

اسکرپ وبسایت با هوش مصنوعی Thunderbit

این مقاله یک راهنمای عملی است که به کاربران نشان می‌دهد چگونه می‌توانند اسکرپ وبسایت با هوش مصنوعی را به‌سادگی و در کوتاه‌ترین زمان ممکن انجام دهند. با استفاده از این روش، امکان استخراج اطلاعات از صفحات مختلف وب، از پروفایل‌های شبکه‌های اجتماعی گرفته تا فروشگاه‌های آنلاین و نتایج املاک، تنها با چند کلیک فراهم می‌شود.

تمام نکات، ترفندها و مثال‌های ارائه‌شده در این مقاله بر اساس تجربه عملی تهیه شده‌اند تا کاربران بتوانند به‌صورت موثر و سریع، با اسکرپ وبسایت با هوش مصنوعی آشنا شوند. ابزارهایی مانند Thunderbit نیز به‌عنوان نمونه‌ای کاربردی برای اجرای این فرایند معرفی شده‌اند.

اسکرپ وبسایت چیست و چه کاربردهایی دارد؟

اسکرپ وبسایت به فرایند خودکار جمع‌آوری داده‌ها از صفحات اینترنتی گفته

می‌شود. در این روش، نرم‌افزارها یا ابزارهای ویژه، محتوای یک وبسایت را بررسی و اطلاعات موردنظر را بدون آنکه نیاز به انجام دستی باشد، استخراج می‌کنند.

این داده‌ها می‌توانند شامل موارد زیر باشند:

- متن
- تصاویر
- قیمت‌ها
- مشخصات محصولات
- اطلاعات تماس
- اخبار
- نوع محتوای ساختاریافته در صفحات وب

کاربردهای اسکریپ وبسایت

کاربردهای اسکریپ وبسایت گسترده و متنوع است:

تحلیل بازار و رقابت

شرکت‌ها با جمع‌آوری قیمت‌ها، مشخصات محصولات و نظرات مشتریان از وبسایت‌های رقیب، می‌توانند استراتژی قیمت‌گذاری و بازاریابی خود را بهینه کنند.

جمع‌آوری لید و اطلاعات مشتریان

کسب‌وکارها می‌توانند با استخراج داده از پروفایل‌ها یا فرم‌های آنلاین، فهرست‌های تماس و مشتریان بالقوه بسازند.

تحقیق و تحلیل داده‌های عمومی

پژوهشگران و تحلیل‌گران می‌توانند داده‌های خبری، آمارهای عمومی و اطلاعات علمی را به صورت خودکار جمع‌آوری و تحلیل کنند.

اتوماسیون و بهینه‌سازی فرایندها

با اسکریپ وبسایت، سازمان‌ها می‌توانند فرایندهای تکراری جمع‌آوری داده را خودکار

کرده و منابع انسانی و زمان را صرفه‌جویی کنند.

با پیشرفت هوش مصنوعی و یادگیری ماشین، اسکریپ وبسایت با هوش مصنوعی امکان تشخیص الگوهای پیچیده، پیشنهاد خودکار ستون‌های داده و استخراج اطلاعات از صفحات پویا و چندلایه را فراهم کرده است. این روش، دقت استخراج را افزایش می‌دهد و توانایی جمع‌آوری حجم بالای داده‌ها را بدون خطاهای انسانی امکان‌پذیر می‌سازد.

در نتیجه، اسکریپ وبسایت با هوش مصنوعی، ابزاری حیاتی برای تحلیل داده، تصمیم‌گیری استراتژیک و اتوماسیون در دنیای دیجیتال امروز محسوب می‌شود.

ابزار Thunderbit

در فرایندهای سنتی اسکریپ وبسایت، کاربران اغلب با چالش‌های متعددی از جمله نیاز به تنظیمات پیچیده، نوشتن اسکریپت‌های اختصاصی، مدیریت پروکسی‌ها و کنترل صفحات چندلایه یا چندصفحه‌ای مواجه می‌شوند. این مسائل باعث می‌شوند انجام استخراج داده‌ها زمان‌بر و خطاپذیر باشد.

ابزار Thunderbit با ارائه قابلیت‌های اسکریپ وبسایت با هوش مصنوعی این فرایند را ساده کرده است و امکان استخراج داده‌ها را به صورت «یک کلیک، دو کلیک» فراهم می‌آورد.

