

به نام خدا

گزارش کار فاز دوم پروژه ی بازیابی اطلاعات

امیرپارسا سلمان خواه

۹۸۳۱۰۳۴

بخش اول: بررسی عملکرد روش tf-idf

نتایج به دست آمده از موتور جستجو را در ۴ حالت زیر بررسی می کنیم. در هر مورد آیدی سند، عنوان سند، لینک سند، جمله های حاوی کلمه و مرتبط بودن سند به کلمه ی مورد جستجو شرح داده می شود. لازم به ذکر است که در هر جستجو ۱۰ سند بازگردانده می شود و برای جستجو از چمپیونز لیست ۲۰ تایی برای هر عبارت استفاده شده است و همچنین محدودیتی برای idf کلمات در نظر گرفته نشده اما میزان Doc coverage روی ۰.۵ تنظیم شده است. به این معنی که فقط سند هایی برگردانده می شوند که حداقل ۵۰ درصد کلمات حاضر در پرسمان را در بر داشته باشند.

الف) یک پرسمان از کلمات ساده و متداول تک کلمه ای (بین الملل):

۱۰۰۹: معاون بین الملل و مدیر کمیته حرفه ای سازی باشگاه استقلال منصوب شدند

<https://www.farsnews.ir/news/۱۴۰۰۰۷۱۹۰۰۰۶۹۵/>معاون-بین الملل-و-مدیر-

کمیته حرفه ای-سازی-باشگاه-استقلال-منصوب-شدند

جملات حاوی کلمه:

۱-مدیرعامل استقلال معاون بین الملل جدید و مدیر کمیته حرفه ای سازی این باشگاه را معرفی کرد.

۲- مصطفی آجورلو، مدیرعامل باشگاه استقلال در حکمی «حامد افضلی» را به عنوان «معاون بین الملل» باشگاه استقلال منصوب کرد.

۳- او سابقه فعالیت در بخش بین الملل باشگاه استقلال در مقاطع مختلف را در کارنامه خود دارد.

۴-نادری سابقه فعالیت در حوزه بین الملل و حرفه ای سازی فوتبال ایران را در کارنامه خود دارد. میزان مرتبط بودن سند:

علی رغم کوتاه بودن سند، کلمه ی بین الملل ۴ بار در این سند ظاهر شده و سند به کلمه ی مورد جستجو مرتبط است.

۵۱۳۴: یک دستاورد جدید جهانی برای اقتصاد ایران

<https://www.isna.ir/news/۱۴۰۰۰۵۱۵۱۰۴۹۱۱/>یک-دستاورد-جدید-جهانی-برای-

اقتصاد-ایران

جملات حاوی کلمه:

۱- به گزارش ایسنا، اتاق بازرگانی بین الملل که نهاد مادر همه اتاق های بازرگانی در سراسر دنیاست،...

۲- این موضوعات از طریق ارتباط اتاق تهران با اتاق بازرگانی بین الملل به مدیران اجرایی آن و سایر اتاق های عضو منعکس شد

۳- در نتیجه این فعالیت ها، اتاق بازرگانی بین الملل، اتاق بازرگانی، صنایع، معادن و کشاورزی تهران را به عنوان اتاق برتر یا فعال ترین اتاق بازرگانی ماه جولای جهان انتخاب کرده است.

میزان مرتبط بودن سند:

کلمه ی بین الملل ۳ بار در سند ظاهر شده و سند در مورد یک دستاورد جهانی صحبت می کند و لذا سند برگردانده شده به کلمه مربوط است.

۲۶۴۵: برگزاری مراسم بزرگداشت مرحوم عسگرخانی در دانشگاه تهران با حضور وزیر ارشاد

<https://www.farsnews.ir/news/۱۴۰۰۰۷۰۴۰۰۰۶۸۴/> برگزاری-مراسم-بزرگداشت-

مرحوم-عسگرخانی-در-دانشگاه-تهران-با-حضور-وزیر

جملات حاوی کلمه:

۱- مراسم بزرگداشت مرحوم ابومحمد عسگرخانی استاد روابط بین‌الملل دانشکده حقوق و علوم سیاسی دانشگاه تهران با حضور وزیر فرهنگ و ارشاد اسلامی برگزار شد.

۲- بنده افتخار شاگردی استاد عسگرخانی را در دانشکده بین‌الملل روابط خارجه داشتم.

۳- صحبت‌های مختلفی را درمورد فضای بین‌الملل مطرح می‌کردند.

۴- نگاه ایشان این بود که کشوری که مقتدر باشد، می‌تواند در فضای بین‌الملل تأثیرگذار باشد.

۵- فقدان ایشان یک صلبه مهم برای جامعه علمی و علوم سیاسی و روابط بین‌الملل بود.

میزان مرتبط بودن سند:

سند در مورد درگذشت مرحوم عسگرخانی است که استاد روابط بین‌الملل بوده است و کلمه ی بین‌الملل ۵ بار در آن تکرار شده است. پس نتیجه ی جستجو مرتبط است و همچنین از آن جا که طول این سند نسبت به دو سند قبلی بیشتر است، علی رغم تعداد بیشتر کلمه ی بین‌الملل در این سند، رتبه ی آن از دو سند دیگر پایین تر قرار گرفته است.

۴۷۶۹: استفاده از روبات‌های قاتل، تهدیدی برای امنیت بین‌المللی است

<https://www.isna.ir/news/۱۴۰۰۰۷۲۳۱۶۲۵۵/> استفاده-از-روبات-های-قاتل-تهدیدی-

برای-امنیت-بین-المللی-است

جملات حاوی کلمه:

۱- باید محدودیت‌های بین‌المللی وجود داشته باشد،

۲- زیرا این سلاح‌ها تهدیدی برای صلح و امنیت بین‌المللی و نقض اصل تناسب و استانداردهای حقوق بشر هستند.

۳- رشاعقیل دانشجوی دکترای حقوق بین‌الملل دانشگاه قم، در همایش بین‌المللی حقوق بین‌الملل و مخاصمات مسلحانه در مناطق غرب آسیا که در تالار شیخ مفید قم برگزار شد، عنوان کرد

۴- از استفاده از روبات‌های کشنده ، باید محدودیت‌های بین‌المللی وجود داشته باشد،

۵- زیرا این سلاح‌ها تهدیدی برای صلح و امنیت بین‌المللی و نقض اصل تناسب و استانداردهای حقوق بشر هستند.

۶- وی ادامه داد: کمیته بین‌المللی و چندین سازمان در تلاش هستند تا پیامدهای هوش مصنوعی را بررسی کنند

۷- دانشجوی دکترای حقوق بین‌الملل دانشگاه قم بیان کرد

۸- کسانی که با وضع قوانین بین‌المللی دغدغه دارند باید تغییرات رخ داده و ایجادشده و خلل. هارا در نظر بگیرند تا نتوان از آنها سوءاستفاده کرد.

۹- باید با ارزیابی چارچوب‌های قانونی بین‌المللی و داخلی موجود موردبررسی قرار گیرد.

۱۰- لازم است قوانینی برای تعیین رفتار مسئولانه استفاده از روبات‌های کشنده، مطابق با قوانین حقوق بشر دوستانه بین‌المللی و حقوق بین‌الملل حقوق بشر تعیین شود.

میزان مرتبط بودن سند:

در این سند ۱۷ بار کلمه ی بین الملل یا بین المللی تکرار شده است. این سند نسبت به سند های قبلی مرتبط تر به نظر می رسد اما با بررسی دقیق تر متوجه شدم که کلمه ی بین المللی توسط کتابخانه های داده شده به کلمه ی بین الملل تبدیل نمی شود و به همین دلیل این ۱۷ کلمه به طور کامل پیدا نشده اند و تنها ۳ بار حضور این کلمه شمرده شده است.

۲۸۰۴: پیام تسلیت بسیج دانشجویی در پی درگذشت استاد روابط بین الملل دانشگاه تهران

<https://www.farsnews.ir/news/۱۴۰۰۰۶۳۰۰۰۳۶۱/>پیام-تسلیت-بسیج-دانشجویی-
در-پی-درگذشت-استاد-روابط-بین الملل-دانشگاه
جملات حاوی کلمه:

- ۱- بسیج دانشجویی دانشکده حقوق و علوم سیاسی دانشگاه تهران طی پیامی درگذشت ابومحمد عسگرخانی استاد روابط بین الملل این دانشکده را تسلیت گفت .
 - ۲- درگذشت استاد گرانمایه گروه روابط بین الملل دانشکده حقوق و علوم سیاسی دانشگاه تهران، دانشمند دلسوز و فرهیخته جناب آقای دکتر ابومحمد عسگرخانی را خدمت خاندان مکرم ایشان، اساتید و دانشجویان و علاقه مندان آن مرحوم تسلیت و تعزیت عرض می نمایم.
- میزان مرتبط بودن سند:
کلمه ۲ بار در این سند حضور داشته و حضور آن در این رتبه منطقی به نظر می رسد.

۲۸۶۱: عدم حمایت آمریکا از شورای همکاری خلیج فارس و زلزله عربستان، امارات و بحرین / اعراب سیگنال مذاکره فرستاده اند

<https://www.farsnews.ir/news/۱۴۰۰۰۶۲۸۰۰۰۲۴۳/>عدم-حمایت-آمریکا-از-شورای-
همکاری-خلیج-فارس-و-زلزله-عربستان-امارات-و-
بحرین / اعراب سیگنال مذاکره فرستاده اند
جملات حاوی کلمه:

- ۱- حسن هانی زاده کارشناس مسائل بین الملل در گفت و گو با خبرنگار سیاسی خبرگزاری فارس، درباره گستاخی اخیر شورای همکاری خلیج فارس در بیانیه اخیرشان و مداخله در امور داخلی کشورمان اظهار داشت
 - ۲- کارشناس مسائل بین الملل تصریح کرد
- میزان مرتبط بودن سند:
در این جا نیز ۲ بار شاهد این کلمه بودیم و با توجه به طول بسیار بیشتر این سند نسبت به سند قبلی منطقیست که رتبه ی آن پایین تر باشد.

