

## چالش هشتم: ویکی پدیا (100 امتیاز)

از زمانی که ویکی‌پدیا بعنوان یک دایره‌المعارف آنلاین پا به عرصه وجود گذاشت، حدود ۲۰ سال می‌گذرد. امروزه کمتر کسی است که این وبسایت را نشناسد و به آن مراجعه نکرده باشد. ویکی‌پدیا را عمدتاً به صفحات یا مقالات آن، که هر کدام توضیح‌دهنده یک موضوع است می‌شناسند. آنچه کمتر در برابر چشم است، **ویرایش‌هایی** است که کاربران در صفحات مختلف ویکی‌پدیا اعمال می‌کنند. به هر صفحه ویکی‌پدیا (در زبان فارسی) که مراجعه کنید، در گوشه‌ی بالا سمت چپ گزینه‌ای به نام «**نمایش تاریخچه**» وجود دارد. با کلیک روی این گزینه، با لیستی از **مشارکت‌های** کاربرهای مختلف در ویرایش محتوای آن صفحه روبرو خواهد شد، که از جدیدترین به قدیمی‌ترین مرتب شده است. به عنوان مثال، در صفحه **هندسه**، با کلیک بر روی «نمایش تاریخچه»، این موارد در ابتدای لیست ظاهر خواهند شد:

The screenshot shows the 'History' page of the article 'هندسه' (Geometry) on the Persian Wikipedia. The page title is 'هندسه: تاریخچه ویرایش‌ها'. Below the title, there is a search bar and a filter dropdown set to 'مقاله'. The main content is a table of revisions:

تاریخ	ساعت	کاربر
۱۴ ژوئن ۲۰۲۰	۱۶:۲۳	Almohammad.f.z
۲۰ مه ۲۰۲۰	۰۶:۰۲	Ila1383moh.92
۱۰ مه ۲۰۲۰	۰۹:۱۴	Reza luke
۷ مه ۲۰۲۰	۰۵:۳۵	5.124.176.133
۲۱ آوریل ۲۰۲۰	۲۳:۵۸	JamRezaMousavi
۲۱ آوریل ۲۰۲۰	۲۳:۵۷	JamRezaMousavi
۲۱ آوریل ۲۰۲۰	۲۳:۵۶	JamRezaMousavi
۲۱ آوریل ۲۰۲۰	۲۳:۵۵	JamRezaMousavi
۱۰ اکتبر ۲۰۱۹	۱۷:۲۳	Taddah
۱۰ اکتبر ۲۰۱۹	۱۷:۲۳	Taddah
۱۰ اکتبر ۲۰۱۹	۱۷:۲۴	Taddah
۱۰ اکتبر ۲۰۱۹	۱۷:۲۵	Taddah
۱۰ اکتبر ۲۰۱۹	۱۷:۲۳	Taddah

همانطور که ملاحظه می‌شود، برای هر ویرایش، بعد از تاریخ و زمان ویرایش، نام کاربری ویرایش‌گر دیده می‌شود، و در مقابل آن، با کلیک روی «مشارکت‌ها»، می‌توان دید که کاربر چه مقالات دیگری در ویکی‌پدیا را ویرایش کرده است. هم‌چنین، بعد از حجم ویرایش (به بایت)، در داخل پراتنز، یک امتیاز به هر ویرایش کاربر اختصاص داده شده است.

در قدم اول، این سوال از شما می‌خواهد از سه مقاله ویکی‌پدیا (که با تنوع موضوع انتخاب کرده‌ایم: **هندسه**، **پزشکی**، **شیراز**) شروع کنید، و در هر یک، با کلیک بر روی «نمایش تاریخچه»، کاربرهایی از ویرایش‌گران را انتخاب کنید که امتیازی مساوی یا بالاتر از ۵۰۰+ داشته باشند. بعنوان مثال، در صفحه هندسه، یکی از کاربرانی که این ویژگی را برآورده می‌کند کاربر **Ehsan Aslani** است. سپس با کلیک روی **مشارکت‌ها** (در مقابل نام کاربری ویرایش‌گرانی با این ویژگی)، با لیستی از مقالات روبرو می‌شوید که این کاربر آنها را ویرایش کرده است. تصویر زیر، بخش ابتدایی این لیست را نشان می‌دهد:



صفحه اصلی  
رویدادهای کنونی  
مقاله تصادفی  
کتابخانه

همکاری  
تغییرات اخیر  
ویکی‌پدیا در ویکی‌پدیا  
راهنما  
تعمیرات با ویکی‌پدیا

ابزارها  
نظر  
مشارکت‌ها  
سیاست‌ها  
نمایش گروه‌های کاربر

## مشارکت‌ها

برای Ehsan aslani (بحث) | سیاه‌پسته‌شدن‌ها | بارگذاری‌ها | سیاه‌ها

### جستجوی مشارکت‌ها

(جدیدترین | قدیمی‌ترین) نمایش (۵۰ مورد جدیدتر | ۵۰ مورد قدیمی‌تر) [۲۰ | ۵۰ | ۱۰۰ | ۲۵۰ | ۵۰۰]

- ۱۴ دسامبر ۲۰۱۶، ساعت ۱۶:۰۲ (تفاوت | تاریخچه) .. (۱,۱۶۴+) .. ستاره (+) تاریخچه مشاهده: افزون متن (برچسب: ویرایش‌گر دیداری)
- ۱۲ دسامبر ۲۰۱۶، ساعت ۱۸:۳۰ (تفاوت | تاریخچه) .. (۳,۵۳۵+) .. ستاره (+) تاریخچه مشاهده: افزون بخش جدید (برچسب: ویرایش‌گر دیداری)
- ۱۲ دسامبر ۲۰۱۶، ساعت ۱۷:۳۵ (تفاوت | تاریخچه) .. (۰,۲۷) .. جین جهان (برچسب: ویرایش‌گر دیداری)
- ۱۲ دسامبر ۲۰۱۶، ساعت ۱۷:۳۳ (تفاوت | تاریخچه) .. (۰,۲۸) .. جهان (افزون پیشنهاد انجام مقالات جهان و گیش) (برچسب: ویرایش‌گر دیداری)
- ۱۲ دسامبر ۲۰۱۶، ساعت ۹:۴۷ (تفاوت | تاریخچه) .. (۳,۸۰۶+) .. فضا (افزون زیربخش‌های جدید) (برچسب: ویرایش‌گر دیداری)
- ۱۱ دسامبر ۲۰۱۶، ساعت ۲۰:۰۳ (تفاوت | تاریخچه) .. (۱۲,۱۴۴+) .. فضا (افزون بخش های جدید) (برچسب: ویرایش‌گر دیداری)
- ۱۱ دسامبر ۲۰۱۶، ساعت ۱۸:۱۰ (تفاوت | تاریخچه) .. (۱,۴۳۱+) .. فضا (افزون به متن مقدمه) (برچسب: ویرایش‌گر دیداری)
- ۱۱ دسامبر ۲۰۱۶، ساعت ۱۷:۳۳ (تفاوت | تاریخچه) .. (۱,۱۲۶+) .. جز فضا (برچسب: ویرایش‌گر دیداری به ویرایش منبع تغییر داده شده)
- ۱۱ دسامبر ۲۰۱۶، ساعت ۱۷:۳۳ (تفاوت | تاریخچه) .. (۲۵,۰۲۹-) .. گیش (۳-مدلهای کیهان‌شناسی، افزون بخش جدید) (برچسب: ویرایش‌گر دیداری)
- ۱۱ دسامبر ۲۰۱۶، ساعت ۱۱:۳۰ (تفاوت | تاریخچه) .. (۲,۱۹۸+) .. گیش (۳-مدلهای کیهان‌شناسی، افزون بخش جدید) (برچسب: ویرایش‌گر دیداری)

