محمد مهدی نظری – ۹۹۳۱۰۶۱

* برای این پروژه از فایل json اصلی استفاده نشده بدلیل بزرگ و وقت گیر بودن پراکسس فایل و در عوض ۲۰۰ خبر از جاهای مختلف آن در فایل data2.json قرار داده شده و شماره از ۲ تا ۲۰۰ به آن هاداده شده تا اجرای پروژه سریعتر باشد .

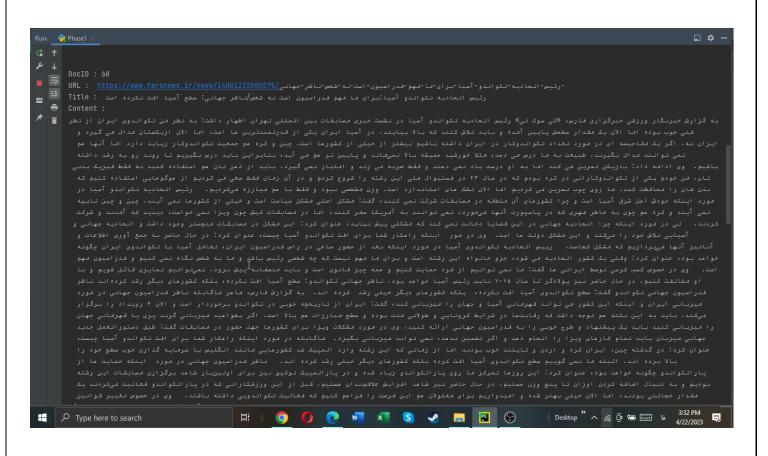
* برای نمایش هر سند مرتبط با کویری ۲ لیست ساخته شده است : اولی docs که همه ی شرط های کویری اعم از NOT یا عبارت در "" در نظر گرفته و شماره سند ها را باهم AND میکند . اگر تعداد داکیومنت های این لیست حداقل ۵ باشد . ۵ تای آن که بیشترین تکرار کلمات را دارند به ترتیب چاپ میکند(sorteddocsand)

و اگر تعدادآن از α کمتر باشد به ازای تعدادی که کم دارد از لیست docsor که صرفا داکیومنت هایی را برمیگرداند که کلمات کویری در آن α 0 شده باشند (چاپ آن هم به نسبت تعداد تکرار توکن در سند است با کمک لیست sorteddocsor)

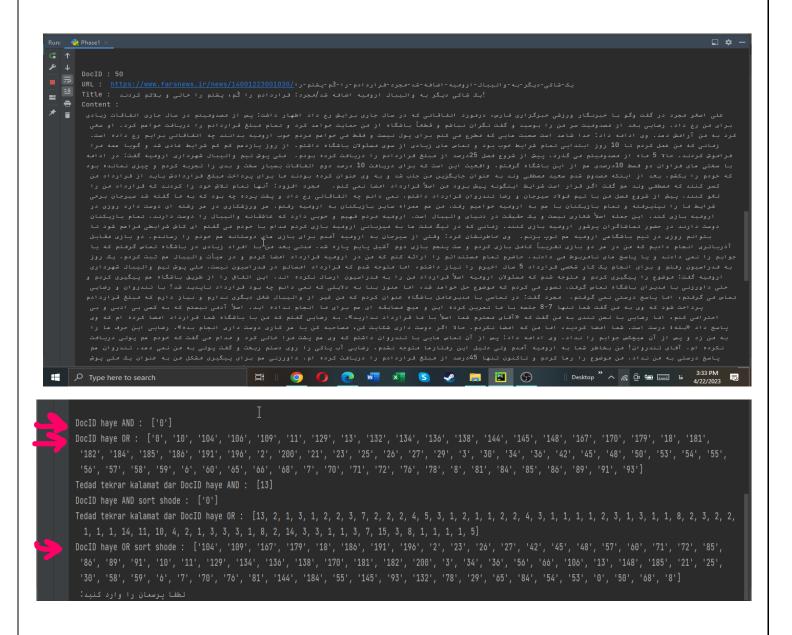
مثال ۱)