با این ابزار، کاربران بدون داشتن مهارت فنی قادر هستند کارهای متنوعی از جمله موارد زیر را خیلی راحت انجام دهند:

- ساختار صفحات وب را شناسایی کنند.
- ستون‌های داده را به صورت خودکار پیشنهاد بگیرند.
- خروجی‌ها را به صورت فایل CSV یا مستقیم به سرویس‌هایی مانند Google Sheets، Airtable و Notion ارسال نمایند.

این ویژگی‌ها، Thunderbit را به یک راهکار سریع و موثر برای اجرای اسکریپ و وبسایت با هوش مصنوعی تبدیل کرده و بسیاری از پیچیدگی‌ها و محدودیت‌های روش‌های سنتی را برطرف می‌کند.

نصب و راه‌اندازی در ۲ دقیقه

مراحل کوتاه و مشخص هستند:

- به وبسایت thunderbit.com بروید.
- اکستنشن Chrome را نصب کنید (Add to Chrome).
- آیکون اکستنشن در گوشه مرورگر ظاهر می‌شود. با کلیک، پنل اسکریپر باز می‌شود.

Thunderbit امکان استفاده رایگان تا سقف ۶ صفحه را فراهم می‌کند که این محدودیت برای تست و تمرین کاربران کافی است. پس از نصب، آماده اجرای ۳ دمو سریع و عملی می‌باشد.

۳ دمو سریع

۱) اسکریپ پروفایل‌ها یا صفحات اجتماعی

فرض کنید صفحه‌ای از نام، عنوان شغلی، تصویر و اطلاعات تماس دارید. با کلیک روی اکستنشن و انتخاب AI web scraper، گزینه AI suggested columns را می‌زنید. Thunderbit به‌صورت خودکار ستون‌هایی مثل نام کامل، عنوان شغلی و تصویر پروفایل را تشخیص می‌دهد.

سپس دکمه Scrape را بزنید و در عرض چند ثانیه داده‌ها را دانلود کنید.

۲) اسکرپ فروشگاه‌های آنلاین با صفحات متعدد (pagination)

برای سایت‌هایی مانند آمازون یا هر فروشگاه که نتایج در چند صفحه مرتب شده‌اند، Thunderbit گزینه «Pagination» دارد. کافی است مراحل زیر را پیش بگیرید:

- گزینه Pagination را فعال کنید و حداکثر صفحات را تنظیم کنید (مثلاً ۵ یا تا ۲۰۰).
- قسمت مربوط به دکمه «صفحه بعدی» را نشانه‌گذاری کنید تا اسکرپ بدانند چگونه به صفحه بعد برود.
- AI suggested columns را انتخاب کنید تا نام محصول، قیمت، امتیاز و تصویر شناسایی شوند.

این کار باعث می‌شود بدون کلیک دستی روی هر صفحه، همه نتایج استخراج شوند. من در مثال عملی بیش از ۲۰۰ کالا همراه تصویر، قیمت و تعداد نقدها استخراج کردم.

۳) اسکرپ مقالات خبری

برای استخراج اطلاعات از مقالات خبری، کاربران می‌توانند از اسکرپ وبسایت با هوش مصنوعی استفاده کنند. Thunderbit قالب‌های آماده‌ای برای این منظور ارائه می‌دهد. مراحل استفاده به شرح زیر است:

۱. انتخاب قالب News Scraper از داخل ابزار
۲. کلیک روی دکمه Scrape برای شروع فرایند استخراج
۳. استخراج خودکار اطلاعات شامل:

- عنوان مقاله
- نام نویسنده
- تاریخ انتشار
- محتوای خلاصه مقاله

۴. دریافت خروجی به صورت فایل CSV برای ذخیره یا تحلیل‌های بعدی

با استفاده از این روش، کاربران می‌توانند خیلی سریع داده‌های مقالات خبری را جمع‌آوری کرده و از آن‌ها برای تحلیل محتوا و تصمیم‌گیری‌های استراتژیک بهره‌مند شوند.

قالب‌ها و اسکریپت‌های آماده

یکی از ویژگی‌های کلیدی Thunderbit، ارائه مجموعه‌ای از قالب‌های از پیش‌ساخته برای اسکریپ و وبسایت با هوش مصنوعی است. کاربران می‌توانند از میان قالب‌های استاندارد مانند LinkedIn Job Scraper یا قالب‌های مربوط به اینستاگرام، آمازون و Zillow انتخاب کنند.