تا این قسمت از سوال، مهارت شما در **crawl** کردن داده سنجیده می‌شود. در مرحله بعدی، هدف، بکارگیری پردازش زبان طبیعی (NLP) بعنوان یکی از ارکان یادگیری ماشین و هوش مصنوعی است. در این مرحله از شما خواسته می‌شود برای هر یک از کاربرانی که امتیاز بزرگتر یا مساوی +۵۰۰ در ویرایش هر یک از ۳ مقاله مذکور (هندسه، پزشکی، شیراز) کسب کرده‌اند، بر مبنای مقالاتی که در ویکی‌پدیا ویرایش کرده‌اند، علاقمندی‌های آنها، با محاسبه میانگین وزنی امتیازها بصورت درصدی اعلام شود. از شما خواسته می‌شود که علاقمندی‌های کاربر را مطابق پنج گروه «ریاضی-فیزیک، علوم تجربی، علوم انسانی، هنر و زبان» (به ترتیبی که در دسته‌بندی کنکور سراسری کارشناسی مرسوم است) قرار دهید. برای این منظور می‌توانید از بخش «رده‌ها» که در پایان تمام مقالات ویکی‌پدیا وجود دارد بهره بگیرید. بعنوان مثال، دسته‌بندی زیر در پایان مقاله «نوسان پلاسما» دیده می‌شود:

## رده‌ها: امواج در پلاسما | پلاسمونیک | فیزیک پلاسما

### نکات:

- مقالاتی که احیاناً «رده‌بندی» بصورت بالا نداشته باشند شامل این چالش نمی‌شوند.
- یکی از مسایلی که در حل این چالش بروز پیدا می‌کند، محدودیت‌هایی است که ویکی‌پدیا (مثل هر سایت دیگر) بر تکرار دسترسی‌های شما به این وبسایت ایجاد خواهد کرد. از شما خواسته می‌شود برای برنخوردن به مشکل از پیشنهادات ویکی‌پدیا (مثل استفاده از API) یا هر راه‌حل ابتکاری دیگر بهره بگیرید. پیشنهاد می‌شود اگر در ویکی‌پدیا اکانت ندارید، یک حساب کاربری در ابتدای کار در ویکی‌پدیا ایجاد کنید.
- بعضی مقالات در ویکی‌پدیا تغییر مسیر داده‌اند، که لازم است در حل این چالش این مسأله لحاظ شود.
- توجه شود که بعضی نام‌های کاربری بصورت آدرس IP در «نمایش تاریخچه» دیده می‌شوند. برای مشاهده مشارکت‌های این کاربران، لازم است روی همین نام کاربری کلیک شود. بخش داخل پروانته بصورت (بحث‌ها | مشارکت‌ها) برای این کاربران وجود ندارد.
- هم‌چنین در نظر داشته باشید که یک مقاله ممکن است در بیش از یک دسته قرار بگیرد. برای مثال مقاله **اقتصاد** در دو دسته علوم انسانی و علوم ریاضی قرار خواهد گرفت. توجه شود امتیاز کاربران در مقالاتی که در چند گروه قرار خواهند گرفت برای تمامی گروه‌ها اعمال می‌شود.

به عنوان مثال، یک کاربر می‌تواند در ۳ صفحه مشارکت داشته باشد که دو صفحه در حوزه علوم تجربی قرار گرفته و برای هر کدام ۱۰۰۰ امتیاز کسب کرده و در صفحه‌ای مثل اقتصاد که در دو دسته‌ی علوم انسانی و ریاضی-فیزیک قرار می‌گیرد ۴۰۰۰ امتیاز دریافت کرده باشد، خروجی مورد انتظار برای این کاربر برابر ۴۰ درصد علوم انسانی و ۴۰ درصد ریاضی-فیزیک و ۲۰ درصد علوم تجربی است.

پاسخ ناقص نیز امتیاز در بردارد.

استفاده از هر زبان و تکنولوژی آزاد است.