

به نام یاور

سید عارف طباطبایی 9831040

TFIDF:

ابتدا توابع مربوط به محاسبه tfidf را پیاده سازی میکنیم:

```
# TFIDF Score
def calculate_tfidf(term_frequency, document_frequency, collection_size):
    if document_frequency != 0:
        return math.log(term_frequency + 1) * math.log(collection_size / document_frequency)
    else:
        return None

def tfidf(term, document, collection):
    # term frequency
    term_frequency = document['content'].count(term)
    # document frequency
    document_frequency = sum([1 for _, doc in collection.items() if term in doc['content']])
```

Index:

سپس باید آن را به index اضافه کنیم:

```
def update_index(ind, collection_size):
    for token, token_dict in ind.items():
        document_frequency = token_dict['num']

        for docID, positions in token_dict['positions'].items():
            term_frequency = positions[0]
            tfidf_score = calculate_tfidf(term_frequency, document_frequency, collection_size)
            positions.append(tfidf_score)
```

در ادامه champion list را نیز اضافه می کنیم:

```
def update_index(ind, collection_size, champion_lists_ratio=0, champion_list_enable=True):
    for token, token_dict in ind.items():
        document_frequency = token_dict['num']
        champion_list_size = int(champion_lists_ratio * document_frequency)

        if champion_list_enable:
            champion_list = []

        for docID, positions in token_dict['positions'].items():
            term_frequency = positions[0]
            tfidf_score = calculate_tfidf(term_frequency, document_frequency, collection_size)
            positions.append(tfidf_score)
            # preprocessed_data[docID]['length'] += tfidf_score ** 2

            if champion_list_enable:
                champion_list.append((docID, tfidf_score))

        if champion_list_enable:
            champion_list.sort(key=lambda x: x[1], reverse=True)
            token_dict['champion list'] = champion_list[:champion_list_size]
```

در نتیجه ساختار index به صورت زیر در میآید:

```
index structure
Token
    num
    positions
        doc
        total num in every doc
        positions
        tfidf
    champion list
```

در ادامه توابع similarity را پیاده سازی میکنیم و مطابق فرمت زیر کوئری ها را هندل میکنیم:

```

# Similarity Funcs

# Cosine Similarity
def cosine_similarity(query, k, champion_list_enable=False):...

# Jaccard Similarity
def calculate_jaccard_similarity(set1, set2):...

def jaccard_similarity(query, k):...

def run_query(query, similarity_type):

    print("query: ", query)
    print("similarity function: ", similarity_type)

    if similarity_type == "Cosine":
        result = cosine_similarity(preprocess(query), 5, False)
    else:
        result = jaccard_similarity(preprocess(query), 5)

    for r in result:
        print(f"Document ID is: {r[0]}")
        print(f"Document Score is: {r[1]}")
        print(preprocessed_data[f'{r[0]}']['title'])
        print(preprocessed_data[f'{r[0]}']['url'])

```

کوئری ها:

یک پرسمان از کلمات ساده و متداول تک کلمه ای:

اسناد:

• معیار شباهت کسینوسی

```
query: اسناد
similarity function: Cosine
Document ID is: 792
Document Score is: 1.1272797865330033
بارگذاری 40 سند مالی سرخایی ها در سایت کدال
https://www.farsnews.ir/news/14001214000395/
Document ID is: 11735
Document Score is: 1.1025472847969588
اسناد مجالس شواری ملی، سنا و اسلامی ساماندهی می‌شود
https://www.farsnews.ir/news/14000809000290/
Document ID is: 7149
Document Score is: 1.0640900794642718
اختصاص اعتبار ۶۰۰ میلیاردی برای تهیه نقشه و اسناد مالکیت جنگاری اراضی کشاورزی
https://www.farsnews.ir/news/14001216000272/
Document ID is: 7339
Document Score is: 1.0238837966626233
متقاضیان اسناد تک برگ در روستاها از پرداخت هزینه برگ سند معاف شدند
https://www.farsnews.ir/news/14001209000779/
Document ID is: 11688
Document Score is: 1.0150058495523142
روتمایی از کتاب «روایت تسخیر» و سامانه اسناد جاسوسی آمریکا به‌مناسبت ۱۳ آبان
https://www.farsnews.ir/news/14000810000216/
```

792:

به گزارش خبرنگار ورزشی خبرگزاری فارس، حسین قربانزاده، رئیس سازمان خصوصی سازی در نشست خبری آغاز واگذاری سهام استقلال و پرسپولیس به مردم در خصوص انتشار **اسناد** دو باشگاه برای شفافیت مالی اعلام کرد: استقلال و پرسپولیس حدود ۴۰ سند را در سایت کدال در مورد **اسناد** مالی خود منتشر کردند تا شفافیت مالی لازم ارائه شود.

