف:** Imagen۳ می‌تواند تصاویری را با سبک‌های هنری مختلف، مانند نقاشی رنگ روغن، طراحی با مداد، عکاسی و غیره تولید کند.

کاربردهای Imagen۳

- **خلق آثار هنری:** هنرمندان می‌توانند از Imagen۳ برای خلق آثار هنری دیجیتال، ایده‌پردازی و آزمایش سبک‌های جدید استفاده کنند.
- **طراحی گرافیک و تبلیغات:** طراحان گرافیک و متخصصان تبلیغات می‌توانند از Imagen۳ برای تولید تصاویر تبلیغاتی، لوگو، بنر و سایر موارد بصری استفاده کنند.
- **طراحی محصول:** طراحان محصول می‌توانند از Imagen۳ برای تجسم ایده‌های خود و تولید نمونه‌های اولیه استفاده کنند.
- **آموزش و سرگرمی:** Imagen۳ می‌تواند در آموزش، برای تجسم مفاهیم پیچیده و همچنین برای سرگرمی و خلق تصاویر خلاقانه استفاده شود.

مزایای استفاده از Imagen۳

- **سرعت و کارایی:** Imagen۳ می‌تواند تصاویر را در عرض چند ثانیه تولید کند، که بسیار سریع‌تر از روش‌های سنتی ایجاد تصویر است.
- **خلاقیت و نوآوری:** Imagen۳ می‌تواند تصاویری را خلق کند که قبلاً هرگز دیده نشده‌اند و ایده‌های جدیدی را برای کاربران ارائه دهد.
- **دسترسی آسان) به ویژه با ادغام با Gemini):** استفاده از Imagen۳ به لطف ادغام آن با Gemini بسیار آسان شده است.

- **مسائل اخلاقی:** استفاده از Imagen۳ می‌تواند مسائل اخلاقی مربوط به مالکیت معنوی، کپی‌رایت و سوءاستفاده از تصاویر را به دنبال داشته باشد.
- **احتمال تولید تصاویر نامناسب یا تبعیض‌آمیز:** اگرچه گوگل تلاش زیادی برای جلوگیری از تولید تصاویر نامناسب توسط Imagen۳ انجام داده است، اما همچنان احتمال وقوع این مشکل وجود دارد.
- **نیاز به پرامپت‌های دقیق:** برای دریافت بهترین نتیجه از Imagen۳، لازم است پرامپت‌ها دقیق و واضح باشند. نوشتن پرامپت‌های مؤثر می‌تواند یک چالش باشد.

دسترسی به Imagen۳

Imagen۳ از طریق Gemini در دسترس است.

Stable Diffusion، تولید تصویر فراتر از تصور

Stable Diffusion یک مدل هوش مصنوعی مولد است که توسط Stability AI توسعه داده شده و قادر به تولید تصاویر دیجیتال از توضیحات متنی است. برخلاف برخی مدل‌های مشابه، Stable Diffusion متن‌باز است و این امکان را به کاربران می‌دهد تا به صورت گسترده از آن استفاده و آن را توسعه دهند.

تاریخچه و نسخه‌های Stable Diffusion

- **Stable Diffusion ۱.۰:** اولین نسخه Stable Diffusion در سال ۲۰۲۲ منتشر شد. این نسخه نشان داد که چگونه یک مدل انتشار پایدار می‌تواند برای تولید تصاویر با کیفیت بالا از متن استفاده شود و به سرعت محبوبیت زیادی پیدا کرد.
- **نسخه‌های بعدی (مانند ۱.۴، ۱.۵، ۲.۰ و غیره):** نسخه‌های بعدی با بهبود چشمگیری در کیفیت، وضوح، و دقت تصاویر نسبت به نسخه اول ارائه شدند. این نسخه‌ها همچنین قابلیت‌های جدیدی مانند بهبود درک پرامپت‌ها، تولید تصاویر واقعی‌تر و با جزئیات بیشتر، و پشتیبانی از ویژگی‌های جدید را ارائه دادند.

Stable Diffusion بر پایه فرایند انتشار پایدار (Diffusion) و با استفاده از معماری ترانسفورمر (Transformer) ساخته شده است. این مدل با آموزش بر روی مجموعه داده‌های عظیمی از تصاویر و متون مرتبط، یاد می‌گیرد که چگونه بین متن و تصویر ارتباط برقرار کند.