نمونه: جستجوی شغل در LinkedIn

برای استخراج داده‌های شغلی در حوزه اتوماسیون AI، مراحل زیر پیشنهاد می‌شود:

۱. جستجو در LinkedIn برای موقعیت‌های شغلی
۲. اعمال فیلترهای مناسب مانند Remote یا مکان مشخص
۳. انتخاب قالب LinkedIn Job Scraper در Thunderbit
۴. استخراج خودکار اطلاعات شامل:

- عنوان شغل
- نام شرکت
- مکان شغل
- لینک آگهی شغلی

با این روش، کاربران می‌توانند بدون نیاز به کلیک دستی روی هر آگهی، داده‌های موردنیاز را جمع‌آوری کرده و از آن‌ها برای تحلیل بازار کار یا برنامه‌ریزی استراتژیک استفاده کنند.

استفاده از Thunderbit باعث می‌شود فرایند اسکریپ وبسایت با هوش مصنوعی سریع، دقیق و کارآمد باشد.

نمونه: رقابت قیمتی و تحلیل محصول در Amazon

فروشنندگان می‌توانند با استفاده از اسکریپ وبسایت با هوش مصنوعی، اطلاعات قیمتی، نام محصولات و امتیازهای کاربران را از صفحات متعدد وبسایت‌هایی مانند Amazon استخراج کنند.

این داده‌ها امکان ایجاد یک Dataset رقابتی را فراهم می‌آورند و به کسب‌وکارها اجازه می‌دهند قیمت‌گذاری خود را بر اساس داده‌های بازار در زمان واقعی (real-time market data) تنظیم کنند.

برای مثال، در جستجوی «digital watch men»، بیش از ۲۰۰ محصول استخراج شد که شامل موارد زیر بود:

- تصویر محصول
- قیمت
- امتیاز (Rating)
- تعداد نقدها

نمونه: جستجوی املاک در Zillow

برای استخراج اطلاعات املاک، کاربران می‌توانند فیلترهای دقیق مانند تعداد اتاق‌خواب و حداکثر قیمت را در خود سایت اعمال کنند و سپس قالب مربوطه را در Thunderbit انتخاب نمایند. با استفاده از اسکریپ وبسایت با هوش مصنوعی، داده‌های زیر به‌صورت خودکار استخراج می‌شوند:

- لیست آگهی‌ها
- قیمت هر ملک
- مشخصات ملک

• لینک هر آگهی

این روش به‌ویژه برای سرمایه‌گذاران و مشاوران املاک مفید است و امکان تحلیل سریع بازار و تصمیم‌گیری هوشمندانه را فراهم می‌کند.

سفارشی‌سازی پیشرفته اسکرپ

Thunderbit صرفاً یک ابزار ساده برای کلیک و استخراج داده نیست؛ بلکه امکانات پیشرفته‌ای برای اسکرپ و وبسایت با هوش مصنوعی ارائه می‌دهد که امکان شخصی‌سازی و انجام عملیات ویژه را فراهم می‌کند.

از جمله قابلیت‌های پیشرفته می‌توان به موارد زیر اشاره کرد:

اسکرپ زیرصفحه‌ها (Subpages)

کاربران می‌توانند از یک لیست پروفایل‌ها وارد هر پروفایل شده و فیلدهای خاص مانند ایمیل یا لینک‌های شبکه‌های اجتماعی را از صفحات داخلی استخراج کنند. این قابلیت برای تولید لیدهای فروش بسیار موثر است.

اسکرپ از URLها

به‌جای استخراج از صفحه فعلی، می‌توان یک لیست از URLها ارائه کرد و اسکرپ به‌صورت خودکار روی هرکدام اجرا شود. این ویژگی برای اتوماسیون‌های زمان‌بندی شده بسیار مناسب است.