نکته جالب این است که رئیس سازمان خصوصی سازی اعلام کرد نظم باشگاه استقلال در زمینه ارائه صورت های مالی بیشتر بوده است و حتی **اسناد** بیشتری را هم ارائه کرده است.

11735:

به گزارش حوزه پارلمانی خبرگزاری فارس، علی اکبر زارع بیدکی درباره طرح ساماندهی **اسناد** مجالس شورای ملی، سنا و اسلامی، گفت: متأسفانه به دلیل کمبود فضا، این **اسناد** در سال‌های گذشته، خارج از استانداردهای لازم در انبار و مکانی دور از کتابخانه نگهداری شده‌اند. این مجموعه مربوط به دهه ۶۰ تا ۹۰، دوره‌های اول تا هشتم **مجلس شورای اسلامی** بوده و ساماندهی اولیه و تفکیک این منابع توسط کارشناسان خبره فراهم‌آوری و ساماندهی منابع، از اواخر مهرماه آغاز شده است.

مدیر مجموعه‌سازی کتابخانه مجلس افزود: در بین این منابع، مجموعه طرح‌ها و لوایح کمیسیون‌های مختلف مجلس و همچنین بولتن‌ها و آمارنامه‌های بسیار مهم سازمان‌های دولتی و خصوصی دیده می‌شود. این منابع جهت تکمیل بخش گزارشات کتابخانه، بسیار مهم و ارزشمند هستند.

زارع بیدکی با بیان اینکه همزمان با ساماندهی این مجموعه عظیم، ادامه ساماندهی **اسناد** کمیسیون اصل نود مربوط به دوره‌های سوم تا ششم مجلس شورای اسلامی در حال انجام است، ادامه داد: تاکنون ساماندهی و برگ شماری بیش از ۱۲ هزار جلد پرونده از این مجموعه به اتمام رسیده و پیش‌بینی می‌شود تا پایان بهمن ماه ۱۴۰۰، کار مجموعه **اسناد** اصل نود به پایان برسد.

وی افزود: این **اسناد** شامل عرایض و شکواییه‌های مردمی بوده و نامه‌هایی است که در آن مردم از شهرهای مختلف و طبقات مختلف جامعه، مطالبات و مشکلات خود را مطرح کرده‌اند. این نوع منابع دست اول به دلیل آنکه از متن توده‌های مردم و مناطق مختلف کشور برخاسته، مشتمل بر اطلاعات بی‌نظیر و منحصر به فردی است که می‌تواند جز برترین **اسناد** جهت استفاده پژوهشگران باشد.

مدیر مجموعه‌سازی کتابخانه مجلس یادآور شد در سال گذشته نیز ۵۵ هزار جلد پرونده از مجموعه **اسناد** مذکور مربوط به دوره‌های سوم تا هفتم قانونگذاری برگ‌شماری و ساماندهی شده و به مخازن مربوطه ارسال شده است.

وی در مورد اسکن این **اسناد** مهم گفت: دیجیتال‌سازی و اسکن منابع **اسناد** مجلس شورای ملی و مجلس شورای اسلامی با سرعت در حال اجرا است. با اسکن بیش از ۹۰۰ هزار برگ سند در سال جاری، اسکن این منابع به بیش از ۲ میلیون و یکصد هزار برگ رسیده است. پیش‌بینی می‌شود در صورت تأمین بودجه، اسکن **اسناد** مجلس شورای ملی تا پایان سال جاری به اتمام برسد.

یکی از مهم‌ترین و معتبرترین **اسناد** در حوزه تاریخ پارلمان، **اسناد** موجود در کتابخانه مجلس شورای اسلامی است. این **اسناد** که بیش از ۲ میلیون برگ است، شامل **اسناد** ادوار دوم تا بیست و چهارم مجلس شورای ملی، هفت دوره مجلس سنا و **اسناد** ادوار مجلس شورای اسلامی است. متأسفانه **اسناد** دوره اول مجلس شورای ملی به دلیل به توپ بستن مجلس به دستور محمد علی شاه از بین رفته است.