به طور خلاصه، Stable Diffusion مراحل زیر را طی می‌کند:

۱. **دریافت پرامپت:** کاربر یک متن (پرامپت) را به Stable Diffusion می‌دهد که تصویری را که می‌خواهد ببیند توصیف می‌کند.
۲. **رمزگذاری متن:** پرامپت متنی توسط یک رمزگذار متن (Text Encoder) به یک نمایش برداری تبدیل می‌شود.
۳. **فرایند انتشار:** یک تصویر کاملاً نویزی (شبیه به نویز تلویزیون) به تدریج و طی مراحل متعدد، با استفاده از اطلاعات حاصل از رمزگذاری متن، پالایش می‌شود تا به تصویر نهایی تبدیل شود. به عبارت دیگر، نویز به تدریج حذف شده و ساختار تصویر شکل می‌گیرد.
۴. **تولید تصویر:** در نهایت، Stable Diffusion تصویری را تولید می‌کند که با پرامپت مطابقت داشته باشد.

قابلیت‌های کلیدی Stable Diffusion

- **تولید تصویر از متن:** اصلی‌ترین قابلیت Stable Diffusion، تولید تصاویر جدید بر اساس توضیحات متنی است.
- **ویرایش تصاویر موجود:** با استفاده از تکنیک‌هایی مانند Inpainting و Outpainting، می‌توان تصاویر موجود را ویرایش کرده و بخش‌هایی از آن‌ها را تغییر داد یا گسترش داد.
- **تولید تصاویر با سبک‌های هنری مختلف:** Stable Diffusion می‌تواند تصاویری را با سبک‌های هنری مختلف، مانند نقاشی رنگ روغن، طراحی با مداد، عکاسی و غیره تولید کند.
- **تغییرات تصویر (Variations):** می‌توان تغییرات مختلفی از یک تصویر را با حفظ عناصر اصلی آن تولید کرد.

کاربردهای Stable Diffusion

- **خلق آثار هنری:** هنرمندان می‌توانند از Stable Diffusion برای خلق آثار هنری دیجیتال، ایده‌پردازی و آزمایش سبک‌های جدید استفاده کنند.
- **طراحی گرافیک و تبلیغات:** طراحان گرافیک و متخصصان تبلیغات می‌توانند از Stable Diffusion برای تولید تصاویر تبلیغاتی، لوگو، بنر و سایر موارد بصری استفاده کنند.
- **طراحی محصول:** طراحان محصول می‌توانند از Stable Diffusion برای تجسم ایده‌های خود و تولید نمونه‌های اولیه استفاده کنند.
- **بازی‌سازی و توسعه محتوای دیجیتال:** Stable Diffusion می‌تواند برای تولید بافت‌ها، مدل‌ها و سایر عناصر بصری در بازی‌ها و سایر پروژه‌های دیجیتال استفاده شود.
- **آموزش و سرگرمی:** Stable Diffusion می‌تواند در آموزش، برای تجسم مفاهیم پیچیده و همچنین برای سرگرمی و خلق تصاویر خلاقانه استفاده شود.

مزایای استفاده از Stable Diffusion

- **متن‌باز بودن:** Stable Diffusion یک پروژه متن‌باز است که به کاربران امکان می‌دهد به کد منبع آن دسترسی داشته باشند، آن را تغییر دهند و به صورت رایگان از آن استفاده کنند.
- **انعطاف‌پذیری:** Stable Diffusion بسیار انعطاف‌پذیر است و می‌توان آن را روی سخت‌افزارهای مختلف، از جمله کامپیوترهای شخصی با کارت گرافیک مناسب، اجرا کرد.
- **سرعت و کارایی:** Stable Diffusion می‌تواند تصاویر را نسبتاً سریع تولید کند، به خصوص با سخت‌افزار مناسب.
- **جامعه کاربری فعال:** Stable Diffusion دارای یک جامعه کاربری بسیار فعال است که منابع، آموزش‌ها و ابزارهای مختلفی را برای کاربران فراهم می‌کنند.

محدودیت‌ها و چالش‌ها

- **نیاز به سخت‌افزار مناسب:** برای اجرای بهینه Stable Diffusion، به یک کارت گرافیک با حافظه VRAM کافی نیاز است.
- **نیاز به پرامپت‌های دقیق:** برای دریافت بهترین نتیجه از Stable Diffusion، لازم است پرامپت‌ها دقیق و واضح باشند. نوشتن پرامپت‌های مؤثر می‌تواند یک چالش باشد.

- **مسائل اخلاقی:** استفاده از Stable Diffusion می‌تواند مسائل اخلاقی مربوط به مالکیت معنوی، کپی‌رایت و سوءاستفاده از تصاویر را به دنبال داشته باشد.