۴۴۴۰: صهیونیست ها به بن بست سیاسی رسیده اند

<https://www.isna.ir/news/۱۴۰۰۰۲۱۴۰۹۷۱۷/>صهیونیست-ها-به-بن-بست-سیاسی-
رسیده-اند
جملات حاوی کلمه:

- ۱- دستیار ویژه رئیس مجلس در امور بین الملل افزود
- ۲- دستیار ویژه رئیس مجلس در امور بین الملل با بیان این که جمهوری اسلامی ایران در تمام...
- ۳- دستیار ویژه رئیس مجلس در امور بین الملل در پایان گفت

میزان مرتبط بودن سند:
در این سند ۳ بار شاهد حضور کلمه ی بین‌الملل هستیم اما به دلیل طول بالاتر بردار آن نسبت به سند قبل، در رتبه ی پایین تری قرار گرفته است که منطقی است.

۶۲۸۰: جشنواره قصه‌گویی فراخوان داد

<https://www.isna.ir/news/۱۴۰۰۱۲۱۱۰۸۴۰/> جشنواره-قصه-گویی-فراخوان-داد

جملات حاوی کلمه:

۱- به گزارش ایسنا به نقل از اداره کل روابط عمومی و امور بین‌الملل کانون، بر اساس فراخوان منتشرشده، بیست‌وسومین جشنواره بین‌المللی قصه‌گویی کانون پرورش فکری کودکان و نوجوانان

به دلیل شیوع کرونا، به صورت مجازی برگزار خواهد شد

۲- خوش بین‌الملل شامل رقابت صحنه‌ای (مجازی) قصه‌گویان برگزیده‌ی ۲۲ دوره جشنواره قصه‌گویی در مراحل منطقه‌ای، ملی یا بین‌الملل، قصه‌گویان و مجریان حرفه‌ای دارای سابقه اجرا در رسانه ملی و تئاتر کشور و قصه‌گویان منتخب سایر کشورها، دو بخشی است که همانند گذشته و با تغییراتی جزئی برگزار خواهد شد.

۳- برگزیده شدن در حداقل ۳ دوره از جشنواره‌های قصه‌گویی کانون در سطح ملی یا بین‌الملل یا جشنواره‌های معتبر قصه‌گویی بین‌الملل، سابقه حداقل ۵ اجرای قصه‌گویی به صورت حرفه‌ای در رسانه ملی، فعالیت موثر اجتماعی در رخدادهای قصه‌گویی، همکاری با خیریه‌ها، ویژه‌برنامه‌ها و...
میزان مرتبط بودن سند:

در این سند ۵ بار کلمه ی بین‌الملل و چند بار نیز کلمه ی بین‌المللی آمده که همانند مورد ذکر شده در بخش های قبلی، کلمه ی بین‌المللی به بین‌الملل تبدیل نشده و شمارش نشده است. همچنین به دلیل بزرگ بودن بردار این سند، رتبه ی آن از سند قبلی کمتر است.

۴۷۶۷: تحریم‌های آمریکا علیه طالبان آثار مخرب بر مردم افغانستان دارد نه حاکمان افغانستان

<https://www.isna.ir/news/۱۴۰۰۰۷۲۲۱۵۸۰۰/> تحریم-های-آمریکا-علیه-طالبان-آثار-

مخرب-بر-مردم-افغانستان-دارد

جملات حاوی کلمه:

۱- دانشیار حقوق بین‌الملل دانشگاه علامه طباطبائی تهران گفت

۲- پوریا عسکری، دانشیار حقوق بین‌الملل دانشگاه علامه طباطبائی تهران، در همایش حقوق بین‌الملل و مخاصمات مسلحانه در منطقه غرب آسیا که در تالار شیخ مفید دانشگاه قم برگزار شد، عنوان کرد

۳- دانشیار حقوق بین‌الملل دانشگاه علامه طباطبائی تهران گفت

میزان مرتبط بودن سند:

در این سند نیز دفعات زیادی کلمه ی بین‌المللی ظاهر شده که شمارش نشده است. همچنین ۴ بار کلمه ی بین‌الملل آمده و با توجه به طول بردار آن، رتبه ی آن منطقی به نظر می‌رسد.

۱۵۴: توضیحات مسؤول مسابقات لیگ یک درباره شایعه سخته ناظر بازی

<https://www.farsnews.ir/news/۱۴۰۰۰۸۰۱۰۰۰۷۲۲>/توضیحات-مسؤول-مسابقات-

لیگ-یک-درباره-شایعه-سخته-ناظر-بازی

جملات حاوی کلمه:

- ۱- تلاش‌های خانم اسدی در بخش روابط بین‌الملل فدراسیون فوتبال و همچنین پیگیری‌ها و تلاش‌ها روزه و وثوق احمدی مشکل پنجره نقل و انتقالاتی باشگاه پارس جنوبی جم حل شد.
- ۲- وثوق احمدی پیگیر پرونده خارجی این باشگاه بود و با نفوذ و تجربه‌ای که در حوزه بین‌الملل داشت به عنوان یک هموطن کمک بسیار شایانی در مورد این مساله کرد که جا دارد از این دو بزرگوار تشکر ویژه کنم.

میزان مرتبط بودن سند:

در این سند ۲ بار کلمه ی بین الملل ظاهر شده است و با توجه به اینکه کلمه ی بین‌المللی محاسبه نمی شود، سندی بهتر از این سند برای این رتبه پیدا نشده است.

بنابراین در کل می توان الگوریتم جستجو عملکرد خوبی را در این بخش از خود ارائه داده است و مشکلات محاسبه نشدن کلمه ی بین‌المللی مربوط به بخش پیش پردازش داده ها است.

ب) یک پرسمان از کلمات ساده و متداول چند کلمه ای (سازمان ملل متحد):

۲۳۰۲: هیئت پارلمانی ایران به رم سفر کرد

<https://www.farsnews.ir/news/۱۴۰۰۰۷۱۷۰۰۰۶۰۷/>هیئت-پارلمانی-ایران-به-رم-

سفر-کرد

جملات حاوی کلمه:

- ۱- نشست مقدماتی کاپ ۲۶ با حضور رؤسا و نمایندگان پارلمان‌های ۷۲ کشور و ۸ سازمان بین‌المللی به میزبانی مجلس نمایندگان ایتالیا و با مشارکت اتحادیه بین‌المجالس جهانی IPU و سازمان ملل متحد در روزهای ۱۶ و ۱۷ مهر ماه برگزار خواهد شد.
- ۲- قرار است پیش‌نویس اولیه سند نهایی کنفرانس پارلمانی سازمان ملل متحد در خصوص تغییرات اقلیم در این کنفرانس به تصویب برسد.

میزان مرتبط بودن سند:

در این سند دو بار عبارت سازمان ملل متحد به شکل دقیق تکرار شده است. همچنین کلمه ی سازمان نیز یک بار به تنهایی آمده است و در کل می توان گفت سند به عبارت جستجو مرتبط است.

۶۸۰۰: شعرهای میوه‌ای

<https://www.isna.ir/news/۱۴۰۰۰۴۰۹۰۶۹۶۱/>شعرهای-میوه-ای

جملات حاوی کلمه:

- ۱- سازمان ملل متحد با هدف ترویج عدالت غذایی و اصلاح سبک تغذیه، امسال را سال میوه نامیده است.
- ۲- مجمع عمومی سازمان ملل متحد سال جاری (۲۰۲۱ میلادی) را سال جهانی میوه و سبزیجات نام‌گذاری کرده است .

میزان مرتبط بودن سند:
در این سند نیز ۲ بار عبارت سازمان ملل متحد ظاهر شده است و به دلیل طول بیشتر بردار سند، در رتبه ای پایین تر از سند قبل قرار گرفته است.

۲۷۴۴: راستینه: اظهارات آیت الله رئیسی مطالبه گری انقلابی در سازمان ملل بود / تاکید رئیس جمهور بر منافع ملت

<https://www.farsnews.ir/news/۱۴۰۰۰۶۳۱۰۰۰۷۸۸/> راستینه-اظهارات-آیت-الله-

رئیس-مطالبه-گری-انقلابی-در-سازمان-ملل-بود-

جملات حاوی کلمه:

۱- با اشاره به سخنرانی ویدیو کنفرانسی آیت الله رئیسی در هفتاد و ششمین نشست سالانه مجمع عمومی سازمان

ملل متحد اظهار داشت

۲- وجود رؤسای جمهور در اجلاس سازمان ملل متحد فرصت بسیار ارزشمندی است که حقوق ملت مطالبه شود

همچنین در ۴ جمله ی دیگر عبارت سازمان ملل تکرار شده است.

میزان مرتبط بودن سند:

این سند نسبت به سند های قبلی مرتبط تر به نظر می رسد و تعداد کلمات مرتبط بیشتری دارد. این موضوع را از طریق کد نیز بررسی کردم و برقرار بود. اما به دلیل طول بیشتر بردار سند نسبت به سند های قبلی، رتبه ی این سند پایین تر است.

۶۹۱۲: انتشار «کودکان و حقوق بشر» برای بچه ها

<https://www.isna.ir/news/۱۴۰۰۰۴۲۴۱۷۳۴۶/> انتشار-کودکان-و-حقوق-بشر-برای-بچه-

ها

جملات حاوی کلمه:

۱- کنوانسیون حقوق کودک سازمان ملل متحد با تصویب پیمان نامه ای عناوین مهمی از حقوق کودک را تصویب کرده که در پی آن کشورهای زیادی با امضای این پیمان نامه حقوق کودک را محترم شمردند .

۲- این مجموعه دوجلدی با تکیه بر پیمان نامه حقوق کودک سازمان ملل متحد بر آن است کودکان را بر حقوق فردی و اجتماعی خود آشنا کند.

میزان مرتبط بودن سند:

این سند نیز به سازمان ملل متحد مربوط است و این عبارت ۲ بار در آن تکرار شده است و به دلیل طول بیشتر بردار آن نسبت به سند دوم، در جایگاه چهارم قرار گرفته است.