این مجموعه **اسناد** که در سال‌های گذشته توسط مجلس شورای اسلامی به کتابخانه واگذار گردیده، شامل **اسناد** کاربرداری، **اسناد** حسابداری، **اسناد** کارگزینی، دفاتر ثبت، مکاتبات داخلی مجلس، نامه‌های مهم ارسالی از شخصیت‌های سیاسی کشورهای مختلف به رؤسای وقت مجلس شورای اسلامی و **اسناد** انتخابات شورایها است.

:7149

به گزارش خبرنگار پارلمانی خبرگزاری فارس، نمایندگان مجلس در نشست علنی صبح امروز دوشنبه (۱۶ اسفند) مجلس شورای اسلامی در روند بررسی بخش هزینه‌ای لایحه بودجه سال ۱۴۰۱ کل کشور، با بند (د) تبصره (۱۰) این لایحه موافقت کردند.

براساس بند (د) تبصره (۱۰) ماده واحده لایحه بودجه ۱۴۰۱ کل کشور، سازمان ثبت **اسناد** و املاک کشور مکلف است با همکاری وزارت جهاد کشاورزی (سازمان امور اراضی) به منظور تهیه نقشه و اطلاعات توصیفی و مکانی اراضی کشاورزی، در صورت درخواست کشاورزان متقاضی صدور سند مالکیت، به ازای هر هکتار مبلغ پانصد هزار (۵۰۰.۰۰۰) ریال از متقاضی صدور سند مالکیت مفروز اراضی کشاورزی اخذ و به حساب خزانه داری کل کشور واریز و معادل صد در صد (۱۰۰٪) تا سقف شش هزار میلیارد (۶.۰۰۰.۰۰۰.۰۰۰.۰۰۰) ریال طی ردیف ۲۰-۷۳۰۰۰۰۰ در اختیار سازمان ثبت **اسناد** و املاک کشور قرار داده تا با استفاده از ظرفیتهای بخش غیر دولتی نسبت به تهیه نقشه‌های مربوط اقدام و **اسناد** مالکیت حدنگاری اراضی کشاورزی را صادر نماید.

سازمان ثبت و **اسناد** املاک کشور موظف است همزمان با صدور **اسناد** تک برگ، نسبت به ارائه اطلاعات مکان محور این **اسناد** به شرکت پست جمهوری اسلامی ایران اقدام کند.

به گزارش خبرنگار ارلمانی خبرگزاری فارس، نمایندگان مجلس شورای اسلامی در جلسه علنی امروز (دوشنبه ۹ اسفند) در بررسی بودجه ۱۴۰۱ در بخش درآمدی با الحاق یک بند به تبصره ۱۰ موافقت کردند.

این بند می گوید:

به منظور تسهیل اجرای ماده ۱۱ قانون جامع حدنگار (کاداستر) کلیه متقاضیان تبدیل **اسناد** دفترچه ای به تک برگ در روستا ها، صرفا از پرداخت هزینه برگ سند معاف باشند.

ماده ۱۱ قانون جامع حدنگار (کاداستر) نیز می گوید:

ماده ۱۱. سازمان مکلف است ظرف مهلت پنج سال پس از لازم الاجراء شدن این قانون کلیه **اسناد** مالکیت دفترچه ای را به **اسناد** مالکیت حدنگار تبدیل کند. برای این منظور کلیه دارندگان **اسناد** مالکیت دفترچه ای اعم از دستگاه ها و اشخاص خصوصی مکلفند در زمان بندی مشخصی که در هر منطقه به وسیله سازمان اعلام می شود؛ سند خود را به سازمان ارائه کنند.

این مصوبه پیشنهاد حجت الاسلام علیرضا سلیمی عضو هیات رئیسه مجلس بود.

به گزارش خبرنگار تشکلهای دانشگاهی خبرگزاری فارس، همزمان با ۱۳ آبان و روز ملی استکبارستیزی از کتاب روایت تسخیر و سامانه ای که **اسناد** لانه جاسوسی سابق آمریکا در آن بارگذاری شده، رونمایی می شود.

در این کتاب و سامانه **اسناد** لانه جاسوسی سابق آمریکا منتشر شده است.

این مراسم رونمایی به همت سازمان بسیج دانشجویی و با حضور جمعی از دانشجویان و مسئولان در مجتمع ۱۳ آبان (لانه جاسوسی سابق آمریکا) برگزار خواهد شد.