دسترسی به Stable Diffusion

Stable Diffusion به صورت متن‌باز در دسترس است و می‌توان آن را از منابع مختلفی مانند GitHub دانلود کرد. همچنین، وبسایت‌ها و پلتفرم‌های مختلفی وجود دارند که امکان استفاده آنلاین از Stable Diffusion را فراهم می‌کنند.

مقایسه چت‌بات‌ها، موتورهای جستجوی سنتی و موتورهای جستجوی هوش مصنوعی

در دنیای دیجیتال امروز، یافتن اطلاعات به سرعت و به آسانی از اهمیت بالایی برخوردار است. سه ابزار اصلی برای این منظور وجود دارند: موتورهای جستجوی سنتی (مانند گوگل)، چت‌بات‌ها و موتورهای جستجوی هوش مصنوعی. هر کدام از این ابزارها ویژگی‌ها، نقاط قوت و ضعف خاص خود را دارند. در این بخش به بررسی تفاوت‌های اساسی آن‌ها می‌پردازیم:

۱. موتورهای جستجوی سنتی (مانند گوگل و بینگ)

- **نحوه عملکرد:** این موتورها با استفاده از الگوریتم‌های پیچیده، صفحات وب را بررسی و فهرست‌بندی می‌کنند. وقتی شما عبارتی را جستجو می‌کنید، موتور جستجو لیستی از صفحات وب مرتبط را بر اساس میزان ارتباط و اعتبار آن‌ها به شما نمایش می‌دهد.
- **نوع پاسخ‌ها:** پاسخ‌ها به صورت لیستی از لینک‌های وب‌سایت‌ها، تصاویر، ویدئوها و سایر انواع محتوا نمایش داده می‌شوند. شما باید خودتان در این صفحات به دنبال پاسخ مورد نظر بگردید.
- **منابع اطلاعاتی:** منبع اصلی اطلاعات این موتورها، صفحات وب عمومی هستند.
- **تعامل:** تعامل با این موتورها معمولاً به صورت وارد کردن کلمات کلیدی و دریافت لیستی از نتایج است. تعامل چندانی به صورت محاوره‌ای وجود ندارد.
- **کاربردها:** پیدا کردن اطلاعات عمومی، تحقیق، خرید آنلاین، یافتن آدرس و شماره تلفن و غیره.
- **مثال:** گوگل، بینگ، یاهو

۲. چت‌بات‌ها مانند Gemini، ChatGPT

- **نحوه عملکرد:** چت‌بات‌ها با استفاده از مدل‌های زبانی بزرگ (LLM) آموزش داده شده‌اند تا بتوانند زبان انسان را درک کنند و به صورت محاوره‌ای پاسخ دهند. آن‌ها می‌توانند به سؤالات پاسخ دهند، متن تولید کنند، کد بنویسند و کارهای دیگری را نیز انجام دهند.

- **نوع پاسخ‌ها:** پاسخ‌ها به صورت متن، کد یا سایر فرمت‌های دیجیتال در قالب یک مکالمه ارائه می‌شوند. چت‌بات‌ها سعی می‌کنند پاسخی مستقیم و جامع به سؤال شما بدهند.
- **منابع اطلاعاتی:** چت‌بات‌ها دانش خود را از مجموعه داده‌های عظیمی که در طول آموزش دیده‌اند، به دست می‌آورند. این داده‌ها می‌تواند شامل متن‌های وب، کتاب‌ها، کدها و غیره باشد. برخی چت‌بات‌ها مانند Gemini به اطلاعات لحظه‌ای از وب نیز دسترسی دارند و دیگران نیز در حال افزودن این امکان به خودشان هستند.
- **تعامل:** تعامل با چت‌بات‌ها به صورت محاوره‌ای و شبیه به گفتگو با یک انسان است. شما می‌توانید سؤالات خود را به زبان طبیعی بپرسید و پاسخ‌های مرتبط را دریافت کنید.
- **کاربردها:** پاسخ به سؤالات، تولید محتوا، کمک در برنامه‌نویسی، ارائه ایده‌های خلاقانه، آموزش و غیره.
- **مثال:** Claude، Gemini، ChatGPT