۲۷۴۲: آصفری: اظهارات رئیسی در سازمان ملل اقدامی برای رسوا کردن غربی ها و احقاق حقوق مردم بود

<https://www.farsnews.ir/news/۱۴۰۰۰۶۳۱۰۰۰۸۱۲/> آصفری-اظهارات-رئیس-در-

سازمان-ملل-اقدامی-برای-رسوا-کردن-غربی-ها-و

جملات حاوی کلمه:

۱- با اشاره به سخنرانی ویدیوکنفرانسی آیت‌الله رئیسی در هفتاد و ششمین نشست سالانه مجمع عمومی سازمان ملل متحد اظهار داشت

میزان مرتبط بودن سند:

در این سند فقط یک بار عبارت سازمان ملل متحد آمده است و با توجه به طول بردار آن که تقریباً برابر با سند قبل‌بست، منطقی است که در جایگاه پنجم باشد.

۲۶۹۰: نوروزی: در ۸ سال قبل دلمان برای سخنرانی‌های انقلابی رئیس‌جمهور در سازمان ملل تنگ شده بود

<https://www.farsnews.ir/news/۱۴۰۰۰۷۰۳۰۰۰۳۸۰/> نوروزی-در-سال-قبل-دلمان-

برای-سخنرانی‌های-انقلابی-رئیس-جمهور-در

جملات حاوی کلمه:

در این سند عبارت سازمان ملل متحد به شکل کامل نیامده است اما ۵ بار عبارت سازمان ملل ظاهر شده است.

۱- با تمجید از مواضع و سخنرانی رئیس‌جمهور کشورمان در مجمع عمومی سازمان ملل گفت
۲- دلمان برای سخنرانی انقلابی رئیس‌جمهور کشورمان در اجلاس سالیانه مجمع عمومی سازمان ملل تنگ شده بود

۳- مواضع و صحبت‌های آقای رئیسی در هفتصد و ششمین مجمع عمومی سازمان ملل همان چیزی بود که مورد انتظارمان بود و تحقق یافت.

۴- آقای رئیسی در مجمع عمومی سازمان ملل بحث حمایت از ملل مظلوم منطقه و همچنین بحث مواضع ایران در برجام و مسائل منطقه‌ای را بیان کرد

۵- سخنرانی آقای رئیسی در مجمع عمومی سازمان ملل استحکام و اقتدار نظام جمهوری

اسلامی ایران را به همگان نشان داد

میزان مرتبط بودن سند:

سند به سازمان ملل متحد مربوط است اما با توجه به عدم حضور کلمه ی متحد در سند، زاویه ی برداری آن فاصله ی بیشتری با بردار سازمان ملل متحد دارد و در این رتبه قرار گرفته است.

۲۲۹۹: اعتراض هیات پارلمانی ایران به نماینده منصور هادی در اجلاس تغییرات آب و هوایی / همکاری دولت مستعفی یمن با ائتلاف سعودی علیه یمنی‌ها

<https://www.farsnews.ir/news/۱۴۰۰۰۷۱۷۰۰۰۶۳۹/> اعتراض-هیات-پارلمانی-ایران-

به-نماینده-منصور-هادی-در-اجلاس-تغییرات-آب-و

جملات حاوی کلمه:

۱- ایران آمادگی کامل خود را برای رسیدگی و حل این مشکل در همکاری با سازمان ملل اعلام کرده.

۲- با مشارکت اتحادیه بین‌المجالس جهانی IPU و سازمان ملل متحد در روزهای ۱۶ و ۱۷ مهرماه برگزار خواهد شد.

در نشست پایانی کنفرانس قرار است پیش‌نویس اولیه سند نهایی کنفرانس پارلمانی سازمان ملل متحد در خصوص تغییرات اقلیم به تصویب برسد.

میزان مرتبط بودن سند:
در این سند ۲ بار عبارت سازمان ملل متحد و یک بار سازمان ملل ظاهر شده است اما طول بردار آن خیلی بیش تر از سند قبلی است و این باعث شده تا در رتبه ی پایین تری قرار بگیرد.

۲۸۵۵: آزمون دشوار شورای حقوق بشر / آیا تحریم های آمریکا لغو می شوند؟

<https://www.farsnews.ir/news/۱۴۰۰۰۶۲۸۰۰۰۳۳۵/> آزمون-دشوار-شورای-حقوق-

بشر-آیا-تحریم های-آمریکا-لغو-می شوند

جملات حاوی کلمه:

در این سند ۵ بار عبارت سازمان ملل، ۳ بار کلمه ی سازمان و ۱ بار کلمه ی متحد آمده است.
میزان مرتبط بودن سند:
این سند نسبت به سند قبلی کمی مرتبط تر است اما باز هم به دلیل طول بیشتر بردار، در رتبه ی پایین تر قرار گرفته است.

۶۸۷۴: پیشنهادی برای کم کردن تنش بین کشورهای حوزه نوروژ

<https://www.isna.ir/news/۱۴۰۰۰۴۲۰۱۴۵۰۵/> پیشنهادی-برای-کم-کردن-تنش-بین-

کشورهای-حوزه-نوروژ

جملات حاوی کلمه:

- ۱- از همین رو، انجمن ایرانی شورای مشورتی سازمان ملل متحد پیشنهاد کرد
 - ۲- انجمن فرهنگی افراز که دارای مقام مشورتی در حوزه فرهنگ از شورای اقتصادی و اجتماعی سازمان ملل متحد است
 - ۳- ر پاسخ به فراخوان سازمان ملل متحد درباره پیشنهادها یا اقدام های انجام شده در حوزه توسعه پایدار در بیانیه ای موضوع لزوم همکاری کارگروه های ثبت میراث معنوی جهانی این کشورها با یکدیگر را پیشنهاد داد
 - ۴- در این بیانیه که به فارسی نیز ترجمه شده و در کتابچه مجموعه بیانیه های انجمن های مردم نهاد دارای مقام مشورتی شورای اقتصادی و اجتماعی سازمان ملل متحد از سراسر جهان منتشر شده
 - ۵- بیش از یک دهه پیش در سال ۲۰۱۰، روز جهانی نوروژ برابر با ۲۱ مارس (اعتدال بهاری) با عنوان «فرهنگ صلح» توسط مجمع عمومی سازمان ملل متحد به تصویب رسید
- میزان مرتبط بودن سند:
این سند نسبت به دو سند خیلی مرتبط تر است زیرا ۴ بار عبارت سازمان ملل متحد در آن آمده است و همچنین یک بار هم عبارت سازمان ملل. اما طول بردار آن خیلی زیاد است و این موجب شده رتبه ی آن پایین تر باشد.

۳۲۵۵: تأکید مقامات IPU و سازمان ملل در توسعه همکاری‌ها برای مقابله با کرونا
<https://www.farsnews.ir/news/۱۴۰۰۶۱۷۰۰۰۹۴۸>/تأکید-مقامات-IPU-و-سازمان-ملل-در-توسعه-همکاری‌ها-برای-مقابله-با-کرونا
جملات حاوی کلمه:

- ۱- مراسم اختتامیه پنجمین اجلاس بین‌المجالس جهانی با سخنرانی صدراعظم اتریش و رئیس مجلس این کشور، رئیس مجمع عمومی سازمان ملل، دبیرکل سازمان ملل، رئیس بین‌المجالس جهانی برگزار شد.
 - ۲- همچنین در این مراسم دبیرکل و رئیس مجمع عمومی سازمان ملل در سخنرانی‌های خود بر حمایت سازمان ملل از بین‌المجالس جهانی تأکید و آمادگی خود را برای این منظور اعلام کردند. میزان مرتبط بودن سند:
- در این سند ۴ بار عبارت سازمان ملل آمده است و باز هم به دلیل طول بردار منطقی است که در این رتبه قرار بگیرد.

در نهایت می‌توان گفت که با توجه به اعمال شدن چمپیونز لیست و بخش index elimination که باعث می‌شود تا سند های خروجی حداقل تعدادی از کلمات پرسمان را در خورد داشته باشند، سند های خروجی سند های مرتبطی به کوئری هستند و اگر این پردازش ها وجود نداشتند ممکن بود سند های غیر مرتبط تری را در خروجی ببینیم.

(پ) یک پرسمان از کلمات دشوار و کم تکرار تک کلمه ای (ژیمناستیک):
۱۳۶۸: هشدار هیات ژیمناستیک تهران در خصوص سالن های مختلط و اقدامات غیراخلاقی

<https://www.farsnews.ir/news/۱۴۰۰۰۷۱۴۰۰۰۲۳۱/>هشدار-هیات-ژیمناستیک-تهران-در-خصوص-سالن-های-مختلط-و-اقدامات-غیراخلاقی
جملات حاوی کلمه:

- ۱- رضا شاکرمی در خصوص انتشار خبری مبنی بر فعالیت سالن مختلط ژیمناستیک در یکی از مناطق جنوب شهر تهران اظهار داشت
- ۲- این سالن ژیمناستیک تحت اجاره فردی خاص است
- ۳- این فرد که متأسفانه هم از پیشکسوتان ژیمناستیک و هم متصدی بخشی از ژیمناستیک تهران است، عنوان کرده از تمرین مختلط دختران و پسران در سالن ژیمناستیک مطلع نبوده است،
- ۴- سرپرست هیات ژیمناستیک استان تهران با هشدار به خانواده‌ها گفت
- ۵- متأسفانه چند وقت قبل هم یک دختر ۱۱ ساله که در رشته ژیمناستیک فعالیت می‌کند اسیر افکار شیطانی یکی از همین افراد شده بود.
- ۶- راه اندازی سالن های ژیمناستیک به صورت مختلط، اقدامی غیراخلاقی است که در پشت پرده آن مقاصد ناسالمی است و به همین خاطر پیگیر برخورد با مسببان آن هستیم.
- ۷- خانواده‌ها می‌توانند برای اطلاع از سالن های ژیمناستیک مجاز که مورد تایید فدراسیون ژیمناستیک هستند با این فدراسیون تماس بگیرند

۹- یا با تماس با هیئات ژیمناستیک استان محل سکونتشان، لیست سالن‌های ژیمناستیک دارای مجوز را دریافت کنند.
میزان مرتبط بودن سند:
در این سند ۱۲ بار کلمه ی ژیمناستیک آمده است و بسیار به عبارت جستجو مرتبط است.

