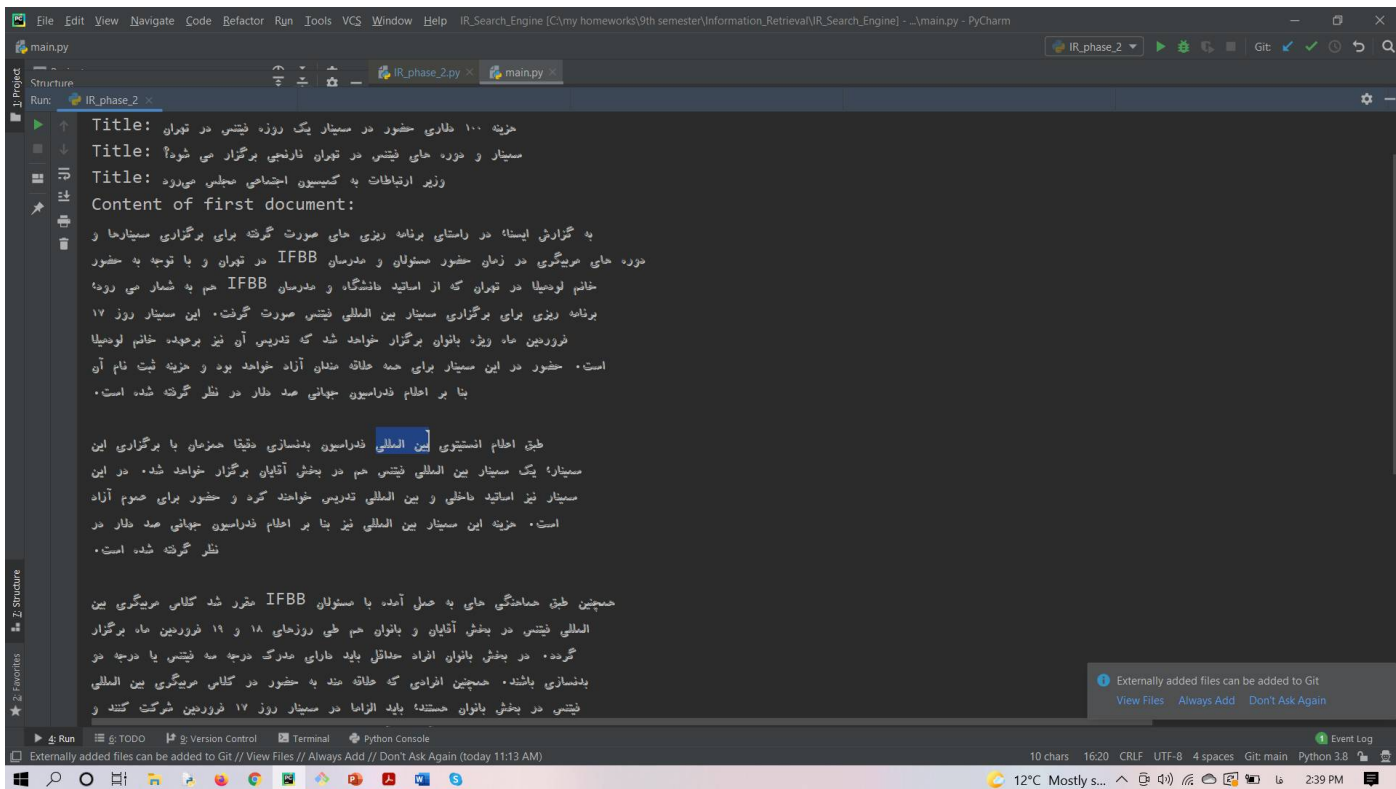