۳. موتورهای جستجوی هوش مصنوعی مانند Perplexity AI

- **نحوه عملکرد:** این موتورهای ترکیبی از ویژگی‌های موتورهای جستجوی سنتی و چت‌بات‌ها هستند. آن‌ها از هوش مصنوعی برای درک بهتر سؤالات کاربر و ارائه پاسخ‌های دقیق‌تر و مرتبط‌تر استفاده می‌کنند. این موتورها علاوه بر ارائه لینک به منابع، سعی می‌کنند پاسخی مستقیم به سؤال کاربر بدهند و منابع مورد استفاده را نیز ذکر می‌کنند.
- **نوع پاسخ‌ها:** پاسخ‌ها هم به صورت متن و هم به صورت لینک به منابع ارائه می‌شوند. تمرکز بر ارائه پاسخ مستقیم به سؤال کاربر و همچنین شفافیت در مورد منابع اطلاعاتی است.
- **منابع اطلاعاتی:** این موتورها از ترکیبی از منابع وب و مدل‌های زبانی بزرگ استفاده می‌کنند.
- **تعامل:** تعامل با این موتورها معمولاً به صورت پرسش و پاسخ است، اما به اندازه چت‌بات‌ها محاوره‌ای نیست.
- **کاربردها:** تحقیق، پیدا کردن اطلاعات دقیق و معتبر، پاسخ به سؤالات پیچیده، خلاصه‌سازی اطلاعات.
- **مثال:** Perplexity AI

ویژگی	موتور جستجوی سنتی	چت‌بات‌ها	موتور جستجوی هوش مصنوعی
نحوه عملکرد	فهرست‌بندی وب	مدل‌های زبانی بزرگ	ترکیبی از وب و LLM
نوع پاسخ‌ها	لینک‌ها	متن، کد، محاوره	متن + لینک
منابع اطلاعاتی	صفحات وب	داده‌های آموزشی + وب (گاهی)	وب + داده‌های آموزشی
تعامل	کلمات کلیدی	محاوره‌ای	پرسش و پاسخ
کاربردها	اطلاعات عمومی	تولید محتوا، پاسخ به سؤالات	تحقیق، اطلاعات دقیق

تفاوت در تعامل

مهم‌ترین تفاوت در نحوه تعامل با این سه دسته ابزار است. در موتورهای جستجوی سنتی، شما با وارد کردن کلمات کلیدی، لیستی از لینک‌ها را دریافت می‌کنید. در چت‌بات‌ها، تعامل به صورت محاوره‌ای و شبیه به گفتگو با یک انسان است. در موتورهای جستجوی هوش مصنوعی، ترکیبی از این دو رویکرد وجود دارد.

کاربردهای تخصصی و عمومی

موتورهای جستجوی سنتی برای یافتن اطلاعات عمومی و انجام کارهای روزمره بسیار مناسب هستند. چت‌بات‌ها در تولید محتوا، پاسخ به سؤالات پیچیده و انجام کارهای خلاقانه کاربرد بیشتری دارند. موتورهای جستجوی هوش مصنوعی برای تحقیقات تخصصی و یافتن اطلاعات دقیق و معتبر مناسب‌تر هستند.

محدودیت‌ها و چالش‌های استفاده از ابزارهای هوش مصنوعی مولد

محدودیت‌ها و چالش‌های استفاده از ابزارهای هوش مصنوعی مولد از زاویه کاربری (یعنی دیدگاه کسی که از این ابزارها برای انجام کارهای خود استفاده می‌کند) به شرح زیر است:

۱. محدودیت‌های مربوط به کیفیت و صحت خروجی

- خروجی‌های نامناسب یا بی‌ربط: گاهی اوقات، ابزارهای هوش مصنوعی مولد خروجی‌هایی تولید می‌کنند که با درخواست کاربر همخوانی ندارند یا کیفیت مطلوبی ندارند.
 - مثال: از یک ابزار تولید تصویر می‌خواهید تصویری از یک "گره با کلاه آبی" تولید کند، اما تصویر گره بدون کلاه یا با کلاهی به رنگ دیگر تولید می‌شود.
- اطلاعات نادرست یا گمراه‌کننده: ابزارهای تولید متن ممکن است اطلاعات نادرست، قدیمی یا بی‌اساس را به عنوان پاسخ ارائه دهند.
 - مثال: از یک چت‌بات در مورد تاریخچه یک رویداد خاص سؤال می‌کنید و پاسخ دریافتی شامل اطلاعات اشتباه یا تحریف شده است.
- عدم انسجام و پیوستگی در متون طولانی: در تولید متون طولانی، ممکن است ابزار نتواند انسجام و پیوستگی لازم را حفظ کند و متن دچار گسستگی یا تناقض شود.
 - مثال: در نوشتن یک داستان بلند با استفاده از هوش مصنوعی، ممکن است شخصیت‌ها یا خط داستانی به طور ناگهانی تغییر کنند یا تناقضاتی در داستان به وجود آید.
- مشکل در درک ظرافت‌های زبانی و فرهنگی: ابزارهای هوش مصنوعی ممکن است در درک ظرافت‌های زبانی، اصطلاحات، کنایه‌ها و مفاهیم فرهنگی دچار مشکل شوند و پاسخ‌های نامناسب یا بی‌دقتی ارائه دهند.
 - مثال: یک عبارت طنزآمیز را به ابزار می‌دهید و آن را به صورت جدی تفسیر می‌کند.