پاپان این رقابت انتخاباتی پیمان رضایی یا کسب 20 رای به عنوان رئیس میات والیبال ارومیه انتخاب شد. معجنین رحمانی 9 رای آورد و اعضای حاضر به احمدی رای ندادند. انتخاب پیمان رضایی به عنوان رئیس جمدید حیات والیبال ارومیه در حالی است که وی در صدت حضورش در سعت صدیر عامل <mark>باشگاه</mark> شهرداری ارومیه نتوانست به توفیق خاصی دست پیدا کند و بهترین هقام اروهیه نایب قهرمانی در لیگ برتر است. علاوه بر این موضوعات جای سوال دارد که چرا در انتخابات ریاست میأت والیبال شهری مانند اروهیه خبری بالا است. من در جایگامی نیستم که بخوامم درمورد مردم و والیبال ارومیه حرف بزنم، اما استال این آقا اینقدر <u>به مردم دروغ گف</u>تهاند و خرابکاری کردهاند، عملاً لطعهماى بزرگى به واليبال اروميه زدهاند وگرنه آن شهر بايد تاك<mark>نون</mark> <mark>ق</mark>هرمانی لیگ برتر را تجربه میکرد و به جام <mark>باشگاهمای ^{آسیا} و جهان میرفت، نه اینکه میزبانی</mark> . جام باشگاهمای آسیا را مم از آن شهر بگیرند، برخیما از آن باش 🛶 من گفتند که ما منوز بدمی تراویسای پدر و پسر و بازیکن فنلاندی سابق خود را مع ندادهایم؛ قرارد<mark>اد تو مم نهایتاً به آنها اضافه خوامد شد. درمعین راستا چندی پیش مم فدراسیون جهانی و کنفدراسیون والیبال <mark>آسیا ط</mark>یزبانی هسایقات جام <mark>باشگاههای آسیا</mark></mark> 2022 را به ایران امدا کردند و فدراسیون کشورهان مم تصمیم گرفت اروسیه را به عنوان هیزبان این رقابتما انتخاب کند، اسا هوضوع سهمی که وجود داد شهرداری اروسیه به عنوان میزبان این بازیما حق حضور نداشت، چراکه این <mark>باشگا:</mark> بعد از سالما بدمی خود را تراویسا را تسویه نکرده بود. گفته میشد <mark>باشگا: شهر</mark>داری ارومیه مبلغی حدود 15 میلیارد تومان را باید به تراویسای پدر و پسر پرداخت کند تا شاید محرومیتش بخشیده شود؛ اما میچ عزمی برای پرداخت این بدمی از سوی مسئولان اروهیه دیده نشد تا اینکه هیزیانی از اروهیه گرفته شد و به تهران رصید. در همان فقطع پیمان رضایی، هدیرعامل <mark>باشگاه شه</mark>رداری اروهیه پس از پیگیریمای *ه*کرر خبرتگار ورزشی خبرگزاری فارس برای اطلاع از آخرین شرایط این پرونده، اظهار داشت: فعلاً جلسه دارم و مر زنمانی که خودم صلاح ابدانم با شعا تعاس خواهم گوفت! در آن زفان اداران والبيال در اروسيه حاضر بودند دست به دست مم داده و از جيب خودشان مم كه شده اين بدمي را جعع آوري و پرداخت كنند تا شامد درخشش تيمشان در سسابقات جام اشگاههای آسیا باشند، اها رضایی و همکارانش ترجیح هیدادند چلسات بدون خروجی برگزار کرده و از پاسخگویی به اصحاب رسانه طفره بروند. به این ترتیب پیمان رضایی حالا قرار است در میأت والیبال فعالیت خود را دنبال کند. این اتفاق در شرایطی در حال رخ دادن است که تیعداری شهرداری اروسیه در فصل آینده لیگ برتر والیبال مم در پاشند. با این شرایط باید هنتظر ساند و دید، آیا رضایی که در دوران هدیریتش لطعه مای زیادی به والیبال اروسیه وارد شد، سیتواند تغییر نقش داده و ناجمی باشد یا