۳۶۱۶: دبیر مجمع فدراسیون ژیمناستیک مشخص شد

<https://www.isna.ir/news/۱۴۰۰۲۰۹۰۶۴۰۸/> دبیر-مجمع-فدراسیون-ژیمناستیک-
مشخص-شد

جملات حاوی کلمه:

- ۱- علیرضا باقری از سوی وزارت ورزش به عنوان دبیر مجمع ژیمناستیک انتخاب شد.
- ۲- به گزارش ایسنا، ثبت نام از کاندیداهای پست ریاست فدراسیون ژیمناستیک از ۷ اردیبهشت ماه آغاز شده و به مدت ده روز کاری تا ۲۱ اردیبهشت ماه ادامه خواهد داشت.

میزان مرتبط بودن سند:

این سند با اینکه به ژیمناستیک مربوط است اما فقط ۲ بار شاهد تکرار این کلمه در آن هستیم. اما به دلیل اینکه طول بردار آن خیلی کم است، در رتبه ی دوم قرار گرفته است.

۶۳۳: خیرخواه: برخی به دنبال فلج کردن ژیمناستیک هستند / با بایکوت فدراسیون موفقیت‌ها بیشتر شد

<https://www.farsnews.ir/news/۱۴۰۰۷۲۵۰۰۱۵۲/> خیرخواه-برخی-به-دنبال-فلج-کردن-ژیمناستیک-هستند-با-بایکوت-فدراسیون
جملات حاوی کلمه:

- ۱- رضا خیرخواه در گفت‌وگو با خبرنگار ورزشی خبرگزاری فارس، درباره عدم اعزام تیم ملی ژیمناستیک هنری به مسابقات جهانی ژاپن اظهار داشت
- ۲- فدراسیون از مدت‌ها قبل درخواست اعزام تیم ملی ژیمناستیک را به شورای برون مرزی وزارت ورزش داده بود
- ۳- در دوره قبلی این رقابت‌ها مهدی احمدکهنی بهترین نتیجه تاریخ ژیمناستیک کشور را کسب کرده بود
- ۴- متأسفانه برخی به دنبال این هستند که به هر نحو ممکن ژیمناستیک ایران را فلج کنند تا از این طریق حرف‌شان را به کرسی بنشانند.
- ۵- سرمربی تیم ملی ژیمناستیک هنری گفت
- ۶- از آقای سجادی خواهش می‌کنم صدای اهالی ژیمناستیک کشور را بشنود
- ۷- اینگونه اقدامات علیه ژیمناستیک کشور انجام نمی‌شد اما خدا جای حق نشسته است.
- ۸- دیدیم که هر چه بیشتر ژیمناستیک را بایکوت کردند، موفقیت‌ها بیشتر شد.
- ۹- چنین اتفاق تاریخی نشان داد که ژیمناستیک ایران به کلی متحول شده است
- ۱۰- ژیمناستیک ثابت کرد به لحاظ مدال‌آوری می‌تواند جزو ۱۰ رشته ورزشی کشورمان باشد.
- ۱۱- سرمربی تیم ملی ژیمناستیک هنری خاطرنشان کرد
- ۱۲- از حذف ژیمناستیک از المپiad استعدادیابی و جلوگیری از اعزام ملی‌پوشان به مسابقات جهانی تا قطع بودجه فدراسیون.

میزان مرتبط بودن سند:
در این سند ۱۲ بار کلمه ی ژیمناستیک تکرار شده است و یکی از مرتبط ترین سند ها به ژیمناستیک به شمار می رود. با این وجود به دلیل طول بالای بردار آن، نسبت به سند قبلی کم ارزش تر شناخته شده است.

۳۶۶۵: ثبت نام ۱۳ نامزد برای پست ریاست فدراسیون ژیمناستیک + اسامی
<https://www.isna.ir/news/۱۴۰۰۰۲۲۱۱۴۸۶۲/> ثبت نام-۱۳-نامزد-برای-پست-ریاست-فدراسیون-ژیمناستیک-اسامی

جملات حاوی کلمه:
۱- ۱۳ نامزد برای پست ریاست فدراسیون ژیمناستیک نام نویسی کردند.
۲- به گزارش ایسنا، ثبت نام از کاندیداهای پست ریاست فدراسیون ژیمناستیک از ۷ اردیبهشت ماه آغاز شد .
۳- ۱۳ نفر برای پست ریاست فدراسیون ژیمناستیک ثبت نام کرده اند .
۴- تاکنون تاریخ برگزاری انتخابات فدراسیون ژیمناستیک اعلام نشده و پیش بینی می شود بلافاصله پس از تایید صلاحیت نامزدها انتخابات این فدراسیون برگزار شود.

میزان مرتبط بودن سند:
در این سند ۴ بار کلمه ی ژیمناستیک آمده است و منطقی است که پایین تر از سند های بالاتر قرار بگیرد.

۴۱۸۹: جزییات تعطیلی های ورزش ایران تا ۹ مهر ۱۴۰۰ / کدام فعالیت های ورزشی در تهران ممنوع است؟
<https://www.isna.ir/news/۱۴۰۰۰۷۰۲۰۰۸۰۸/> جزییات-تعطیلی-های-ورزش-ایران-تا-۹-مهر-۱۴۰۰-کدام-فعالیت-های-ورزشی

جملات حاوی کلمه:
میزان مرتبط بودن سند:

۳۸۷۹: جزییات تعطیلی ورزش ایران تا پایان تیرماه + تصویر
<https://www.isna.ir/news/۱۴۰۰۰۴۲۷۱۹۸۳۳/> تیرماه-تصویر

جملات حاوی کلمه:
۳ بار کلمه ی ژیمناستیک (سالنی) در این متن تکرار شده است.
میزان مرتبط بودن سند:

این متن به طور کلی در مورد تعطیلی ورزش ها صحبت می کند و ۳ بار در آن کلمه ی ژیمناستیک به چشم می خورد و با توجه به طول زیاد بردار آن طبیعی است که در این جایگاه قرار بگیرد.

۴۰۵۷: دبیر: اگر من در مباحث فنی ۱۰ باشم، درستکار ۱۰۰ است / بنا کاملاً بر اساس چرخه انتخابی عمل کرد!

<https://www.isna.ir/news/۱۴۰۰۰۶۰۵۰۳۴۰۷/>دیبر-اگر-من-در-مباحث-فنی-۱۰-باشم-

درستکار-۱۰۰-است-بنا-کاملا

جملات حاوی کلمه:

۱- یک نفر از دوستانی که در ژیمناستیک بوده می آید و درباره کشتی نظر می دهد.
میزان مرتبط بودن سند:

در این سند که مربوط به کشتی است، فقط یک بار کلمه ی ژیمناستیک آمده است و به نظر بی ربط ترین سند به ژیمناستیک از بین سند های بالا همین سند است. بنابراین جایگاه آن منطقی است.

لازم به ذکر است حاصل خروجی فقط ۷ سند مرتبط بود و ۳ سند دیگر به ترتیب سند های ۱ و ۲ و ۳ با امتیاز صفر بودند که نشان گر این است که دیگر کلمه ی ژیمناستیک در سایر سند ها وجود ندارد.

در نتیجه می توان گفت که عملکرد موتور جستجو برای پیدا کردن این کلمه ی خاص بسیار خوب بوده است و اسناد هم تا حد خوبی به ترتیب اولویت چیده شده بودند.

ت) یک پرسمان از کلمات دشوار و کم تکرار چند کلمه ای (واکسن آسترازنکا):

۱۶۷: ستاره مسلمان بهترین فوتبالیست تاریخ را چه کسی می داند؟

<https://www.farsnews.ir/news/۱۴۰۰۰۸۰۱۰۰۰۴۹۰/>ستاره-مسلمان-بهترین-

فوتبالیست-تاریخ-را-چه-کسی-می-داند

جملات حاوی کلمه:

۱- محمد صلاح در مصاحبه با اسکای اسپورت در خصوص آینده خود در لیورپول صحبت کرد.

۲- صلاح اظهار داشت: در ذهن من، من بهترین بازیکن تاریخ فوتبال هستم .

میزان مرتبط بودن سند:

نام محمد صلاح یک بار در این سند ظاهر شده و علاوه بر آن یک بار هم کلمه ی صلاح آمده است و سند با موضوع مرتبط است.

۱۲۴: صلاح یا رونالدو؟ تیم کدام ستاره امشب برنده می شود+عکس

<https://www.farsnews.ir/news/۱۴۰۰۰۸۰۲۰۰۰۲۷۵/>صلاح-یا-رونالدو-تیم-کدام-

ستاره-امشب-برنده-می-شود-عکس

جملات حاوی کلمه:

۱- به همین بهانه مقایسه ای میان عملکرد محمد صلاح ستاره مصری لیورپول و کریس رونالدو فوق ستاره پرتغالی توسط سایت هواسکورد انجام شده است.

۲- محمد صلاح تاکنون ۷ گل و ۴ پاس گل برای لیورپول ثبت کرده است و کریس رونالدو در لیگ برتر ۳ گل زده و پاس گلی را ثبت نکرده است.

۳- صلاح در مجموع نمره ۸.۲۱ را کسب کرده و کریس رونالدو هم نمره ۶.۹۴ را به خود اختصاص داده است.

میزان مرتبط بودن سند:

در این سند ۲ بار عبارت محمد صلاح ظاهر شده است و یک بار نیز کلمه ی صلاح. اما با توجه به طول بیشتر بردار آن نسبت به سند قبلی، در رتبه ی دوم قرار گرفته است.

۶۹۰: لیگ برتر انگلیس | آتش بازی لیورپول در دیدار مقابل واتفورد/ هت تریک فیرمینو

<https://www.farsnews.ir/news/۱۴۰۰۰۷۲۴۰۰۰۶۹۹/> لیگ-برتر-انگلیس | آتش-بازی-

لیورپول-در-دیدار-مقابل-واتفورد-هت-تریک

جملات حاوی کلمه:

۱- در این بازی مانه ۹، فیرمینو ۳۷، ۵۲ و ۹۰+۱ و محمد صلاح ۵۴ گلزنی کردند.