انتهای پیام/

اولویت و ارجحیت به داک های با طول کمتر داده شده است. دلیل این امر، نرمال سازی امتیاز داک ها می باشد. چرا که هر چه تعداد ترم های استفاده شده در داک کمتر باشد، نسبت امتیاز کسب شده توسط کوئری به سائز داک بیشتر خواهد بود.

- معیار شباهت ژاکارد

```
query: اسناد
similarity function: Jaccard
Document ID is: 9893
Document Score is: 0.037037037037037035
با حکم رئیس‌جمهور، مختارپور رئیس سازمان اسناد و کتابخانه ملی شد
https://www.farsnews.ir/news/14000921000552/
Document ID is: 11688
Document Score is: 0.023809523809523808
رونمایی از کتاب «روایت تسخیر» و سامانه اسناد جاسوسی آمریکا به‌مناسبت ۱۳ آبان
https://www.farsnews.ir/news/14000810000216/
Document ID is: 12086
Document Score is: 0.02127659574468085
طرح نمایندگان برای عدم الزام ثبت نقل و انتقال خودرو در دفاتر اسناد رسمی
https://www.farsnews.ir/news/14000727000193/
Document ID is: 792
Document Score is: 0.0196078431372549
بارگذاری 40 سند مالی سرخایی ها در سایت کدال
https://www.farsnews.ir/news/14001214000395/
Document ID is: 7961
Document Score is: 0.017241379310344827
اسناد آغاز عملیات اجرایی ۴۸ پروژه بزرگ صنعتی در حضور رئیس‌جمهور امضا شد
https://www.farsnews.ir/news/14001117000359/
```

همانطور که مشاهده می شود سند 11688 در هر دو مشترک است.

یک پرسمان از عبارات ساده و متداول چند کلمه ای:

باشگاه های فوتبال آسیا:

• معیار شباهت کسینوسی

```
query: باشگاه های فوتبال آسیا
similarity function: Cosine
Document ID is: 5738
Document Score is: 0.3904276293336304
قائدی و مغانلو رقیب ستاره الهلال نشدند+عکس و فیلم
https://www.farsnews.ir/news/14001010000151/
Document ID is: 5291
Document Score is: 0.38745277197897504
با اعلام AFC: استقلال و پرسپولیس از لیگ قهرمانان حذف شدند
https://www.farsnews.ir/news/14001017000056/
Document ID is: 409
Document Score is: 0.3838895935074921
کمیته صدور مجوز حرفه ای فدراسیون ایران را تعلیق کرد AFC
https://www.farsnews.ir/news/14001218000584/AFC
Document ID is: 1604
Document Score is: 0.3806907704361262
تعطیلی لیگ عربستان در ماه رمضان/میزبان لیگ قهرمانان آسیا مشخص شد
https://www.farsnews.ir/news/14001203000465/
Document ID is: 6626
Document Score is: 0.37335750809057994
بازیکن سابق پرسپولیس نامزد بهترین مهاجم سال آسیا شد+عکس
https://www.farsnews.ir/news/14000928000213/
```

به گزارش خبرگزاری فارس، کنفدراسیون فوتبال آسیا به طور رسمی از حذف ۳ باشگاه استقلال، پرسپولیس و گل گهر از لیگ قهرمانان آسیای سال ۲۰۲۲ خبر داد.

کنفدراسیون فوتبال آسیا همچنین خبر داد که درخواست گل گهر رد شده است و این تیم نیز نمی‌تواند به عنوان نماینده کشورمان در فصل آینده حضور پیدا کند.

طبق ماده ۱۴.۴ و ۱۴.۵ اساسنامه مجوز حرفه‌ای، این ۳ باشگاه نتوانستند معیارهای ضروری حضور در لیگ قهرمانان آسیا را بدست آورند و پرونده آنها به کمیته انضباطی و اخلاق کنفدراسیون فوتبال آسیا ارجاع می‌شود.

در تاریخ اول دی ماه، کنفدراسیون فوتبال آسیا ۵ روز به ۳ باشگاه ایرانی مهلت داد تا مدارک خود را برای کسب مجوز حرفه‌ای ارائه دهند اما با وجود این اقدام مدارک جدید مورد تایید قرار نگرفت.

سایت فوتی‌رنکنینگ خبر حذف ۳ نماینده ایرانی را بازتاب داد و نوشت که کشورمان در لیگ قهرمانان آسیای سال آینده تنها ۲ تیم‌های فولاد و سپاهان به صورت مستقیم راهی مرحله گروهی لیگ قهرمانان آسیا خواهند شد.