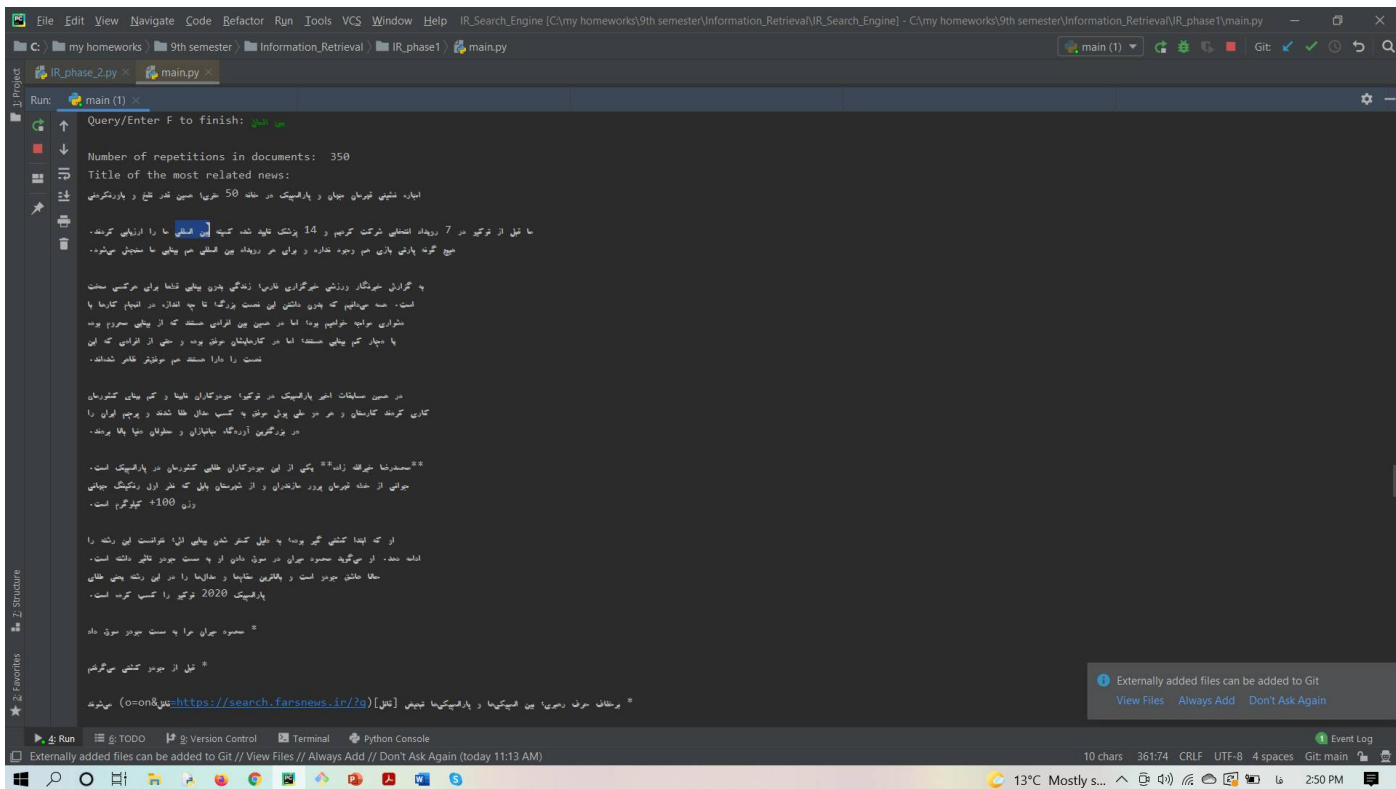
۱. پاسخ به پرسمان در حالت های زیر:

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

نتیجه ی فاز دوم _

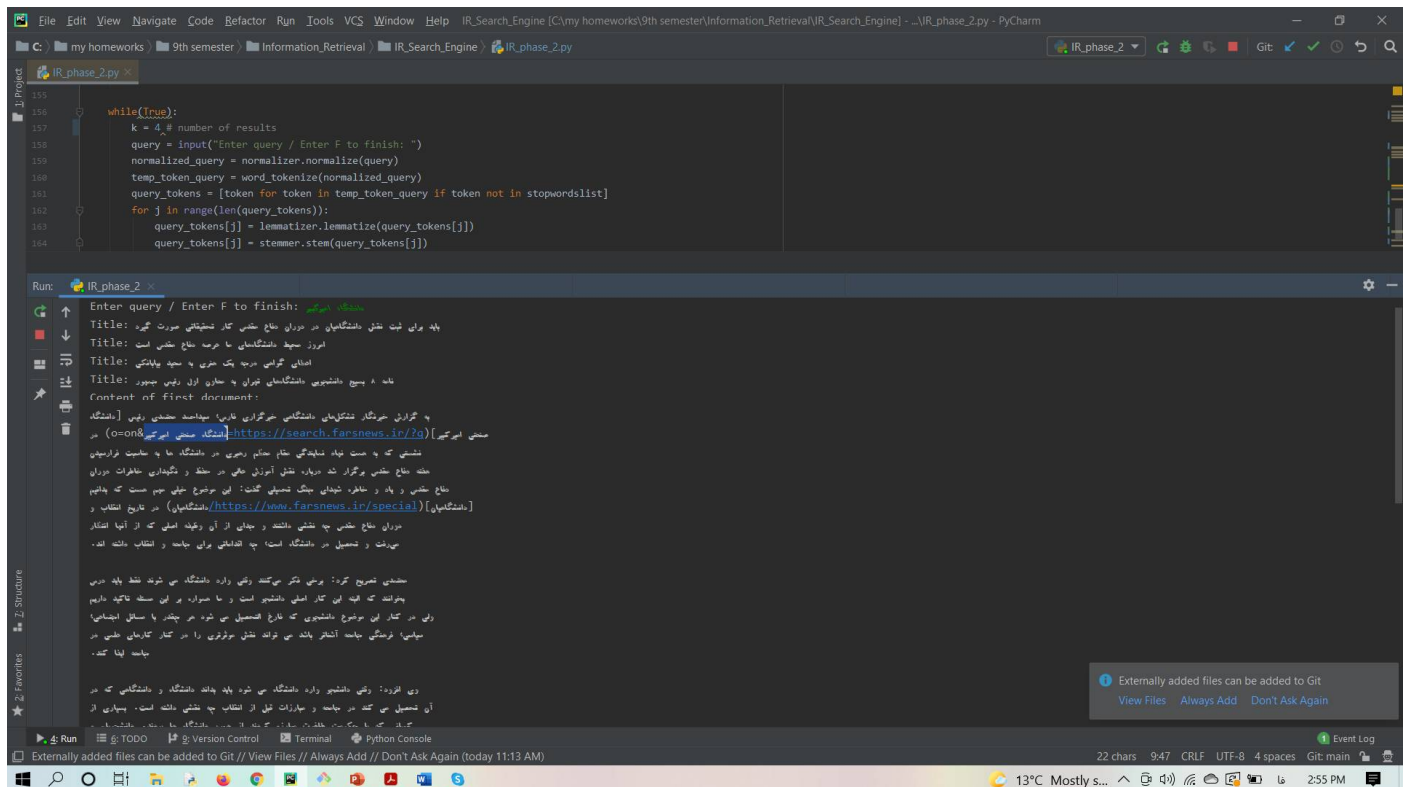