میزان مرتبط بودن سند:

در این سند یک بار عبارت محمد صلاح ظاهر شده است و به موضوع مرتبط است و به دلیل طول کم بردار آن در رتبه ی سوم قرار گرفته است.

۴۰۳: طارمی بهترین بازیکن دیدار با میلان /ستاره مسلمان رکورددار گلزنی در

لیورپول+عکس

<https://www.farsnews.ir/news/۱۴۰۰۰۷۲۸۰۰۰۰۳۹/> طارمی-بهترین-بازیکن-دیدار-

با-میلان-ستاره-مسلمان-رکورددار-گلزنی-در

جملات حاوی کلمه:

۱- در این بازی محمد صلاح ستاره مسلمان و مصری لیورپول موفق به ثبت دبل شد و رکورد جدیدی در جمع قرمزپوشان ثبت کرد.

۲- محمد صلاح همچنین تبدیل به اولین بازیکن تاریخ لیورپول شد که در ۹ دیدار متوالی حداقل یک گل به ثمر رسانده است.

میزان مرتبط بودن سند:

در این سند ۲ بار نام محمد صلاح تکرار شده است اما با توجه طول بیشتر بردار آن نسبت به سند قبلی در رتبه ی پایین تر قرار گرفته است.

۶۲۶: ستاره مسلمان از مسی هم بهتر است+فیلم

<https://www.farsnews.ir/news/۱۴۰۰۰۷۲۵۰۰۰۲۹۰/> ستاره-مسلمان-از-مسی-هم-

بهتر-است-فیلم

جملات حاوی کلمه:

۱- محمد صلاح شب گذشته بار دیگر به سوژه بزرگ رسانه‌های دنیا بدل شد

۲- نمی‌توان از درخشش صلاح در این بازی گذشت.

۳- صلاح چنین بازیکنی است که می‌تواند چنین لحظاتی را مکرر تکرار کند.

۴- کلوپ در پاسخ به سوال خبرنگار که صلاح باید چه کار کند که به بهترین بازیکن دنیا بدل شود، گفت

۵- صلاح اکنون بهترین بازیکن دنیاست ولی فرقی با مسی و رونالدو این است که آنها چندین

سال در این کیفیت بودند

۶- محمد ۵ سال است چنین نمایشی از خودش نشان می‌دهد.

میزان مرتبط بودن سند:
در این سند یک بار عبارت محمد صلاح، چهار بار کلمه ی صلاح و یک بار نیز محمد به تنهایی تکرار شده است. این سند با وجود این که یکی از مرتبط ترین سند ها به نظر می رسد اما به دلیل طول بالای بردار آن رتبه ی پایین تری نسبت به بقیه ی اسناد ذکر شده دارد.

۴۱: صلاح گلزن ترین فوتبالیست آفریقایی تاریخ لیگ برتر انگلیس+عکس
<https://www.farsnews.ir/news/۱۴۰۰۰۸۰۳۰۰۰۷۴۲/>صلاح-گلزن-ترین-فوتبالیست-آفریقایی-تاریخ-لیگ-برتر-انگلیس-عکس

جملات حاوی کلمه:
۱- در این بازی محمد صلاح ستاره مسلمان و مصری باشگاه لیورپول موفق به ثبت هت تریک شد.
۲- صلاح با این گل ها از دیدیه دروگبا ستاره ساحل عاجی و سابق باشگاه چلسی که ۱۰۴ گل در لیگ برتر به ثمر رسانده است، عبور کرد.
۳- هم تیمی صلاح می تواند خود را به این ستاره مصری برساند.
میزان مرتبط بودن سند:
در این سند یک بار محمد صلاح و دو بار صلاح به تنهایی ظاهر شده است و رتبه ی آن منطقی به نظر می رسد.

۱۴۶۴: سایت آمریکایی ستاره مسلمان را بهترین بازیکن دنیا نامید / قول محمد صلاح به لیورپولی ها+فیلم
<https://www.farsnews.ir/news/۱۴۰۰۰۷۱۲۰۰۰۵۳۲/>مسلمان-را-بهترین-بازیکن-دنیا-نامید-قول-محمد-صلاح

جملات حاوی کلمه:
۱- به گزارش خبرگزاری فارس، محمد صلاح شب گذشته در چارچوب رقابت های لیگ برتر انگلیس باز هم شاهکار کرد و توانست نام خود را به عنوان «بهترین بازیکن کنونی جهان» بر سر زبان ها بیاندازد.
۲- این بازی با درخشش محمد صلاح به تساوی ۲ بر ۲ رسید و ستاره مسلمان توانست یکی از زیباترین گل های فصل را به نام خود ثبت کند.
۳- محمد صلاح فوق العاده است.
۴- جیمی اوهارا، کارشناس فوتبال انگلیس محمد صلاح را بهترین بازیکن دنیا دانست و او را نزدیک به لیونل مسی نامید.
۵- سسک فابرگاس، بازیکن سابق چلسی، بارسلونا و آرسنال نیز پس از گل صلاح به تمجید از ستاره مسلمان پرداخت و نوشت: کلاهمان را برای محمد صلاح باید برداریم.
۶- محمد صلاح برای لیورپول این کار را کرد تا این تیم به صدر جدول برسد ولی منچستر سیتی اجازه این کار را نداد.
۷- محمد صلاح تفاوت ها را رقم می زند و لیونل مسی جدید فوتبال است.
میزان مرتبط بودن سند:

این سند به نظر مرتبط ترین سند به نظر می رسد زیرا دفعات زیادی عبارت محمد صلاح در آن تکرار شده است اما باز هم با توجه به طول بالای آن در مکان پایین تری قرار گرفته است.

۱۱۹: برترین گلزنان دربی انگلیس که هنوز بازی می کنند/رونالدو در رده چهارم+عکس
<https://www.farsnews.ir/news/۱۴۰۰۰۸۰۲۰۰۰۳۱۲/> برترین-گلزنان-دربی-انگلیس-
که-هنوز-بازی-می-کنند-رونالدو-در-رده-چهارم

جملات حاوی کلمه:

۱- در رده دوم محمد صلاح ستاره مصری لیورپول ایستاده است. این بازیکن ۴ گل به ثمر رسانده است.

میزان مرتبط بودن سند:

در این سند صرفاً یک بار نام محمد صلاح ظاهر شده است و حضور آن در این رتبه منطقی به نظر می رسد.

۱۰۴: تحقیر منچستر به دست لیورپول / یک نفر به داد شیاطین سرخ برسد

<https://www.farsnews.ir/news/۱۴۰۰۰۸۰۲۰۰۰۴۲۸/> تحقیر-منچستر-به-دست-
لیورپول-یک-نفر-به-داد-شیاطین-سرخ-برسد

جملات حاوی کلمه:

۱- گل های این تیم در این جدال را که در ورزشگاه اولدترافورد برگزار شد، نابی کیتا (دقیقه ۵)، دیگو جوتا (۱۳) و محمد صلاح (۳۸، ۴۵+ و ۵۰) به ثمر رساندند.

میزان مرتبط بودن سند:

در این سند نیز صرفاً یک بار نام محمد صلاح آمده است و حضور آن در این رتبه منطقی است.

۳۸۴: چه کسی رئیس فدراسیون تکواندو می شود؟

<https://www.farsnews.ir/news/۱۴۰۰۰۷۲۸۰۰۰۲۷۸/> چه-کسی-رئیس-فدراسیون-
تکواندو-می-شود

جملات حاوی کلمه:

۱- ثبت نام برای حضور در انتخابات روز شنبه، ۲۴ مهرماه به پایان رسید و شهرام اربابی، علی باوقار، علی محمد بسحاق، عباس پورسلان، سید محمد پولادگر، محسن تهرانی، محمد جمشیدی، حسین حسین زاده، سیروس رضایی، هادی ساعی، محسن سترگی، جمشید سهرابی، آرش فرهادیان، وحید فرهنگ و پیوند گلزار برای حضور نام نویسی کردند
۲- براساس ماده ۱۰ دستورالعمل، لیست نفرات تایید شده مدیریتی برای طی مراحل قانونی مرتبط با اجزای "الف" و "ج"، بند "یک" ماده ۲۲ اساسنامه به مراجع ذی صلاح ارسال می شود.

میزان مرتبط بودن سند:

این سند به محمد صلاح مربوط نیست اما به دلیل ۳ بار تکرار شدن کلمه ی محمد و یک بار حضور صلاح، به عنوان یک سند مرتبط شناخته شده است. این در حالیست که ممکن است اخبار دیگری مرتبط با محمد صلاح وجود داشته باشد و حضور این سند در خروجی اصلاً مطلوب نیست.

بنابراین می توان گفت بجز سند آخر که سند مرتبطی به موضوع مورد نظر نبود، موتور جستجو عملکرد خوبی را از خود نشان داده است.

بخش دوم: مقایسه ی عملکرد Word Embedding در مدل آموزش داده شده توسط خودم با مدل آماده و tf-idf

در این بخش عملکرد موتور جستجو در ۳ حالت بررسی می شود:

۱- حالت word embedding با مدل آموزشی خودم

۲- حالت word embedding با مدل آماده

۳- حالت tf-idf

برای هر یک از حالت ها ۴ جستجوی با عبارات کوتاه، طولانی، پرسشی با کلمات رایج و پرسشی با کلمات نادر بررسی می شود و در هر یک ۵ نتیجه ی اول نمایش داده شده و میزان ارتباط هر جستجو نیز توضیح داده می شود و هر کدام برچسب گذاری می شوند. سپس برای هر یک از آن ها معیار های precision@k محاسبه می شود.

در نهایت برای هر حالت و با در نظر گرفتن تمامی پرسمان ها، معیار های mean average precision و mean reciprocal precision محاسبه می شوند.