۲. محدودیت‌های مربوط به کنترل و شخصی‌سازی

- عدم کنترل کامل بر خروجی: کاربر نمی‌تواند به طور کامل بر خروجی ابزار کنترل داشته باشد و ممکن است نتواند دقیقاً به نتیجه دلخواه خود برسد.
 - مثال: در استفاده از ابزارهای تولید موسیقی، ممکن است نتوانید دقیقاً ملودی یا ریتمی را که در ذهن دارید، تولید کنید.
- مشکل در ویرایش و اصلاح خروجی: ویرایش و اصلاح خروجی‌های تولید شده توسط هوش مصنوعی می‌تواند زمان‌بر و دشوار باشد، به خصوص در متون طولانی یا تصاویر پیچیده.
 - مثال: در یک تصویر تولید شده توسط هوش مصنوعی، می‌خواهید یک شیء کوچک را تغییر دهید، اما انجام این کار به صورت دقیق و بدون افت کیفیت دشوار است.
- وابستگی به ابزار: استفاده بیش از حد به این ابزارها می‌تواند باعث کاهش خلاقیت و مهارت‌های فردی شود.

۳. چالش‌های مربوط به حریم خصوصی و امنیت

- نگرانی‌های مربوط به حریم خصوصی داده‌ها: برای استفاده از برخی از این ابزارها، ممکن است لازم باشد اطلاعات شخصی یا داده‌های حساس خود را با آنها به اشتراک بگذارید که می‌تواند نگرانی‌هایی در مورد حریم خصوصی ایجاد کند.
 - مثال: برای استفاده از یک ابزار تولید تصویر، ممکن است لازم باشد تصاویری از خود یا دیگران را آپلود کنید.
- احتمال سوءاستفاده از داده‌ها: ممکن است داده‌های شما توسط شرکت‌های ارائه‌دهنده این ابزارها برای اهداف دیگری مانند آموزش مدل‌های جدید یا تبلیغات استفاده شود.

۴. چالش‌های مربوط به دسترسی و هزینه

- هزینه بالای برخی از ابزارها: برخی از ابزارهای پیشرفته هوش مصنوعی مولد، به خصوص آنهایی که دارای قابلیت‌های حرفه‌ای هستند، هزینه بالایی دارند و برای همه کاربران قابل دسترسی نیستند.

- نیاز به اینترنت پرسرعت و دستگاه‌های قدرتمند: استفاده از برخی از این ابزارها نیازمند اینترنت پرسرعت و دستگاه‌های قدرتمند با پردازنده‌های گرافیکی قوی است.

۵. چالش‌های مربوط به اخلاق و مسئولیت‌پذیری

- استفاده غیراخلاقی از ابزارها: این ابزارها می‌توانند برای تولید محتوای جعلی، همراه‌کننده یا مضر استفاده شوند.
 - مثال: تولید اخبار جعلی، تصاویر غیراخلاقی یا متون توهین‌آمیز.
- مشکل در تشخیص محتوای تولید شده توسط هوش مصنوعی: تشخیص اینکه یک محتوا توسط انسان تولید شده یا توسط هوش مصنوعی، روزبه‌روز دشوارتر می‌شود که می‌تواند مشکلاتی را در زمینه‌های مختلف مانند آموزش، رسانه و حقوق ایجاد کند.

نکاتی برای کاربران:

- همیشه خروجی‌های ابزارهای هوش مصنوعی را با دقت بررسی کنید و به آنها به عنوان منبع قطعی اطلاعات اعتماد نکنید.
 - در مورد مسائل مربوط به حریم خصوصی و امنیت داده‌های خود آگاه باشید.
 - از این ابزارها به صورت مسئولانه و اخلاقی استفاده کنید.
 - به یاد داشته باشید که این ابزارها، ابزاری برای کمک به شما هستند و نباید جایگزین خلاقیت و تفکر انسانی شوند.
- با آگاهی از این محدودیت‌ها و چالش‌ها، کاربران می‌توانند از ابزارهای هوش مصنوعی مولد به صورت مؤثرتر و مسئولانه‌تر استفاده کنند.