نتیجه ی فاز اول _

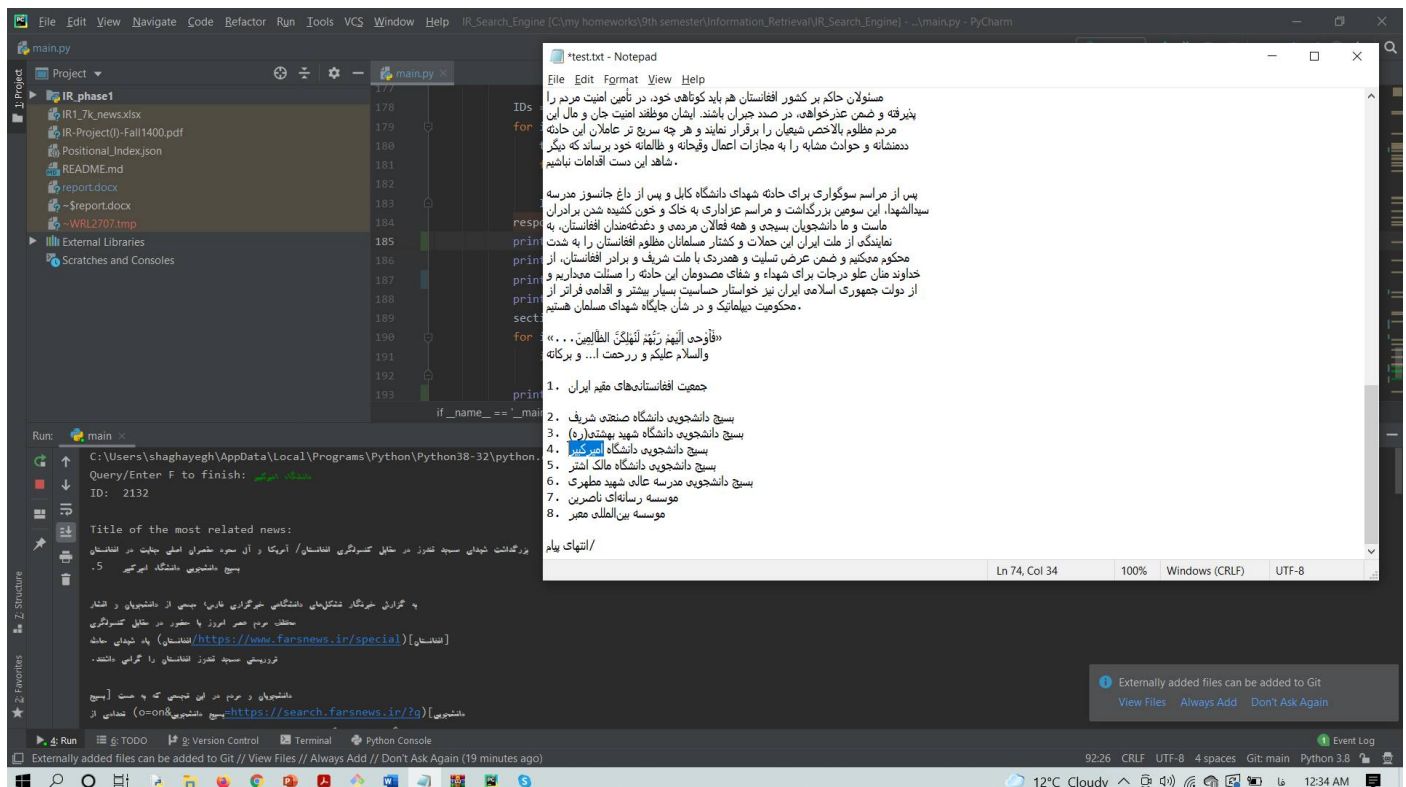


(ب) یک پرسمان از عبارات ساده و متداول دو کلمه ای (دانشگاه امیرکبیر)