الف) یک پرسمان کوتاه(یارانه):

الف (۱) عملکرد Word Embedding با مدل خودم

۴۷۱۲: بررسی طرح های مربوط به اصلاح نظام یارانه ای کشور در فراکسیون انقلاب اسلامی

Similarity: [[۰.۶۱۴۸۴۷۲۷]]

<https://www.isna.ir/news/۱۴۰۰۰۶۲۳۱۷۵۰۷/> بررسی طرح های مربوط به اصلاح

نظام یارانه ای کشور در فراکسیون

برچسب گذاری: صحیح

۵۲۴۳: بررسی طرح های مربوط به اصلاح نظام یارانه ای کشور در فراکسیون انقلاب اسلامی

Similarity: [[۰.۶۱۴۸۴۷۲۷]]

<https://www.isna.ir/news/۱۴۰۰۰۶۲۳۱۷۵۰۷/> بررسی طرح های مربوط به اصلاح

نظام یارانه ای کشور در فراکسیون

برچسب گذاری: صحیح

۳۰۳۱: حقیقت پور: کسانی که برای شرکت در یک کنسرت پول های هنگفت هزینه می کنند مستحق دریافت یارانه نیستند

Similarity: [[۰.۵۷۵۲۱۳۹۵]]

<https://www.farsnews.ir/news/۱۴۰۰۰۶۲۳۰۰۰۶۳۲/> حقیقت پور- کسانی که برای

شرکت در یک کنسرت پول های هنگفت هزینه می کنند

برچسب گذاری: صحیح

۵۳۱۰: پول پاشی ۱۲ هزار میلیاردی برای پردرآمدها!

Similarity: [[۰.۵۵۵۶۸۲۰۳]]

<https://www.isna.ir/news/۱۴۰۰۰۷۲۰۱۴۰۲۰/> پول-پاشی-۱۲-هزار-میلیاردی-برای-

پردرآمدها

برچسب گذاری: صحیح

۵۳۱۳: اوضاع عجیب یارانه‌های پنهان / ۲.۵ برابر بودجه دولت رفت!

Similarity: [[۰.۵۴۵۱۹۵۸۹]]

<https://www.isna.ir/news/۱۴۰۰۰۷۲۱۱۵۱۲۳/> اوضاع-عجیب-یارانه-های-پنهان-۲-۵-

برابر-بودجه-دولت-رفت

برچسب گذاری: صحیح

Precision@1: $1/1 = 1$

Precision@5: $5/5 = 1$

Average Precision = $1/5(1/1+2/2+3/3+4/4+5/5) = 1$

Reciprocal Rank = $1/1 = 1$

الف ۲) عملکرد Word Embedding با مدل آماده

۵۳۱۰: پول پاشی ۱۲ هزار میلیاردی برای پردرآمدها!

Similarity: [[۰.۶۸۹۰۶۱۲۴]]

<https://www.isna.ir/news/۱۴۰۰۰۷۲۰۱۴۰۲۰/> پول-پاشی-۱۲-هزار-میلیاردی-برای-

پردرآمدها

برچسب گذاری: صحیح

۵۲۰۰: یارانه بنزینی این بار در دولت جدید

Similarity: [[۰.۶۷۵۳۴۱۱]]

<https://www.isna.ir/news/۱۴۰۰۰۶۰۸۰۵۶۶۵/> یارانه-بنزینی-این-بار-در-دولت-جدید

برچسب گذاری: صحیح

۴۷۱۲: بررسی طرح‌های مربوط به اصلاح نظام یارانه ای کشور در فراکسیون انقلاب اسلامی

Similarity: [[۰.۶۲۸۴۶۱۷۹]]

<https://www.isna.ir/news/۱۴۰۰۰۶۲۳۱۷۵۰۷/> بررسی-طرح-های-مربوط-به-اصلاح-

نظام-یارانه-ای-کشور-در-فراکسیون

برچسب گذاری: صحیح

۵۲۴۳: بررسی طرح های مربوط به اصلاح نظام یارانه ای کشور در فراکسیون انقلاب اسلامی

Similarity: [[۰.۶۲۸۴۶۱۷۹]]

<https://www.isna.ir/news/۱۴۰۰۰۶۲۳۱۷۵۰۷/> بررسی طرح های مربوط به اصلاح نظام یارانه ای کشور در فراکسیون برچسب گذاری: صحیح

۳۰۳۱: حقیقت پور: کسانی که برای شرکت در یک کنسرت پول های هنگفت هزینه می کنند مستحق دریافت یارانه نیستند

Similarity: [[۰.۶۲۶۸۳۵۲۸]]

<https://www.farsnews.ir/news/۱۴۰۰۰۶۲۳۰۰۰۶۳۲/> حقیقت پور- کسانی که برای شرکت در یک کنسرت پول های هنگفت هزینه می کنند برچسب گذاری: صحیح

Precision@1: $1/1 = 1$

Precision@5: $5/5 = 1$

Average Precision = $1/5(1/1+2/2+3/3+4/4+5/5) = 1$

Reciprocal Rank = $1/1 = 1$

الف (۳) عملکرد tf-idf

۱۹۵۷: عضو هیات رئیسه مجلس: سالانه ۱۶۵ میلیارد دلار در کشور یارانه پرداخت می شود

<https://www.farsnews.ir/news/۱۴۰۰۰۷۲۶۰۰۰۲۱۵/> عضو هیات رئیسه مجلس- سالانه ۶۵ میلیارد دلار در کشور یارانه پرداخت برچسب گذاری: صحیح

۴۷۱۲: بررسی طرح های مربوط به اصلاح نظام یارانه ای کشور در فراکسیون انقلاب اسلامی

<https://www.isna.ir/news/۱۴۰۰۰۶۲۳۱۷۵۰۷/> بررسی طرح های مربوط به اصلاح نظام یارانه ای کشور در فراکسیون برچسب گذاری: صحیح

۵۲۴۳: بررسی طرح های مربوط به اصلاح نظام یارانه ای کشور در فراکسیون انقلاب اسلامی

<https://www.isna.ir/news/۱۴۰۰۰۶۲۳۱۷۵۰۷/> بررسی طرح های مربوط به اصلاح نظام یارانه ای کشور در فراکسیون برچسب گذاری: صحیح

۴۹۱۷: زمان واریز یارانه و کسر قسط یازدهم وام کرونا
<https://www.isna.ir/news/۱۴۰۰۲۱۹۱۳۱۸۷/> زمان-واریز-یارانه-و-کسر-قسط-یازدهم-
وام-کرونا
برچسب گذاری: صحیح

۳۰۳۱: حقیقت پور: کسانیکه برای شرکت در یک کنسرت پول های هنگفت هزینه
می کنند مستحق دریافت یارانه نیستند
<https://www.farsnews.ir/news/۱۴۰۰۰۶۲۳۰۰۰۶۳۲/> حقیقت پور-کسانیکه-برای-
شرکت-در-یک-کنسرت-پول-های-هنگفت-هزینه-می-کنند
برچسب گذاری: صحیح

Precision@1: $1/1 = 1$

Precision@5: $5/5 = 1$

Average Precision = $1/5(1/1+2/2+3/3+4/4+5/5) = 1$

Reciprocal Rank = $1/1 = 1$

بنابراین هر ۳ موتور جستجو عملکرد خوبی را در این بخش از خود نشان داده اند.

(ب) یک پرسمان طولانی (آمار واکسیناسیون در کشور):

ب (۱) عملکرد Word Embedding با مدل خودم
۵۷۶۵: سامانه ثبت نام واکسن کرونا برای متولدین ۱۳۴۰ بازگشایی شد
Similarity: $[[0.51606016]]$

<https://www.isna.ir/news/۱۴۰۰۰۴۲۳۱۷۰۲۷/> سامانه-ثبت-نام-واکسن-کرونا-برای-
متولدین-۱۳۴۰-بازگشایی-شد
برچسب گذاری: اشتباه

۵۷۷۱: سامانه ثبت نام واکسن کرونا برای متولدین ۱۳۴۰ بازگشایی شد
Similarity: $[[0.51606016]]$

<https://www.isna.ir/news/۱۴۰۰۰۴۲۳۱۷۰۲۷/> سامانه-ثبت-نام-واکسن-کرونا-برای-
متولدین-۱۳۴۰-بازگشایی-شد
برچسب گذاری: اشتباه

۵۸۱۷: واکسیناسیون افراد بالای ۵۰ سال در قم آغاز شد
Similarity: $[[0.5181118]]$

<https://www.isna.ir/news/۱۴۰۰۰۵۱۱۰۷۵۳۴/> واکسیناسیون-افراد-بالای-۵۰-سال-در-
قم-آغاز-شد
برچسب گذاری: اشتباه

۶۱۱۸: تداوم واردات واکسن از سوی هلال احمر / ۶ میلیون دُز سینوفارم وارد کشور شد

Similarity: [[0.50911932]]

<https://www.isna.ir/news/140072517310/> تداوم-واردات-واکسن-از-سوی-هلال-
احمر-۶-میلیون-دز-سینوفارم-وارد
برچسب گذاری: صحیح

۵۹۶۴: اعلام استان های صدرنشین در واکسیناسیون کرونا

Similarity: [[0.50885452]]

<https://www.isna.ir/news/140061510988/> اعلام-استان-های-صدرنشین-در-
واکسیناسیون-کرونا
برچسب گذاری: صحیح

Precision@1: 0/1 = 0

Precision@5: 2/5 = 0.4

Average Precision = 1/2(1/4 + 2/5) = 0.325

Reciprocal Rank = 1/4 = 0.25

ب ۲) عملکرد Word Embedding با مدل آماده

۵۹۶۴: اعلام استان های صدرنشین در واکسیناسیون کرونا

Similarity: [[0.62708558]]

<https://www.isna.ir/news/140061510988/> اعلام-استان-های-صدرنشین-در-
واکسیناسیون-کرونا
برچسب گذاری: صحیح

۱۷۶۸: کاهش ۹۵ درصدی تعداد شهدای کادر درمان در پیک پنجم کرونا

Similarity: [[0.59823627]]

<https://www.farsnews.ir/news/140071100772/> کاهش-۹۵-درصدی-تعداد-
شهدای-کادر درمان-در-پیک-پنجم-کرونا
برچسب گذاری: اشتباه