• معیار شباهت ژاکارد

```
query: باشگاه های فوتبال آسیا
similarity function: Jaccard
Document ID is: 4988
Document Score is: 0.07894736842105263
موافق AFC با درخواست پرسپولیس
با-درخواست-پرسپولیس-AFC-موافق/https://www.farsnews.ir/news/14001020000675
Document ID is: 3764
Document Score is: 0.07692307692307693
بازدید ناظران AFC از استادیوم آزادی+عکس
از-استادیوم-آزادی-عکس-AFC-بازدید-ناظران/https://www.farsnews.ir/news/14001104000657
Document ID is: 5071
Document Score is: 0.061224489795918366
درخواست پرسپولیس از AFC برای ارائه دلایل حذف از لیگ قهرمانان آسیا
برای-ارائه-دلایل-حذف-از-لیگ-قهرمانان-آسیا-AFC-درخواست-پرسپولیس-از/https://www.farsnews.ir/news/14001019000762
Document ID is: 1788
Document Score is: 0.06060606060606061
حامد لک با پرسپولیس تمدید کرد
حامد-لک-با-پرسپولیس-تمدید-کرد/https://www.farsnews.ir/news/14001130001170
Document ID is: 2021
Document Score is: 0.06060606060606061
اعلام زمان نشست خبری گل‌محمدی پیش از بازی با آلمینیوم
اعلام-زمان-نشست-خبری-گل‌محمدی-پیش-از-بازی-با-آلمینیوم/https://www.farsnews.ir/news/14001127001339
```

در مقایسه با شباهت کسینوسی، مشاهده می‌شود که نتایج بیشتر در مورد ایران است و نه آسیا.

• روش مکانی فاز 1

```
Query: باشگاه های فوتبال آسیا
Terms: ['باشگاه', 'آسیا', 'فوتبال']
Document: 0 in Title of: اعلام زمان فرعه کشی جام باشگاه های فوتبال آسیا Relevance Score: 3
Document: 797 in Title of: پتانسیل تهران برای میزبانی جام باشگاه‌ها بیشتر است/مافظیه در حد لیگ یک و دو هم نیست Relevance Score: 3
Document: 1306 in Title of: تشریح جزئیات دریافت مجوز حرفه‌ای برای تیم‌های فوتبال و تشکیل کمیته ویژه در هفته‌های پایانی لیگ Relevance Score: 3
Document: 1535 in Title of: چرا تیم ملی فوتبال به تایلند اعزام نمی‌شود؟ Relevance Score: 3
Document: 1671 in Title of: اعلام زمان نشست خبری گل‌محمدی پیش از بازی با آلمینیوم Relevance Score: 3
```


یک پرسمان دشوار و کم تکرار تک کلمه ای:

تعلیق:

• شباهت کسینوسی

```
query: تعلیق
similarity function: Cosine
Document ID is: 1746
Document Score is: 2.127121398387756
تولید کوکائین در آپارتمان یک-المپیک
https://www.farsnews.ir/news/14001201000587/
Document ID is: 6163
Document Score is: 2.092029798475717
تعلیق دوباره در انتظار وزنه برداری مصر
https://www.farsnews.ir/news/14001004000488/
Document ID is: 513
Document Score is: 1.7551665629350772
اجازه فیفا به بازیکنان خارجی شاغل در روسیه و اوکراین
https://www.farsnews.ir/news/14001217000210/
Document ID is: 12164
Document Score is: 1.5712188646690877
مراقب تحریف های تاریخی روحانی و ظریف در برجام باشیم / سناریو جدیدی در خارج از ایران علیه جبهه انقلاب شکل گرفته
https://www.farsnews.ir/news/14000724000632/
Document ID is: 3064
Document Score is: 1.418080932258504
اعلام آرای کمیته اخلاق سرمربی شهرداری همدان محروم شد
https://www.farsnews.ir/news/14001113000359/
```

1746:

به گزارش خبرگزاری فارس، یک ورزشکار المپیک ایرانلی به دلیل اجازه استفاده از آپارتمانش برای تولید کوکائین به حبس **تعلیق** محکوم شد.

لئون رید ۲۷ ساله، روز جمعه در دادگاه بریستول کراون به ۲۱ ماه حبس **تعلیق** محکوم شد. البته این ورزشکار باتوجه به اسلحه گرم پیدا شده در منزلش گناهکار شناخته نشد.