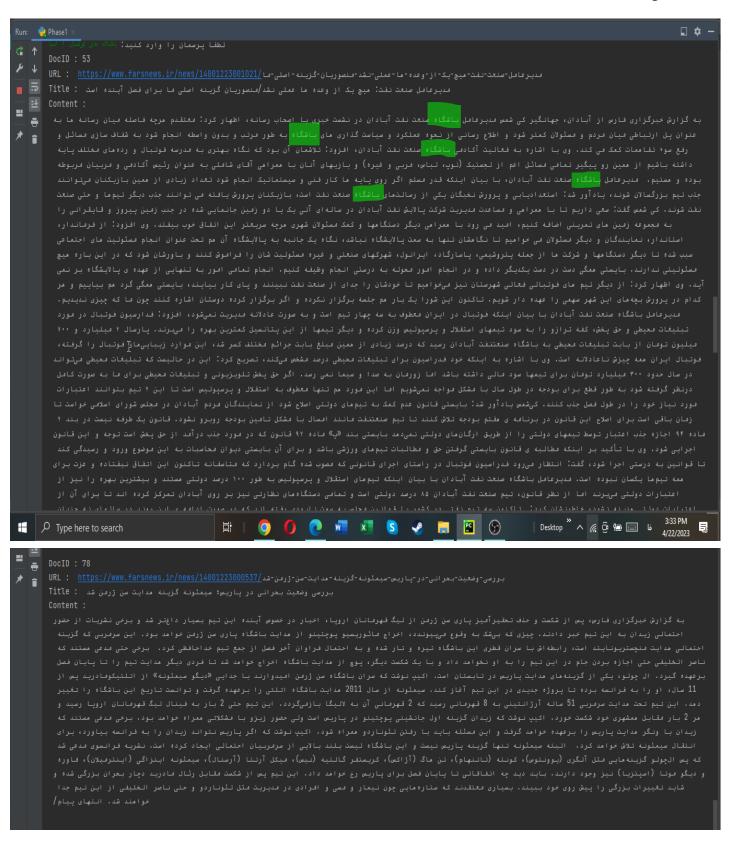


≝ ∄ ∤ i



ملاحظه میکنیم تنها سندی که همه ی کلمات کویری را همزمان داشت سند \cdot است . برای لیست \cdot مشاهده میکنیم که سندهای λ و λ و λ و λ تا با بیشترین تکرار کلمات هستند (لیست ها صعودی مرتب شده اند بر اساس تکرار کلمات) سند λ دوبار تکرار شده چونکه هم در لیست λ هم در لیست λ تعداد تکرارش بالاست. تعدا تکرار هر سند در لیست λ یا λ مشخص است مثلا در سند λ سیزده بار تکرار داشتیم یا در سند λ پانزده تکرار داشتیم.

مثال ۲)



DocID : 132

دليل-عدم-انتقال-آزمون-در-زمستان-به-لوركوزن-فاش-شه./https://www.farsnews.ir/news/14001106000386

دلیل عدم انتقال آزمون در زمستان به لورکوزن فاش شد : Title

Content :

به گزارش خبرگزاری فارس، سردار آزمون مهاجم ملی پوش کشورمان مفته جاری اردوی دبی را ترک کرد و رامی آلمان شد و با باشگاه باپرلورکوژن فراردادی 5 ساله اعضا کرد تا زفعل جدید در پوندسلیگا توپ پزند، نفریه معتبر «کیکر» آلمان دست به افشاگری زد و دلیل منتقل نشدن آزمون در زمستان و شرایط کنونی خبر داد و نوشت: مدیران باشگاه لورزکون با مدیران باشگاه زنیت وارد مذاکره شدند و میخواستند سردار آزمون را اکنون در نقل و انتقالات زمستانی جذب کنند. اما سران باشگاه زنیت خوامان مبلغ 10 میلیون پورو در ازای 6 ماه قرارداد باقی ماننده آزمون از باشگاه لورکوژن شدند اما مدیران باشگاه آلمانی این درخواست روس ما را رد کردند. آزمون که می خواست به لورکوژن برود حاضر شد قراردادش را اکنون امضا کند ولی در تابستان رامی پوندسلیگا شود. ستاره ایرانی در تست جدیدی که داده کرونایش منفی شده و اکنون باشگاه زنیت الورکوژن برود حاضر شد قراردادش را اکنون امضا کند ولی در تابستان رامی پوندسلیگا شود. ستاره ایرانی در تست جدیدی که داده کرونایش منفی شده و اکنون باشگاه زنیت

DocID: 145

URL : https://www.farsnews.ir/news/14001106000309/ أخر-زد والمحال المحالمة الم

کاپیتان پرسپولیس تنها بازیکن فایب در محل تعرین/ سید جلال به <u>سیم آخر زد</u>

Content

به گزارش خبرنگار ورزشی خبرگزاری فارس، بازیکنان تیم پرسپولیس اسروز طبق برنامه در معل تعرین حاضر شدند اسا آنها معان طور که پیش بینی شده بود حاضر به برگزاری تعرین نشدند. در حالی که اعضای کادر فنی معنی در ورزشگاه حاضر شده و حتی بازیکنان مم برای تعرین آماده بودند و حتی برخی از آنها تغیس فوتبال مم انجام دادند اسا به یکباره نظرشان عوض شد و به یحیی گل محمدی اعلام کردند به دلیل اینکه باشگاه زیر قولش زده و پولی به آنها نداده حاضر نیستند تعرین کنند. نکته قابل تامل اینجا است که از بین بازیکنان تنها سبد جلال حسینی کاپینان سرخپوشان در تعرین حاضر نشد و گفته هی شود حتی تعاس او با برخی بازیکنان باعث شد تا بقیه مم تصمیم بگیرند امروز تعرین نکنند. حسینی پیش از این مم نسبت به شرایط باشگاه در میکسدزون ورزشگاه آزادی گلابه کرده بود. باشگاه پرسپولیس از ابتدای فصل جاری فقط 20 درصد قرارداد بازیکنان را پرداخت کرده و در این مدت میچ پاداشی مم به تیم تعلق نگرفته است. با توجه به بدقولی مای جعفر سمیعی و مجید صدری دو مدیرعامل و سرپرست پیشین باشگاه بازیکنان دیگر نسبت به وعده مای رضا درویش مدیرعامل فعلی اعتمادی ندارند. مدیرعامل پرسپولیس به اعضای تیم اعلام کرده پیگیر پرداخت شدن مطالبات است اسا اینکه این بازیکنان دیگر نسبت به وعده مای رضا درویش مدیرعامل فعلی اعتمادی ندارند. مدیرعامل پرسپولیس مم کسی پاسخگوی شرح این شرایط نبست. انتهای پیام /