۵۸۰۲: ۳۰۳ فوتی جدید کرونا در کشور / تزریق بیش از ۴۳۶ هزار دُز واکسن در شبانه
روز گذشته

Similarity: [[0.57606266]]

<https://www.isna.ir/news/140050604257/> ۳۰۳-فوتی-جدید-کرونا-در-کشور-
تزریق-بیش-از-۴۳۶-هزار-دز-واکسن-در-
برچسب گذاری: صحیح

۵۹۷۰: رکوردشکنی دوباره تزریق واکسن کرونا در ایران / تزریق ۱۱۱۲۴۴۰ دُز در شبانه روز گذشته

Similarity: [[۰.۵۷۳۸۷۰۳۶]]

<https://www.isna.ir/news/۱۴۰۰۶۱۷۱۲۹۴۹/> رکوردشکنی-دوباره-تزریق-واکسن-کرونا-در-ایران-تزریق-۱۱۱۲۴۴۰-دز
برچسب گذاری: صحیح

۵۸۴۴: آمارهای جدید کرونا در ایران / ۵۰۸ تن دیگر جان باختند

Similarity: [[۰.۵۷۱۴۰۶۴۳]]

<https://www.isna.ir/news/۱۴۰۰۵۱۹۱۳۷۸۳/> آمارهای-جدید-کرونا-در-ایران-۵۰۸-تن-دیگر-جان-باختند
برچسب گذاری: صحیح

Precision@1: $1/1 = 1$

Precision@5: $4/5 = 0.8$

Average Precision = $1/5(1/1+2/3+3/4+4/5) = 0.643333$

Reciprocal Rank = $1/1 = 1$

ب (۳) عملکرد tf-idf

۵۹۱۱: مرکز آمار: گزارشی از واکسیناسیون و مرگ ۲۰۰۰ نفری ندادیم

<https://www.isna.ir/news/۱۴۰۰۶۰۱۰۰۳۸۴/> مرکز-آمار-گزارشی-از-واکسیناسیون-و-مرگ-۲۰۰۰-نفری-ندادیم
برچسب گذاری: صحیح

۵۹۶۴: اعلام استان‌های صدرنشین در واکسیناسیون کرونا

<https://www.isna.ir/news/۱۴۰۰۶۱۵۱۰۹۸۸/> اعلام-استان-های-صدرنشین-در-واکسیناسیون-کرونا
برچسب گذاری: صحیح

۲۰۸۷: امیدواری مخبر به حمایت نخبگان و همراهی مردم در واکسیناسیون عمومی

<https://www.farsnews.ir/news/۱۴۰۰۷۲۱۰۰۰۸۰۴/> امیدواری-مخبر-به-حمایت-نخبگان-و-همراهی-مردم-در-واکسیناسیون-عمومی
برچسب گذاری: صحیح

۳۱۹۹: تلاش برای افزایش آمار روزانه واکسیناسیون به ۲ میلیون نفر

<https://www.farsnews.ir/news/۱۴۰۰۶۲۰۰۰۰۴۶۱/> تلاش-برای-افزایش-آمار-روزانه-واکسیناسیون-به-۲-میلیون-نفر
برچسب گذاری: صحیح

۵۱۴۹: خودروسازان از وزارت بهداشت واکسن می‌خرند / هر دوز ۲۰۰ هزار تومان!
<https://www.isna.ir/news/۱۴۰۰۰۵۱۹۱۴۰۰۴>/خودروسازان-از-وزارت-بهداشت-واکسن-
می-خرند-هر-دوز-۲۰۰-هزار-تومان
برچسب گذاری: اشتباه

Precision@1: $1/1 = 1$

Precision@5: $4/5 = 0.8$

Average Precision = $1/5(1/1+2/2+3/3+4/4) = 0.8$

Reciprocal Rank = $1/1 = 1$

بنابراین با توجه به مقادیر به دست آمده، به نظر می‌رسد که در این پرسمان طولانی روش‌های مدل آماده و tf-idf عملکرد بسیار بهتری داشته‌اند. اما عملکرد tf-idf با توجه به اینکه سند های مرتبط تر را بالاتر نشان داده است و Average Precision بیشتری داشته است، بهتر بوده.

نتایج منطقی به نظر می‌رسند زیرا در روش tf-idf که چمپیونز لیست و doc coverage در آن اعمال شده، سند هایی برگردانده می‌شوند که حداکثر پوشش کلمات یکسان را با پرسمان داشته باشند و کلمات پرسمان بیشترین تکرار را در این سند ها دارند. از طرفی دو مدل دیگر با توجه به اینکه بر اساس شباهت بین کلمات آموزش داده شده‌اند، ممکن است سند های مشابهی با پرسمان خواسته شده را پیدا کنند اما به طور دقیق خواسته ی مورد نظر ما را پوشش ندهند. البته مدل آماده در این پرسمان به خوبی عمل کرده است و این می‌تواند به دلیل این باشد که با تعداد کلمات بیشتری آموزش داده شده و به همین علت دقت آن افزایش یافته است.

پ) یک عبارت پرسشی با کلمات رایج (آیا مسی بهترین بازیکن جهان است؟)

پ ۱) عملکرد Word Embedding با مدل خودم

۱۱۹۴: خط قرمز مسی روی رونالدو/کدام بازیکن شایسته کسب توپ طلا است؟

Similarity: [[۰.۷۴۸۸۶۱۴۱]]

<https://www.farsnews.ir/news/۱۴۰۰۰۷۱۷۰۰۰۱۴۵/>خط-قرمز-مسی-روی-رونالدو-
کدام-بازیکن-شایسته-کسب-توپ-طلا-است
برچسب گذاری: صحیح

۱۳۷: ال کلاسیکو با رقابت جانشینان مسی و رونالدو+عکس

Similarity: [[۰.۷۴۶۰۸۵۰۱]]

<https://www.farsnews.ir/news/۱۴۰۰۰۸۰۲۰۰۰۱۴۱/>ال-کلاسیکو-با-رقابت-
جانشینان-مسی-و-رونالدو-عکس
برچسب گذاری: اشتباه

۹۵۷: «گاوی» جا پای مسی گذاشت +عکس

Similarity: [[0.۷۴۱۹۰۳۰۷]]

<https://www.farsnews.ir/news/۱۴۰۰۰۷۲۰۰۰۰۳۸۱/> گاوی-جا-پای-مسی-گذاشت-

عکس

برچسب گذاری: اشتباه

۱۱۱۳: هفتمین توپ طلا در انتظار مسی

Similarity: [[0.۷۳۹۸۱۰۶۸]]

<https://www.farsnews.ir/news/۱۴۰۰۰۷۱۸۰۰۰۲۷۰/> هفتمین-توپ-طلا-در-انتظار-

مسی

برچسب گذاری: صحیح

۱۲۳: زمان ظهور اساطیر جدید در ال کلاسیکو فرا رسید+عکس

Similarity: [[0.۷۳۳۱۸۶۸]]

<https://www.farsnews.ir/news/۱۴۰۰۰۸۰۲۰۰۰۲۷۷/> زمان-ظهور-اساطیر-جدید-در-

ال-کلاسیکو-فرا-رسید-عکس

برچسب گذاری: اشتباه

Precision@1: $1/1 = 1$

Precision@5: $2/5 = 0.4$

Average Precision = $1/5(1/1+2/4) = 0.3$

Reciprocal Rank = $1/1 = 1$

پ ۲) عملکرد Word Embedding با مدل آماده

۱۱۹۴: خط قرمز مسی روی رونالدو/کدام بازیکن شایسته کسب توپ طلا است؟

Similarity: [[0.۷۸۵۴۶۸۵۹]]

<https://www.farsnews.ir/news/۱۴۰۰۰۷۱۷۰۰۰۱۴۵/> خط-قرمز-مسی-روی-رونالدو-

کدام-بازیکن-شایسته-کسب-توپ-طلا-است

برچسب گذاری: صحیح

۱۹۶: برتری عجیب مسی مقابل رونالدو+عکس

Similarity: [[0.۷۷۳۱۲۵۲۱]]

<https://www.farsnews.ir/news/۱۴۰۰۰۸۰۱۰۰۰۱۶۷/> برتری-عجیب-مسی-مقابل-

رونالدو-عکس

برچسب گذاری: صحیح

۳۹۹:مسی از رونالدو جلو زد+عکس

Similarity: [[0.۷۷۱۱۴۰۴۴]]

<https://www.farsnews.ir/news/۱۴۰۰۰۷۲۸۰۰۰۰۵۷/>مسی-از-رونالدو-جلو-زد-عکس

برچسب گذاری: صحیح

۱۰۰۷:چرا مسی مستحق کسب توپ طلای هفتم است؟

Similarity: [[0.۷۶۴۵۸۸۱۳]]

<https://www.farsnews.ir/news/۱۴۰۰۰۷۱۹۰۰۰۰۶۸۷/>چرا-مسی-مستحق-کسب-توپ-

طلای-هفتم-است

برچسب گذاری: صحیح

۹۷۲:آمار جالب مسی در گلزنی با آلبی سلسه

Similarity: [[0.۷۶۰۲۵۷۸۸]]

<https://www.farsnews.ir/news/۱۴۰۰۰۷۲۰۰۰۰۰۰۸۲/>آمار-جالب-مسی-در-گلزنی-با-

آلبی-سلسه

برچسب گذاری: اشتباه

Precision@1: $1/1 = 1$

Precision@5: $4/5 = 0.8$

Average Precision = $1/5(1/1+2/2+3/3+4/4) = 0.8$

Reciprocal Rank = $1/1 = 1$

پ ۳) عملکرد tf-idf

۳۹۷:بهترین بازیکن جهان از دید امبایه

<https://www.farsnews.ir/news/۱۴۰۰۰۷۲۸۰۰۰۰۰۹۲/>بهترین-بازیکن-جهان-از-دید-

امبایه

برچسب گذاری: صحیح

۱۱۱۳:هفتمین توپ طلا در انتظار مسی

<https://www.farsnews.ir/news/۱۴۰۰۰۷۱۸۰۰۰۰۲۷۰/>هفتمین-توپ-طلا-در-انتظار-

مسی

برچسب گذاری: صحیح

۸۶۲:در تعداد هت تریک مسی برتر است یا رونالدو؟+عکس

<https://www.farsnews.ir/news/۱۴۰۰۰۷۲۱۰۰۰۰۶۸۳/>در-تعداد-هت-تریک-مسی-

برتر-است-یا-رونالدو-عکس

برچسب گذاری: صحیح

۹۳۴: فهرست برترین گلزنان ملی فعال دنیای فوتبال / یک اماراتی و هندی در مکان سوم و چهارم+عکس