استخراج از فایل‌ها و تصاویر

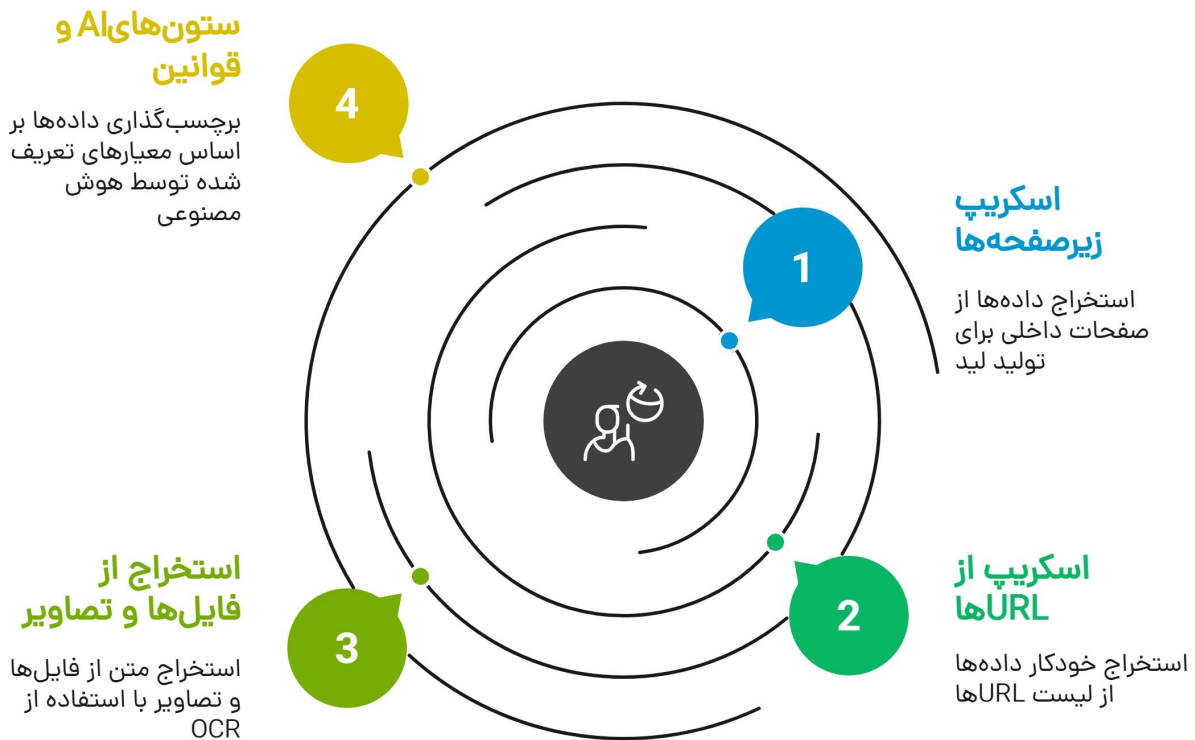
در صورتی‌که اطلاعات در قالب عکس یا PDF باشند، Thunderbit با استفاده از فناوری OCR متن موردنظر را استخراج می‌کند؛ به‌عنوان مثال: «لیست نویسندگان یک مقاله چاپی».

ستون‌های AI و قوانین (ICP Tagging)

کاربران می‌توانند به هوش مصنوعی دستور دهند هر ردیف را بر اساس معیارهای

مشخص برچسب‌گذاری کند؛ به‌عنوان مثال: «شناسایی شرکت‌های SMB در حوزه AI با بیش از ۵ کاربر برای تشخیص مشتریان ایده‌آل».

در عمل، با افزودن ستون «ICP» و تعریف قواعد ساده، خروجی داده‌ها طوری تنظیم شد که هر ردیف با Y یا N مشخص شود. این فرایند باعث سرعت و دقت بیشتر در پرورش لید می‌شود. استفاده از این قابلیت‌ها، اسکرپ و بسایت با هوش مصنوعی را به یک ابزار قدرتمند برای تحلیل و جمع‌آوری داده‌های هدفمند تبدیل می‌کند.



خروجی‌ها

پس از فرایند اسکرپ و بسایت با هوش مصنوعی، داده‌های استخراج شده می‌توانند به روش‌های زیر ذخیره یا استفاده شوند:

- دانلود داده‌ها به‌صورت فایل CSV برای ذخیره محلی یا تحلیل‌های بعدی
- ارسال مستقیم داده‌ها به Airtable، Google Sheets یا Notion با اتصال حساب

کاربری به Thunderbit

• امکان به‌روزرسانی خودکار داده‌ها در پایگاه‌های داده کاری تیم‌های فروش و عملیات

این قابلیت باعث می‌شود تیم‌ها بتوانند اطلاعات استخراج شده را بدون نیاز به مراحل دستی، مستقیم وارد سیستم‌های کاری خود کنند و از مزایای اسکریپ وبسایت با هوش مصنوعی برای تصمیم‌گیری سریع و دقیق بهره‌مند شوند.