Ĭ

DocID: 144

URL : https://www.farsnews.ir/news/14001106000302/كرد/https://www.farsnews.ir/news/14001106000302/كلگهر-با-درخو است-پرسپولیس-مخالفت-کرد

گلگهر با درخواست پرسپولیس مخالفت کرد : Title

Content :

به گزارش خبرنگار ورزهی خبرگزاری فارس، با شروع نقل و انتقالات زمستانی فوتبال، باشگاه های هختلف فعالیت خودشان را برای تقویت هرچه بهتر جهت حضور در ادامه بازیهای نیم فصل دوم آفاز کردند. در این بین باشگاه پرسپولیس درخواست خودش را برای جذب مهدی تیکدری سدافع و سعید صادفی هافیک تیم گل گهر، ارائه کرد. با اینحال باشگاه گل گهر موافق انتقال این دو بازیکن نیست با این درخواست مخالفت کرد چراکه با توجه به اهمیت بازیها در نیم فصل دوم سدیران باشگاه قصد دارند با حفظ مهردهای فعلی، برای حضوری بهتر و قویتر تیم را در مناطق هورد نظر کادرفنی تقویت کند. انتهای پیام /

DocID haye AND: ['10', '109', '132', '134', '136', '138', '144', '145', '148', '167', '170', '18', '185', '186', '196', '2', '200', '21', '23', '25', '42', '48', '53', '57', '58', '59', '6', '66', '7', '72', '78', '81', '85', '89']

DociD haye OR: ['0', '10', '104', '100', '109', '129', '13', '132', '134', '136', '138', '144', '145', '148', '167', '170', '18', '181', '182', '181', '185', '186', '191', '190', '2', '200', '21', '25', '26', '27', '29', '3', '30', '34', '36', '42', '48', '50', '53', '54', '55', '56', '57', '58', '59', '60', '65', '66', '68', '7', '70', '71', '72', '76', '78', '81', '84', '85', '86', '89', '91', '93']

Tedad tekrar kalamat dar DocID haye AND: [2, 1, 7, 2, 2, 2, 4, 5, 3, 1, 2, 1, 3, 1, 1, 1, 2, 3, 1, 3, 1, 1, 11, 1, 3, 3, 3, 2, 3, 1, 3, 7, 3, 1, 1]

DocID haye AND sort shode: ['109', '167', '18', '186', '196', '2', '23', '42', '48', '57', '72', '85', '89', '10', '134', '136', '138', '170', '200', '66', '148', '185', '21', '25', '58', '59', '6', '7', '76', '81', '144', '145', '132', '78', '53']

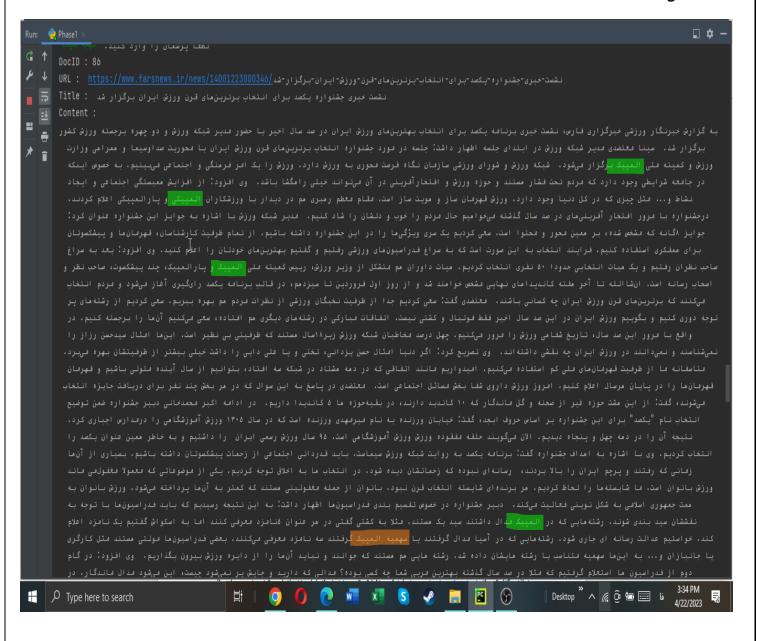
Tedad tekrar kalamat dar DocID haye OR : []

DocID haye OR sort shode : []

ملاحظه میکنیم که لیست AND بیشتر از ۵ تاست و در موارد نشان داده شده خبری از کلمه آسیا نیست.