نتیجه ی فاز دوم –

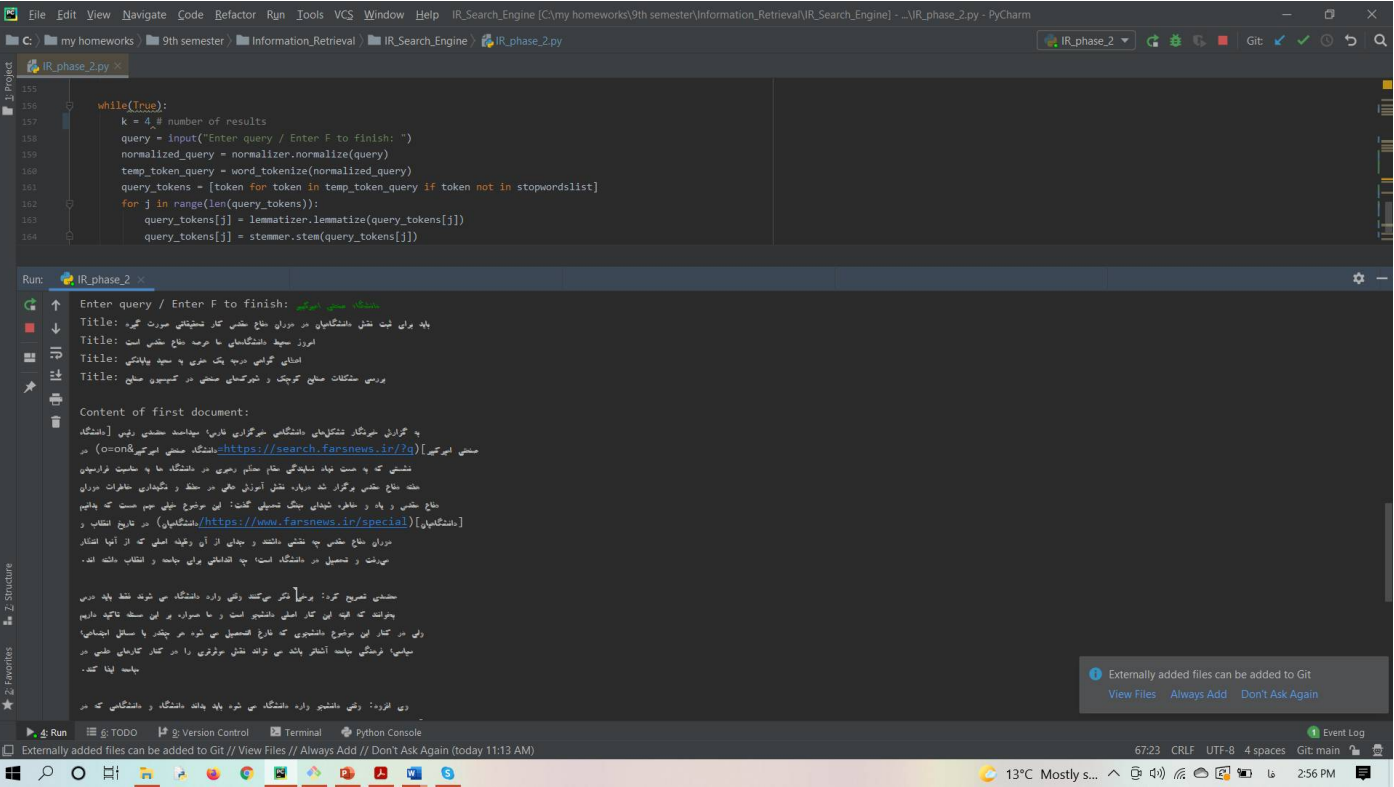


نتیجہ ی فاز اول –

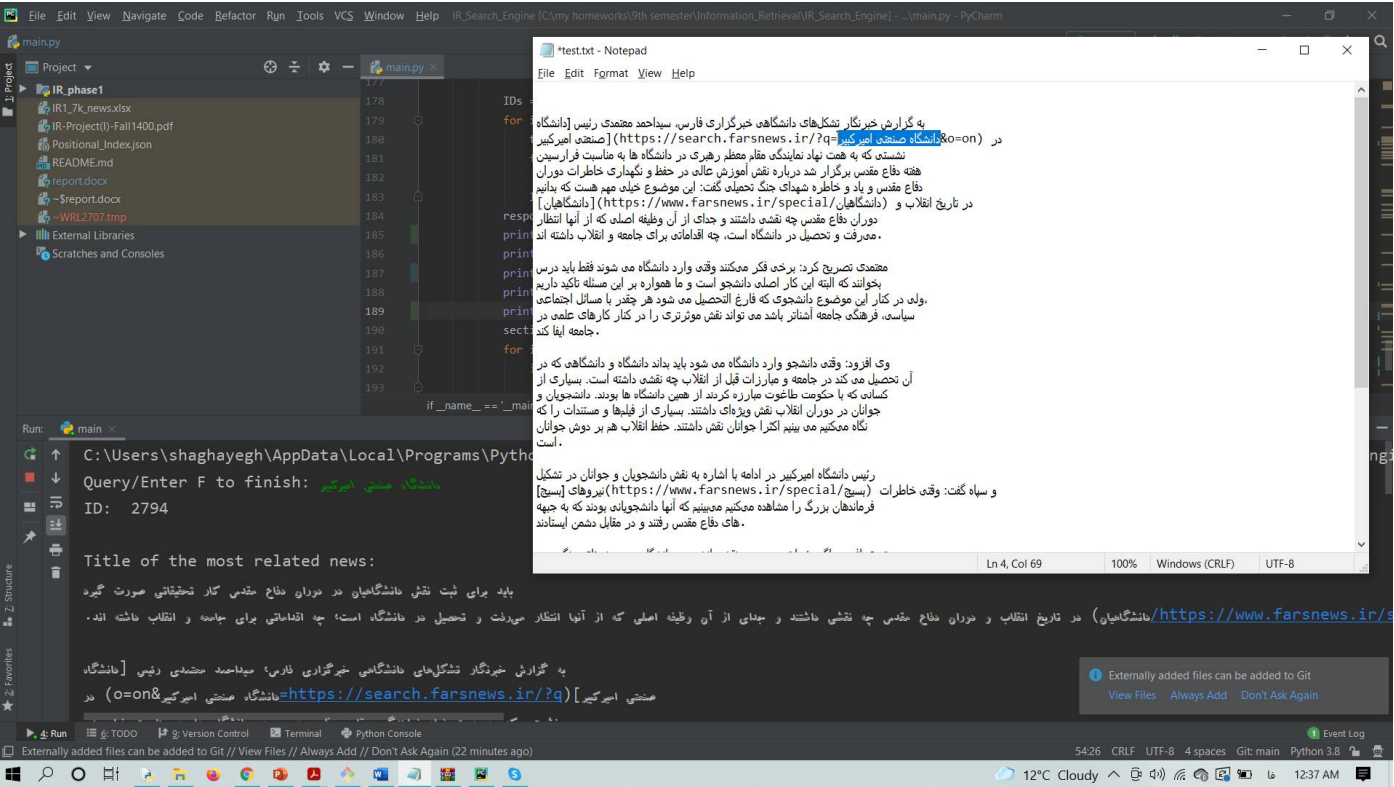


پ) یک پرسمان از عبارات ساده و متداول چند کلمه ای

نتیجه ی پرسمان از عبارات ساده و متداول چند کلمه ای فاز دوم دانشگاه صنعتی امیرکبیر



نتیجه ی فاز اول _ دانشگاه صنعتی امیرکبیر



نتیجه ی فاز دوم _ سازمان ملل متحد

```

155 while(True):
156     k = 4 # number of results
157     query = input("Enter query / Enter F to finish: ")
158     normalized_query = normalizer.normalize(query)
159     temp_token_query = word_tokenize(normalized_query)
160     query_tokens = [token for token in temp_token_query if token not in stopwordslist]
161     for j in range(len(query_tokens)):
162         query_tokens[j] = lemmatizer.lemmatize(query_tokens[j])
163         query_tokens[j] = stemmer.stem(query_tokens[j])

```

Run: IR_phase_2

Enter query / Enter F to finish: سازمان ملل متحد

Title: هیئت پارلمانی ایران به دوم سنتر کرده