<https://www.farsnews.ir/news/۱۴۰۰۰۷۲۰۰۰۰۷۵۹/>فهرست-برترین-گلزنان-ملی-
فعال-دنیای-فوتبال-یک-اماراتی-و-هندی-در-مکان
برچسب گذاری: اشتباه

۱۳۱۳: جنجال آفرینی مادر امپایه؛ پایان شایعه اختلاف با مسی تا آب پاکی روی دست
رئال

<https://www.farsnews.ir/news/۱۴۰۰۰۷۱۵۰۰۰۲۰۱/>جنجال-آفرینی-مادر-
امپایه-پایان-شایعه-اختلاف-با-مسی-تا-آب-پاکی-روی-دست
برچسب گذاری: اشتباه

Precision@1: $1/1 = 1$

Precision@5: $3/5 = 0.6$

Average Precision = $1/5(1/1+2/2+3/3) = 0.6$

Reciprocal Rank = $1/1 = 1$

با توجه به این که نتایج به هم نزدیک بود، k را به ۱۰ افزایش دادیم و در نهایت دو موتور دارای word embedding عملکرد نسبتاً بهتری به tf-idf داشتند. این اتفاق ممکن است بخاطر این باشد که آن موتورهای جستجو، کلماتی مثل توپ طلا را مرتبط با پیرسمان دانسته اند و سند هایی مشابه به پیرسمان مورد جستجو را پیدا کرده اند.

ت) یک عبارت پرسشی با کلمات نادر (آیا سالن های ژیمناستیک غیر قانونی هستند؟)

ت (۱) عملکرد Word Embedding با مدل خودم

۳۸۷۹: جزییات تعطیلی ورزش ایران تا پایان تیرماه + تصویر

Similarity: $[[0.58197522]]$

<https://www.isna.ir/news/۱۴۰۰۰۴۲۷۱۹۸۳۳/>جزئیات-تعطیلی-ورزش-ایران-
تا-پایان-تیرماه-تصویر

برچسب گذاری: اشتباه

۴۱۸۹: جزییات تعطیلی های ورزش ایران تا ۹ مهر ۱۴۰۰ / کدام فعالیت های ورزشی در
تهران ممنوع است؟

Similarity: $[[0.56244746]]$

<https://www.isna.ir/news/۱۴۰۰۷۰۲۰۰۸۰۸/> جزئیات-تعطیلی-های-ورزش-
ایران-تا-۹-مهر-۱۴۰۰-کدام-فعالیت-های-ورزشی
برچسب گذاری: اشتباه

۱۳۶۸: هشدار هیات ژیمناستیک تهران در خصوص سالن های مختلط و اقدامات
غیراخلاقی

Similarity: [[۰.۵۲۸۸۱۷۵۷]]

<https://www.farsnews.ir/news/۱۴۰۰۷۱۴۰۰۲۳۱/> هشدار-هیات-
ژیمناستیک-تهران-در-خصوص-سالن های-مختلط-و-اقدامات-غیراخلاقی
برچسب گذاری: صحیح

۳۵۳۳: تعطیلی تمامی مجموعه های ورزشی شهرداری تهران

Similarity: [[۰.۴۹۱۴۲۸۹۹]]

<https://www.isna.ir/news/۱۴۰۰۱۱۶۰۷۴۰۹/> تعطیلی-تمامی-مجموعه-های-
ورزشی-شهرداری-تهران
برچسب گذاری: اشتباه

۳۴۸۵: رشته ورزشی زیر تیغ پروتکل های کرونا

Similarity: [[۰.۴۸۰۱۳۶۱۴]]

<https://www.isna.ir/news/۰۰۰۶۱۶۳۵۳۰۶/> ۲۳-رشته-ورزشی-زیر-تیغ-
پروتکل-های-کرونا
برچسب گذاری: اشتباه

Precision@1: 0/1 = 0

Precision@5: 1/5 = 0.2

Average Precision = 1/1(1/5) = 0.2

Reciprocal Rank = 1/5 = 0.2

ت ۲) عملکرد Word Embedding با مدل آماده

۴۱۸۹: جزئیات تعطیلی های ورزش ایران تا ۹ مهر ۱۴۰۰ / کدام فعالیت های ورزشی در
تهران ممنوع است؟

Similarity: [[۰.۶۱۷۱۸۰۶]]

<https://www.isna.ir/news/۱۴۰۰۷۰۲۰۰۸۰۸/> جزئیات-تعطیلی-های-ورزش-
ایران-تا-۹-مهر-۱۴۰۰-کدام-فعالیت-های-ورزشی
برچسب گذاری: اشتباه

۳۸۷۹: جزئیات تعطیلی ورزش ایران تا پایان تیرماه + تصویر

Similarity: [[0.61695249]]

<https://www.isna.ir/news/140042719833/> جزئیات-تعطیلی-ورزش-ایران-
تا-پایان-تیرماه-تصویر
برچسب گذاری: اشتباه

۲۳۳۴۸۵: رشته ورزشی زیر تیغ پروتکل های کرونا

Similarity: [[0.57599153]]

<https://www.isna.ir/news/0061635306/> ۲۳-رشته-ورزشی-زیر-تیغ-
پروتکل-های-کرونا
برچسب گذاری: اشتباه

۱۳۶۹۲: اعضای مجمع انتخاباتی فدراسیون زورخانه ای مشخص شدند

Similarity: [[0.52720701]]

<https://www.isna.ir/news/140030403266/> اعضای-مجمع-انتخاباتی-
فدراسیون-زورخانه-ای-مشخص-شدند
برچسب گذاری: اشتباه

۳۶۱۶: دبیر مجمع فدراسیون ژیمناستیک مشخص شد

Similarity: [[0.52580548]]

<https://www.isna.ir/news/140020906408/> دبیر-مجمع-فدراسیون-
ژیمناستیک-مشخص-شد
برچسب گذاری: اشتباه

Precision@1: 0/1 = 0

Precision@5 0/5 = 0

Average Precision = 0

Reciprocal Rank = 0/1 = 0

ت ۳) عملکرد tf-idf

۱۳۶۸: هشدار هیات ژیمناستیک تهران در خصوص سالن های مختلط و اقدامات
غیراخلاقی

<https://www.farsnews.ir/news/140071400231/> هشدار-هیات-
ژیمناستیک-تهران-در-خصوص-سالن های-مختلط-و-اقدامات-غیراخلاقی
برچسب گذاری: صحیح

۳۶۱۶: دبیر مجمع فدراسیون ژیمناستیک مشخص شد

<https://www.isna.ir/news/۱۴۰۰۲۰۹۰۶۴۰۸/> دبیر-مجمع-فدراسیون-

ژیمناستیک-مشخص-شد
برچسب گذاری: اشتباه

۳۸۷۹: جزییات تعطیلی ورزش ایران تا پایان تیرماه+ تصویر
<https://www.isna.ir/news/۱۴۰۰۴۲۷۱۹۸۳۳/> جزییات-تعطیلی-ورزش-ایران-
تا-پایان-تیرماه-تصویر
برچسب گذاری: اشتباه

۴۱۸۹: جزییات تعطیلی های ورزش ایران تا ۹ مهر ۱۴۰۰ / کدام فعالیت های ورزشی در
تهران ممنوع است؟
<https://www.isna.ir/news/۱۴۰۰۷۰۲۰۰۸۰۸/> جزییات-تعطیلی-های-ورزش-
ایران-تا-۹-مهر-۱۴۰۰-کدام-فعالیت-های-ورزشی
برچسب گذاری: اشتباه

۱۴۵۷: سوت زنی | تذکر و برخورد با سالن های ورزشی مختلط توسط وزارت ورزش
<https://www.farsnews.ir/news/۱۴۰۰۷۱۲۰۰۶۵۸/> سوت زنی | تذکر-و-
برخورد-با-سالن های-ورزشی-مختلط-توسط-وزارت-ورزش
برچسب گذاری: صحیح

Precision@1: $1/1 = 1$

Precision@5: $2/5 = 0.4$

Average Precision = $1/2(1 + 0.4) = 0.7$

Reciprocal Rank = $1/1 = 1$

در این پرسمان به وضوح عملکرد tfidf بهتر بوده است و با اینکه کلا ۲ سند مربوط به پرسمان وجود داشته است هر دو را برگردانده است. با توجه به اینکه کلمه ی ژیمناستیک نادر است و کلماتی نظیر سالن و غیر قانونی نیز در پرسمان وجود دارند طبیعی است که عملکرد tfidf بهتر باشد چون به طوری عمل می کند که اکثر کلمات پرسمان در سند باشند.

محاسبه ی Mean Reciprocal Rank و Mean Average Precision:

الف) برای مدل خودم

$MAP = \frac{1}{4}(1+0.325+0.3+0.2) = 0.45625$

$MRR = \frac{1}{4}(1+0.25+1+0.2) = 0.6125$

ب) برای مدل آماده

$$\text{MAP} = \frac{1}{4}(1+0.643+0.8+0) = 0.61075$$

$$\text{MRR} = \frac{1}{4}(1+1+1+0) = 0.75$$

ج) برای tfidf

$$\text{MAP} = \frac{1}{4}(1+0.8+0.7+0.6) = 0.775$$

$$\text{MRR} = \frac{1}{4}(1+1+1+1) = 1$$

در نهایت به نظر می رسد به ترتیب موتور جستجوهای tfidf، مدل آماده و مدل خودم عملکرد بهتری از خود نشان داده اند.