در دوخط آخر میبینیم که دیگر لیست ORرا سورت نکرده دلیل آن حفط حافطه است تا زمانی که تعداد لیست ANDکم نباشد سراغ لیست OR نمیرویم.

مثال ۳)



DocID : 13

URL : https://www.farsnews.ir/news/14001224000731/پولادگر-درخششمای-مندبال-صابه-آبرو-و-اعتبار-بینالعللی-است-پاکدل زودتر از برناضه به امداف خود رسیدیم : Title :
پولادگر: درخششمای مندبال مایه آبرو و اعتبار بینالعللی است/ پاکدل: زودتر از برناضه به امداف خود رسیدیم : Content

به گزارش جبرنگار ورزشی خبرگزاری فارس، مجمع فرق العاده فدراسیون مغنبال به ریاست دکتر صید محمد پولادگر معاون وزیر ورزش و جراتان و با حضور دیگر اعشا امروز استنده کرد مالی استاد مارس آکادس میلی استیاد کمیته علی العبید نم مورد خبر و آمایش برگزار شد. در این مجمع که با حضور مید محمد پولادگر به عنوان نماینده وزارش و جوانان، فضالله بافترزاده به عنوان نماینده کمیته علی العبید، نمسرالله پردیچهره عنون فادر استون فلور و نواد ورزش و دیگر اعتان مجمع براه امروز در سرد و دیگر اعتان مجمع قدار گوشت. در این مجمع اساس نامه فدرامیون مای ورزشی، دستورالعمل کمیمیون مای ورزشکاران، دربیان و دیگر اعتان مجمع شاس نامه فدرامیون مای ورزشی، دستورالعمل کمیمیون مای ورزشکاران، دربیان و دیگر اعتان مجمع قدار گرفت. پاکدل: وودتر از برنامه به اعداله خود رسیدیم علیرها پاکدل رئیس فدرامیون مندیال در این مجمع با اشاره آب الثانات الله المورد مورد برزسی و معنوا بالا بالا کشورت آزا مورد ماله بازی مندین الله المهار کرد: در طول سالی که کشفت شامه رخ دادن اثناتات میان و استان این مجمع با اشاره آب به اعداله مورد را در خود ماله بازی در خود سالی که کشفت شامه نو دادن در الفاقات بصیار مهمی در مندیال ایران بودیام بازی بازی در مول مالی که کشفت شامه نو دادن دختر ایران توانستند تاریخ ماین کشور بازی میان بازی در در خود است بازی میان بازی در مول میان است که بهمید میان و به بازی و در نوان تاید به جامعه مندیال برزی در مای کشور بازی در در خود ایدان بازی بازی در در خود ایدان بازی در در خود بازی در در خود ایدان بازی در در خود بازی در در خود بازی در در خود بازی در در خود ایدان بازی در خود ایدان بازی در در خود ایدان بازی در در در خود بازی در در در میان بازی در در در خود در در بخش مایی دادن بازی در در در خود بازی نوان ازی در خود بازی در خود بازی میان بازی در در در در در در در در خود میان بین العملی در در در خود میان بین العملی ایدان در در در خود بازی در خود در این خوان و بازی در خود در در در خود در در در میان بین العملی است میدی بازی میدی در در خود در در خود در در در خود در در در میان در میان بین العملی در در خود در در خود در در در در