Title: شرمه‌ای می‌دانی

Title: انتشار «کودکان و حقوق بشر» برای بهیچدا

Title: راستیه: افکار آیت الله رئیسی مخالفه‌گری افغانی، دو سازمان ملل بود / تأکید رئیس جمهور بر حفاظت ملت

Content of first document:

به گزارش سوله پارلمانی خیرگزاری دایمی، هیئت پارلمانی جمهوری اسلامی ایران به ستور شرکت در نشست مقدماتی کاپ ۲۶ (کنفرانس تشریفات آب و هوا و اقلیمی) به شهر دوم سنتر کرده.

نشست مقدماتی کاپ ۲۶ با حضور رؤسا و نمایندگان پارلمانی‌های ۷۲ کشور و ۸ سازمان بین‌المللی به میزبانی مجلس نمایندگان ایتالیا و با مشارکت آستانه بین‌المللی جهانی IPU و سازمان ملل متحد در روزهای ۱۶ و ۱۷ مهر ماه برگزار خواهد شد.

در این کنفرانس دو خصوصی اهداف جهانی جهت مقابله با تشریفات اقلیمی، روزگرمیز جهت پیرویه وضعیت زمینی در دوران کووید-۱۹، تأمین مالی میاست‌های جهانی در سوله، تشریفات اقلیم و نقش پارلمانی‌ها در دستیابی به اهداف کاپ ۲۶ گنجانده خواهد شد.

قراردست پیل‌نویس اولیه سند نوایی کنفرانس پارلمانی سازمان ملل متحد دو خصوصی تشریفات اقلیم دو این کنفرانس به تصویب برسد.