تحلیل داده‌ها با ChatGPT

پس از ایجاد یک Dataset تمیز با استفاده از اسکریپ وبسایت با هوش مصنوعی، مرحله بعدی تحلیل داده‌ها و استخراج بینش‌های عملی است. روند معمول تحلیل شامل مراحل زیر است:

۱. دانلود دیتای CSV

داده‌های استخراج شده از Thunderbit به‌صورت CSV ذخیره می‌شوند.

۲. استفاده از ChatGPT یا ابزار تحلیل متصل

با بهره‌گیری از پرامپت‌های آماده، گراف‌ها، روندها و نقاط خارج از قاعده شناسایی می‌شوند.

۳. تصمیم‌گیری بر اساس تحلیل

نتایج حاصل از تحلیل برای تصمیم‌گیری‌های قیمتی، انتخاب محصولات یا هدف‌گذاری لیدها به‌کار گرفته می‌شوند.

برای مثال، در مجموعه ساعت‌های Casio که با Thunderbit اسکریپ شده بود، یک پرامپت تحلیلی چهار نمودار ارائه داد که شامل موارد زیر بود:

• روند تغییرات قیمت

• ارتباط بین قیمت و امتیاز (Rating)

• توزیع تعداد نقدها

• شناسایی محصولات outlier

این تحلیل‌ها به‌طور فوری قابل استفاده برای تعیین استراتژی قیمت‌گذاری و تصمیم‌گیری هوشمندانه در بازار هستند.

نکات و بهترین روش‌ها برای اسکریپ وبسایت با هوش مصنوعی

برای بهره‌وری حداکثری و کاهش ریسک‌ها هنگام اسکریپ وبسایت با هوش مصنوعی، رعایت نکات زیر توصیه می‌شود:

۱. بررسی قوانین و سیاست‌های سایت

قبل از شروع استخراج داده، قوانین استفاده از داده و سیاست‌های سایت را مطالعه کنید تا با مشکلات حقوقی مواجه نشوید.

۲. استفاده از Pagination و محدودیت‌گذاری

برای حجم‌های بزرگ داده، از تنظیمات Pagination و محدودیت صفحات استفاده کنید تا ریسک بلاک شدن کاهش یابد.

۳. بازاستفاده از قالب‌ها

یک‌بار قالب مناسب ایجاد کنید و برای صفحات مشابه از همان قالب استفاده نمایید؛ این کار زمان بر نیست ولی کارایی استخراج داده را به‌طور چشمگیری افزایش می‌دهد.

۴. استفاده از ستون‌های AI

برای برچسب‌گذاری و پاکسازی خودکار داده‌ها از ستون‌های AI بهره ببرید؛ به‌عنوان مثال، ترجمه، عادی‌سازی قیمت یا انجام محاسبات ساده.

رعایت این نکات باعث می‌شود فرآیند اسکریپ وبسایت با هوش مصنوعی سریع‌تر،

دقیق‌تر و امن‌تر انجام شود.

با توجه به اهمیت روزافزون هوش مصنوعی در کسب‌وکارهای امروز، بهره‌گیری از ابزارهایی مانند Thunderbit برای اسکریپ و بسایت با هوش مصنوعی، نه تنها سرعت جمع‌آوری و تحلیل داده‌ها را افزایش می‌دهد؛ بلکه امکان تصمیم‌گیری هوشمند و مبتنی بر داده را نیز فراهم می‌کند.

برای کسانی که دنبال رشد سریع و جهشی در کسب‌وکار خود هستند، توصیه می‌شود مقاله «[جهش در کسب‌وکار با هوش مصنوعی](#)» را در وبسایت مدیرسبز مطالعه کنند؛ این مقاله با نگاهی کاربردی و عملی نشان می‌دهد چگونه استفاده هوشمندانه از هوش مصنوعی می‌تواند مسیر توسعه کسب‌وکار را شتاب دهد و فرصت‌های جدید خلق کند.