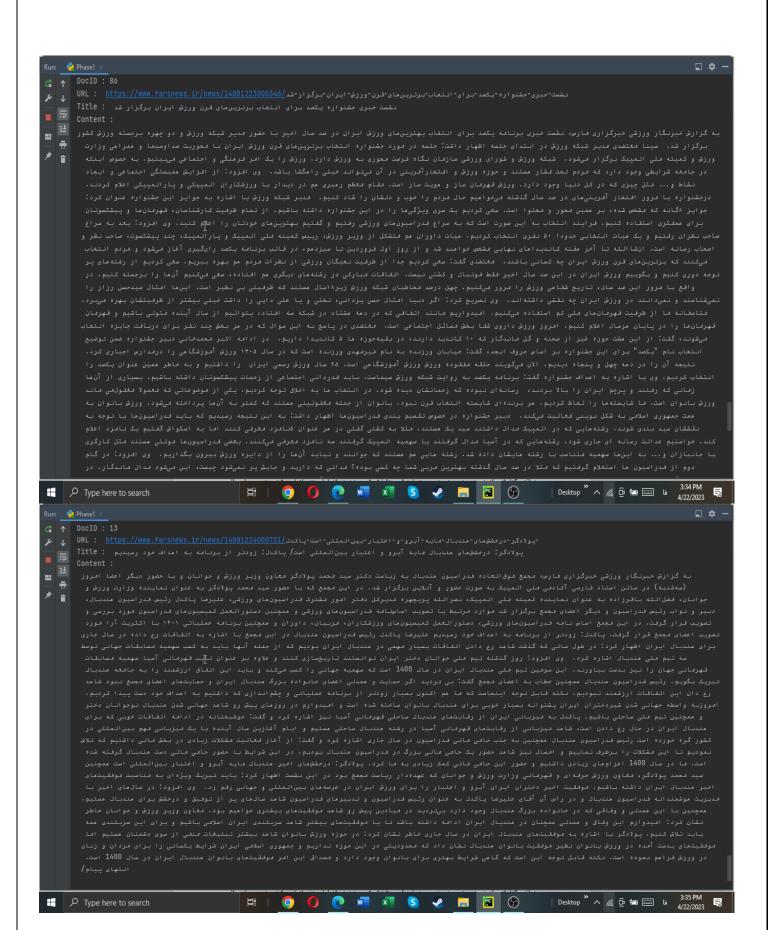
DocID : 3

ماجدى-در-نشست-با-صالحي-اميري-اميدوارم-در-اين-دوره-نيم-العبيك-موفق/https://www.farsnews.ir/news/14001224000964

ماجدی در نشست با صالحی امیری: امیدوارم در این دوره تیم العبیک موفق شود : Title

Content :