این دهنده در بازی های کشورهای مشترک المنافع که سال ۲۰۱۸ به میزبانی استرالیا برگزار شد برنز گرفت. وی در المپیک توکیو در دوی ۲۰۰ متر راهی نیمه نهایی شد.

انتهای پیام/

6163:

به گزارش خبرگزاری فارس، مصر بعد از **تعلیق** بحث برانگیز ۲ ساله به مسابقات جهانی وزنه برداری بازگشت اما حالا به دلیل دوبینی که ۲ سال و نیم پیش مشخص شده با محرومیت دیگری روبه روست.

آژانس بین المللی تست (ITA) که کارهای اخذ نمونه های دوپینگ فدراسیون جهانی وزنه برداری را انجام می دهد، تاکید تست دوپینگ ۶ وزنه بردار مصری در جریان بازی های آفریقایی ۲۰۱۹ که در مراکش برگزار شد، مثبت شده است.

به محض اینکه این اتفاق بیفتد و تحریم ها اجرایی شود، فدراسیون وزنه برداری مصر بار دیگر در معرض **تعلیق** قرار می گیرد. وزنه برداری جزو رشته های موفق مصر محسوب می شود اما حالا با مشکل جدی به نام دوبینگ روبه روست.

در طول **تعلیق** دو ساله که اوایل ماه جاری به پایان رسید، یک تحقیق دولتی صورت گرفت، فعالیت فدراسیون ملی متوقف شد و دو مربی برای مدت کوتاهی زندانی شدند اما ورزشکاران در ازبکستان تمرین کردند و عملکرد خوبی داشتند.

هنوز اسامی ورزشکاران دوبینی به طور رسمی اعلام نشده است. براساس اعلام فدراسیون جهانی، کشورهایی که ۲ مورد یا بیشتر نمونه مثبت دوپینگ داشته باشند به مدت ۲ سال **تعلیق** می شوند.

```

query: تعلیق
similarity function: Jaccard
Document ID is: 895
Document Score is: 0.018867924528301886
محرومیت دو مدیر فوتبالی از سوی کمیته اخلاق
https://www.farsnews.ir/news/14001212000304/
Document ID is: 1746
Document Score is: 0.017857142857142856
تولید کوکائین در آپارتمان یک المپیک
https://www.farsnews.ir/news/14001201000587/
Document ID is: 6800
Document Score is: 0.016666666666666666
رای شکایت شیمیدر از ملیپوش بسکتبال مشخص شد
https://www.farsnews.ir/news/14000925000002/
Document ID is: 409
Document Score is: 0.015625
کمیته صدور مجوز حرفه ای فدراسیون ایران را تعلیق کرد AFC
https://www.farsnews.ir/news/14001218000584/AFC/
Document ID is: 2939
Document Score is: 0.015384615384615385
المپیک زمستانی پکن برگزار می‌شود - مراسم افتتاحیه و رژه اسکی‌بازان ایران
https://www.farsnews.ir/news/14001115000447/

```

895:

به گزارش خبرگزاری فارس و به نقل از سایت رسمی فدراسیون فوتبال، نصرالله خواجه گیری مدیر اسبق تیم بادران تهران به پرداخت دویست میلیون ریال جزای نقدی و محرومیت شش ماه از هرگونه فعالیت ورزشی در زمینه فوتبال، فوتسال و فوتبال ساحلی (تعلیق نصف محرومیت باقیمانده برای مدت یک سال) محکوم شد.

در دیگر رای صادره نیز محمدتقی مشایخی مالک اسبق تیم پیشگامان فنون پارس-لیگ دسته سوم کشور به پرداخت دویست میلیون ریال جزای نقدی و دوازده ماه محرومیت از هرگونه فعالیت ورزشی در زمینه فوتبال، فوتسال و فوتبال ساحلی محکوم شد.

6800:

به گزارش خبرنگار ورزشی خبرگزاری، کمیته انضباطی فدراسیون بسکتبال پس از تشکیل جلسه در خصوص شکایت باشگاه شیمیدر از محمد جمشیدی رای این شکایت را صادر کرد.

براساس این رای محمد جمشیدی به دلیل توهین به باشگاه شیمیدر و انتشار ویدیوی آن در فضای مجازی به مدت یک ماه محروم شد که این محرومیت تعلیقی خواهد بود.

پیش از این هم یک بار دیگر شیمیدر به دلیل زیر پا گذاشتن قرارداد از سوی جمشیدی از این بازیکن شکایت کرده بود که به نتیجه ای نرسید و کار را به مراجع قضایی کشاند.