به گزارش خبرگزاری فارس، سید رضا صالحی امیری رئیس کعیته ملی العپیک در دیدار با سید میرشاد ماجدی سرپرست فدراسیون فوتبال بر آمادگی این کعیته به منظور حمایت معه جانبه از نیم العپیک ایران و ضرورت پیش بینی برنامه های آماده سازی مناصب برای این تیم تاکید کرد و گفت: ضمن آرزوی موفقیت برای آفای ماجدی به عنوان چهره با اخلاق فوتبال و ورزش تعام تلاشمان این خوامد بود تا تیم امید و العپیک در معیطی آرام و بدون دفدفه آماده حضور در بازی های آمیایی شود، ضمن اینکه پیش از این مم اعلام کردیم دخالتی در امور فنی تیم ما نداشته و نخوامیم داشت. سرپرست فدراسیون فوتبال نیز با تاکید بر ضرورت تعامل و معکاری معه جانبه در خصوص برنامه ریزی و آماده سازی مناسب تیم العبیک گفت:تیم فوتبال زیر 25 سال ایران با حضور مهدی مهدوی کیا و اعضای کادر فنی معچنین از ظرفیت بسیار خوبی برای موفقیت در این دوره برخوردار است و معامنگی و معکاری نزدیک کعیته ملی العپیک و فدراسیون فوتبال در راستای حمایت از این تیم با تامین خواسته مای مورد نظر اعضای کادر فنی می تواند نشیدی به معراه داشته باشد. امیدواریم در این دوره شامد موفقیت تیم العپیک و کسب نتایجی در شان و جایگاه فوتبال ملی ایران باشیم. در این نشست که با حضور نصرالله سجادی مشاور رئیس کمیته و پیمان فخری سرپرست کاروان ورزشی اعزامی ایران به بازی مای آمیایی ۲۰۲۲ مانگزو برگزار شد معامنگی لازم در خصوص معکاری دو جانبه نصرالله سجادی مشاور رئیس کمیته و پیمان فوتبال و کمیته ملی العپیک در خصوص برنامه ما و اردومای آماده سازی آنی تیم العپیک صورت پذیرفت. انتهای پیام /



```
DocID haye AND : ['13', '86']

DocID haye OR : ['100', '106', '114', '13', '171', '181', '184', '191', '29', '3', '36', '50', '56', '62', '65', '68', '70', '76', '85', '86']

Tedad tekrar kalamat dar DocID haye AND : [5, 7]

DocID haye AND sort shode : ['13', '86']

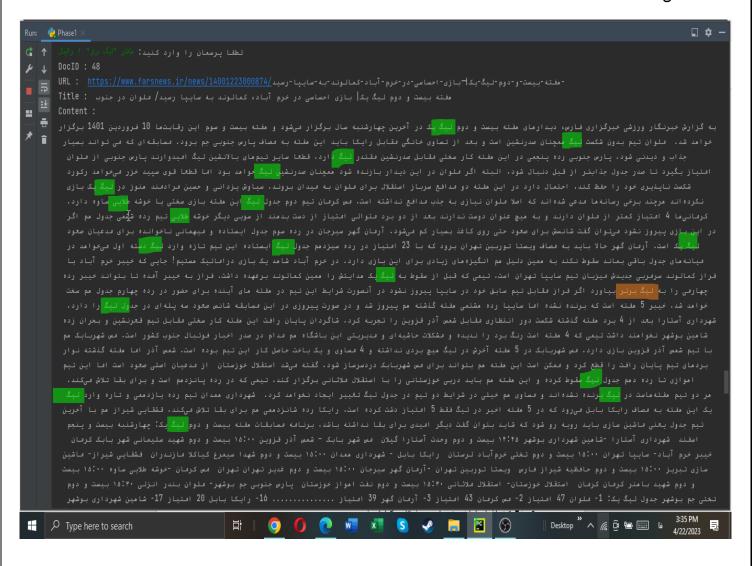
Tedad tekrar kalamat dar DocID haye OR : [3, 4, 2, 5, 2, 1, 3, 1, 2, 8, 3, 1, 2, 1, 4, 1, 1, 1, 7]

DocID haye OR sort shode : ['181', '191', '50', '62', '68', '70', '76', '85', '114', '171', '29', '56', '100', '184', '36', '106', '65', '13', '86', '3']

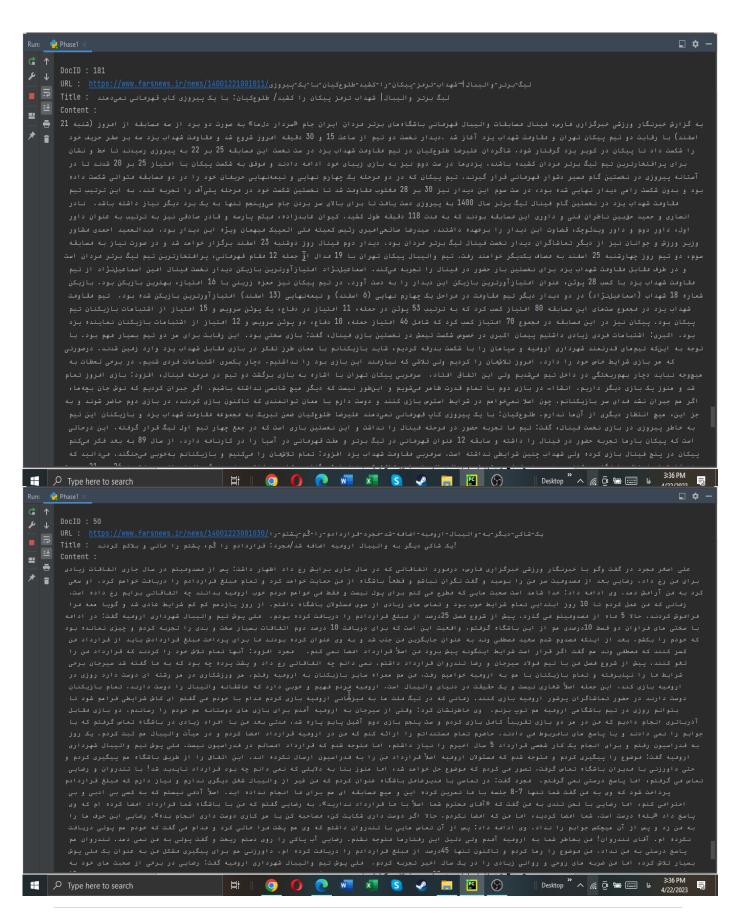
: كففا برسمان را وارد كنبد:
```

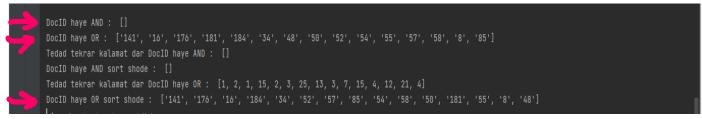
ملاحظه میکنیم فقط در دو لیست عبارت "سهمیه المپیک" وجود دارد (هایلایت نارنجی)

مثال ۴)