شیمیدری‌ها پس از این اتفاق در مصاحبه هایی به نحوه رای دادن فدراسیون در خصوص جمشیدی اعتراض کرده بودند.

همچنین سند شماره 1746 مشترک است.

یک پرسمان دشوار و کم تکرار چند کلمه ای:

اهداف شوم:

- شباهت کسینوسی:

```
query: اهداف شوم
similarity function: Cosine
Document ID is: 10949
Document Score is: 0.9963171221997812
مسئول بنیاد بین‌المللی استخبارت‌شناسی منصوب شد
https://www.farsnews.ir/news/14000825000886/
Document ID is: 5958
Document Score is: 0.9701258941575961
رفت: امیدوارم بتوانم شهرداری آستارا را از این شرایط نجات دهم/قبل از پیوستن به این باشگاه جواب رد دادم
https://www.farsnews.ir/news/14001007000345/
Document ID is: 4149
Document Score is: 0.9535185633398179
از ملاقات با رونالدو تا حضور در جمع 10 دروازه بان برتر جهان؛ بانو، فوتبالیستی که در آرزوی لژیونر شدن هستند
https://www.farsnews.ir/news/14001030000161/
Document ID is: 6843
Document Score is: 0.9423066544432213
عده‌ای در حال برنامه‌ریزی برای گرفتن عزت‌ورزشکاران ایرانی هستند
https://www.farsnews.ir/news/14001224001230/
Document ID is: 10564
Document Score is: 0.9257759847870254
سختگوی هیأت رئیسه مجلس: خشکی رودخانه‌ها میراث شوم ژنرال‌های دولت سازندگی و اصلاحات است
https://www.farsnews.ir/news/14000907000086/
```

10949:

به گزارش خبرنگار تشکلهای دانشگاهی خبرگزاری فارس؛ طی حکمی از سوی رئیس سازمان بسیج دانشجویی محمدمهدی رمضان زاده به عنوان مسئول بنیاد بین‌المللی استخبارت‌شناسی این سازمان منصوب شد.

گفتنی است، ایجاد شناخت در باره ماهیت استخبار و ایجاد احساس تبری و نفرت نسبت به مظاهر استخبار، یادآوری جنایات استخبار جهانی برای جلوگیری از غفلت یا فراموشی آنها توسط مردم، ایجاد شناخت درباره ساز و کارها و روش‌های استخبار برای پیاده کردن نقشه‌های **شوم** و شیطانی، آگاه‌سازی نسبت به وظیفه و نقش هر فرد در نهضت مبارزه با استخبار و جریان سازی‌های داخلی و خارجی نسبت به استخبار و نهضت همگانی در مقابل آن از مهم‌ترین **اهداف** این بنیاد است.

رمضان‌زاده مسئولیت‌هایی همچون معاون روابط عمومی و فضای مجازی سازمان بسیج دانشجویی، معاون ارتباطات و فضای مجازی بسیج دانشجویی تهران بزرگ، دبیر شورای تبیین مواضع بسیج دانشجویی تهران بزرگ و مسئول بسیج دانشجویی دانشگاه شهیدبهبشتی را در کارنامه دارد.

4149:

فائزه کمالی دیگر میهمان شبکه جام جم درباره شرایطش اظهار داشت: متولد ۱۳۷۵ هستم و اصالتا شیرازی هستم، اما در کرج زندگی می‌کنم. در حال حاضر هم ترم آخر رشته تربیت بدنی در مقطع کارشناسی هستم و پست دروازه بانی را در فوتسال دارم. خانواده من ورزشی بودند و مادر من کوهنورد و خواهر دوقلوی من هندبالیست است. من در پنجم دبستان از طریق یکی از دوستان مادرم وارد فوتسال استان شدم. اول به صورت دانش آموزی و حالا چند سال است به صورت حرفه بازی می‌کنم.

کمالی توضیح داد: در حوزه دروازه بانی، به دلیل اینکه توپ زیادی سمت دروازه ارسال می‌شود هیچان زیادی در طول بازی برای من وجود دارد و به نظر من به همین دلیل جذابیت فوتسال بیشتر از فوتبال است.