یه گزارش جبرنگار ورزقی جبرگزاری فارس، تیم مقاومت شهداب پزد در دور برگفت فینال مسابقات والبیال قهرمانی باشگاه مای برتر مردان ایران جام «سردار دلاما» با تنییه سه بر یک مقابل پیکان تهران پیان تهران برای و افتاد به چبروزی رسید و در تحصینی قهرمانی کور و را لیگ برتر جنن گرفت. علیرها طلوعکیان در ارزیابی عملکرد تیمش در بازی دوره غیال، گفت: در درجه اول باید حدا را شکر کنیم، برای تحصیه به ایرا دست که چنین اتفاقی پرای والبیال کشور و داده و به قصعه به خاطر تداوم برای تحصیه با پرد برای اولین پار چنین اتفاقی را رفم زد و یا خداد به به به خاطر تداوم برای برای برای والبین پار چنین اتفاقی را رفم زد و یا دارد به مجموعه قهداب برکیم. شیرینتر است نوخ جانهان، بچه با بسیار زحمت گفیدت و من کمترین مهم را دافتم، صربی مقاوت شهداب پرد برای اولین پار چنین اتفاقی با دده. پیکان تیمی با کادولنی بسیار خوب و تجربه کانی برخوردار بود و با بیش از 10 ارزم متفاوت فقابل شهداب بازی کرد. اما برای مر ارنج این تیم برنامه داشتیم و مقابلت به مثل کردیم، حتی تعام اتفاقات را با بازیکنانهان پیفربیای میکردیم، موفقیت تیم ما به دلیل معدلی و انکا بر تاکیکما بود، اگر مربی ایزار لازم را در اختیار نداشتهانات با بازیکنانها را برای موفقیت تیم ما به دلیل معدلی و انکا بین تعریع کرد؛ باید از مین لحمه که استان پرد برای تحصین بار صاحب طلال لیگ برتر والبیال شده، باید به دنبال حفظ این قهرمانی باشد. به دست آوردن قهرمانی یک بخش کار است ولی تکهداری و خطع آلهرمانی پرد برای تحصین بار صاحب مهرم این باید برای تحصین بار صاحب می موفقی بازیکنان مین تعمر این بازیکنان باز برد موجه مصند که به اتفاقی برد برای تحصین بار می مینی خواده است. طبوع به مواند مین در باری تحصین بار ساحب که برخوان داده مید خود تعلی برگزی ایا بازیکنان و تیم معیر برد مید بازی تحوید و شعوت بازی است. بازیکنان بازیکنان خوب به این دادی بازیکنان خوب به مواند بازیکنان دایع که مربولی جوان داده بید بازیکنان خوب بازیکنان و تیم که بازیکنان دایع که میرون در ویان برخود و فایستان دایع که روزد داری که بیش داری بینان میازی بازیکنان بازیکنان در بین ازیکنان دایع که برخواد بازیکنان بازیکنان خوب که در بازیکنان خوب که انتان دایع کوبی براید بین در بودی با داخید بازیک بازیکنان بازیکنان بازیکنان داید بازی در مواند کر در با نمان دادند و منوا





مشاهده میکنیم که هیچ سندی همه ی شرایط خواسته شده کویری را ندارد پس از لیست OR برای نمایش جواب استفاده میکنیم.

مثال ۵)



مشاهده میکنیم که برای این کویری نادر لیست ORهم کافی نبوده و فقط ۲ سند پیدا کرده است و به ارور index out of range بر خوردیم (در چاپ سند ها فرض بر این بوده که ۵ سند چاپ میشود در اینجا چون مجموع و لیست کمتر از ۵ بوده پس به ارور خورده)

چونکه عبارت جردن هم در سند موجود است پس لیست AND خالی بوده و سند ۸۵ که پرینت شده تنها عضو لیست OR بوده است.

بطور کلی برای همه کویری ها اگر لیست AND با تعداد تکرار بالاتری موجود باشد کویری به سند مرتبط تر است اما هر چه به سمت لیست OR برویم ارتباط کمتر میشود.

برای مثال اول اولین سند و کویری مرتبط اند چونکه هم در لیست AND هست و هم تعداد تکرار الایی دارد (۱۳).

در مثال دوم ار تباط کمی دارند علارغم اینکه لیست AND لیست طویلی است اما تعداد تکرار ها پایین است و سند بازگرداندنه شده بیشتر در موضوع فوتبال است تا فوتسال. در مثال سوم ارتباط خوبی داریم به همان دلایل گفته شده (لیست AND و تکرار به نسبت خوب)

در مثال ۴ ارتباط خوبی نداریم چونکه هیچ سندی همه ی شرایط کویری را رعایت نمیکند و سراغ لیست OR رفتیم اما بد هم نیست (از جمله موارد ناخوشایند باز گرداندن خوشه طلایی برا عبارت طلای است که میدانیم منظور مدال لیگ برتر است نه نام تیم)

در مثال پنجم هم بدلایل گفته شده هم تعداد سند کم است هم به دلیل خلاصه بودن کویری صرفا متن با اسم مایکل جودن ملاک خوبی نیست بلکه باید بیشتر در حوزه بسکتبال باشد (سند بازگردانده شده در مورد مقایسه در آمد ورزشکاران است)