وی اظهار داشت: **هدف** من رفتن به تیم ملی است و اینکه می‌خواهم جزء ۱۰ گلر برتر جهان **شوم** و همچنین می‌خواهم لژیونر **شوم**. یکی از مشکلات زنان فوتبالیست دستمزدهای کم و قرارداد‌های پایین در ایران است برای همین من می‌خواهم در کشور های عربی لژیونر **شوم**.

```

query: اهداف شوم
similarity function: Jaccard
Document ID is: 174
Document Score is: 0.022727272727272728
مربیان اهداف پروازی مشخص شدند
https://www.farsnews.ir/news/14001221001049/
Document ID is: 10949
Document Score is: 0.022222222222222223
مسئول بنیاد بین المللی استخبارشناسی منصوب شد
https://www.farsnews.ir/news/14000825000886/
Document ID is: 9600
Document Score is: 0.02127659574468085
بمباران اهداف زمینی توسط جنگنده های سوخو 22 هو افضای سپاه
https://www.farsnews.ir/news/14000930000213/
Document ID is: 9603
Document Score is: 0.02040816326530612
شلیک و اصابت همزمان 5 فروند موشک کروز در رزمایش پیامبر اعظم سپاه
https://www.farsnews.ir/news/14000930000200/
Document ID is: 8439
Document Score is: 0.0196078431372549
وزیر دفاع جمهوری آذربایجان فردا به تهران سفر میکند
https://www.farsnews.ir/news/14001104000868/

```

174:

به گزارش خبرگزاری فارس، از سوی مسئولان فدراسیون تیراندازی مربیان **اهداف** پروازی زنان و مردان معرفی شدند. براین اساس، پیام رجایی به عنوان مربی تراپ بانوان، علی افسری تراپ آقایان، هنریک آواکیان مربی اسکیت و همچنین حسین ضیایی به عنوان مدیر فنی تیم های ملی **اهداف** پروازی منصوب شدند. به گزارش فارس، پیش از این مسئولان فدراسیون مربیان تپانچه را معرفی کرده بودند. هنوز تغییراتی در کادر فنی تفنگ ایجاد نکرده اند.

همچنین سند 10949 در دو روش مشترک است.

به طور کلی در هنگام سرچ کلمات پیچیده با روش شباهت کسینوسی به دلیل خاص بودن یک کلمه، اگر یک داک آن کلمه را داشته باشد از امتیاز بالایی برخوردار می باشد. اما از طرفی گاهی اوقات نتایج خروجی مرتبط نیستند و برای مثال در اینجا که کوئری جنبه سیاسی دارد، گاها خروجی مربوط به ورزشی نشان داده شده است.

• روش مکانی فاز 1:

```
Query: اهداف شوم
Terms: ['اهداف', 'شوم']
Forbidden Terms: []
phase words: {}

Document: 8989 in Title of: ماجرای فرمانده لشکر صدام که در رکاب شهید سلیمانی به شهادت رسید / علت کمرنگ شدن رابطه حاج قاسم با هاشمی
Relevance Score: (2, 7)
https://www.farsnews.ir/news/14001017000334/

Document: 10745 in Title of: سخنان تنه‌گین در دیدار با رهبر انقلاب / سال ۴۸ حضرت آیت الله خامنه‌ای به شهید ستاری چه نوشتند؟
Relevance Score: (2, 6)
https://www.farsnews.ir/news/140009020000005/

Document: 2831 in Title of: افشای دلایل جدایی یزوقسور از هندیال ایران / فرمانداز: نمی‌خواهم وارد جنگ شوم / با همسر قدامت‌یون و برخی تصمیم‌ها موافق نبودم
Relevance Score: (2, 5)
https://www.farsnews.ir/news/1400116000751/

Document: 7622 in Title of: آمریکا درآمد کشورها را محدود می‌کند و ژنت انسان دوستی می‌گیرد
Relevance Score: (2, 5)
https://www.farsnews.ir/news/1400121000750/

Document: 6843 in Title of: عده‌ای در حال برنامه‌ریزی برای گرفتن عزت ورزشکاران ایرانی هستند
Relevance Score: (2, 4)
https://www.farsnews.ir/news/14001224001230/
```

همانطور که مشاهده می شود، ارتباط محتوایی داک سوم به موضوع کمتر می باشد. همچنین در این حالت، برخلاف مدل مبتنی بر tfidf لزوما داک های با ساینز کوچک تر ارجح شمرده نمی شود و داک های طولانی تر نیز به عنوان پاسخ بازگردانده می شود. همچنین در برخی کلمات پیچیده تر مشاهده شد که این روش توانایی پاسخ به آن کوئری ها را دارا نبود و به همین دلیل در خروجی هیچ داکي برگردانده نمی شد.