نتیجه ی فاز اول _ سازمان ملل متحد

```

177 IDs = []
178 for i in range(len(binary_response)):
179     temp = []
180     for j in binary_response[i]:
181         temp.append(j[0])
182     IDs.append(temp)
183 response = list(set.intersection(*map(set, IDs)))
184 print("ID: ", response[0])
185 print("")
186 print("Title of the most related news: ")
187 print(documentCollection[response[0] - 1].title)
188 print("")
189 sections = sent_tokenize(documentCollection[response[0] - 1].content)
190 for i in sections:
191     if query in i:
192         print(i)
193 if __name__ == '__main__': while (True): if (len(query_words) > 1)

```

Run: main

Query/Enter F to finish: سازمان ملل متحد

ID: 6912

Title of the most related news:

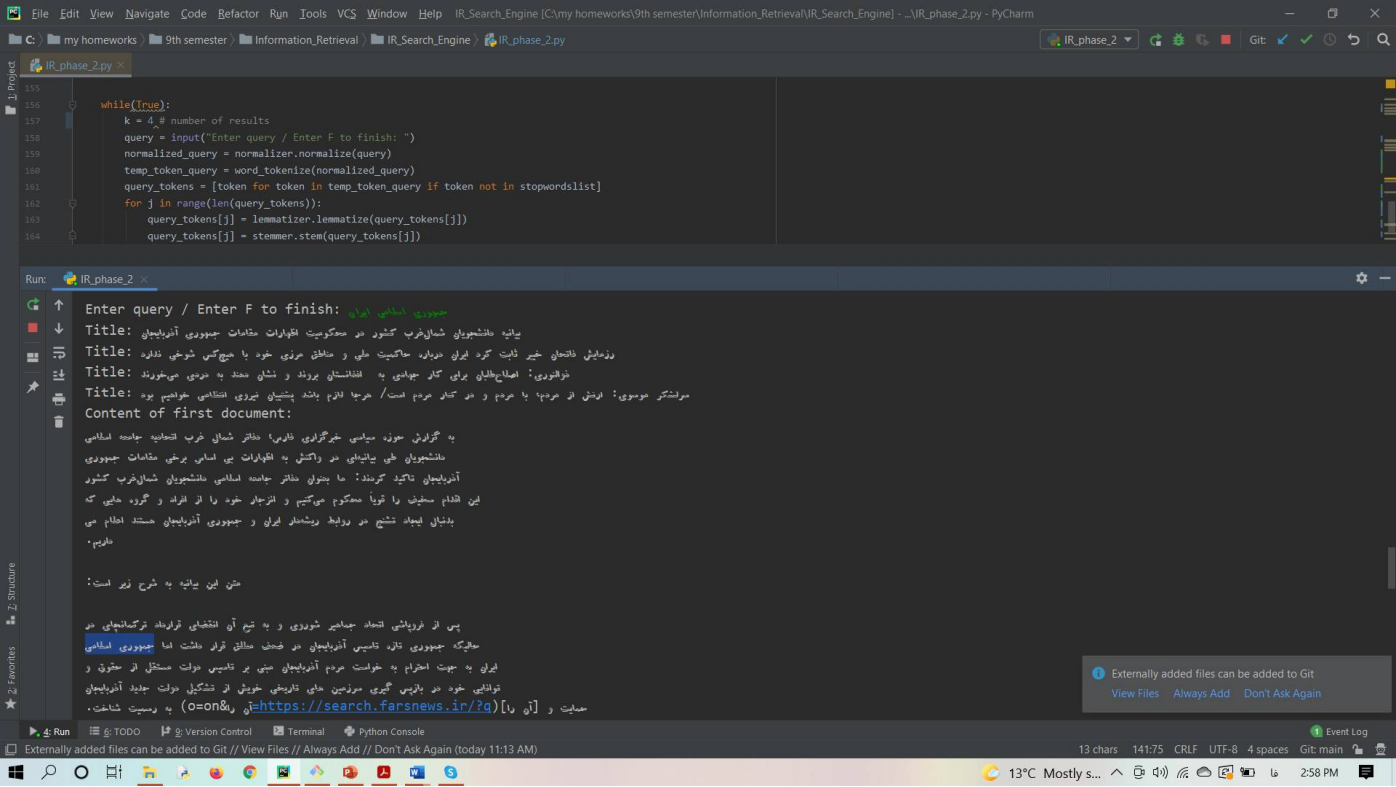
انتشار «کودکان و حقوق بشر» برای بهیچدا

کنوانسیون حقوق کودک سازمان ملل متحد با تصویب پیمان‌نامه‌ای خنارین موسی از حقوق کودک را تصویب کرده که در پی آن کشورهای زیادی با اضایی این پیمان‌نامه حقوق کودک را محترم شمرند.

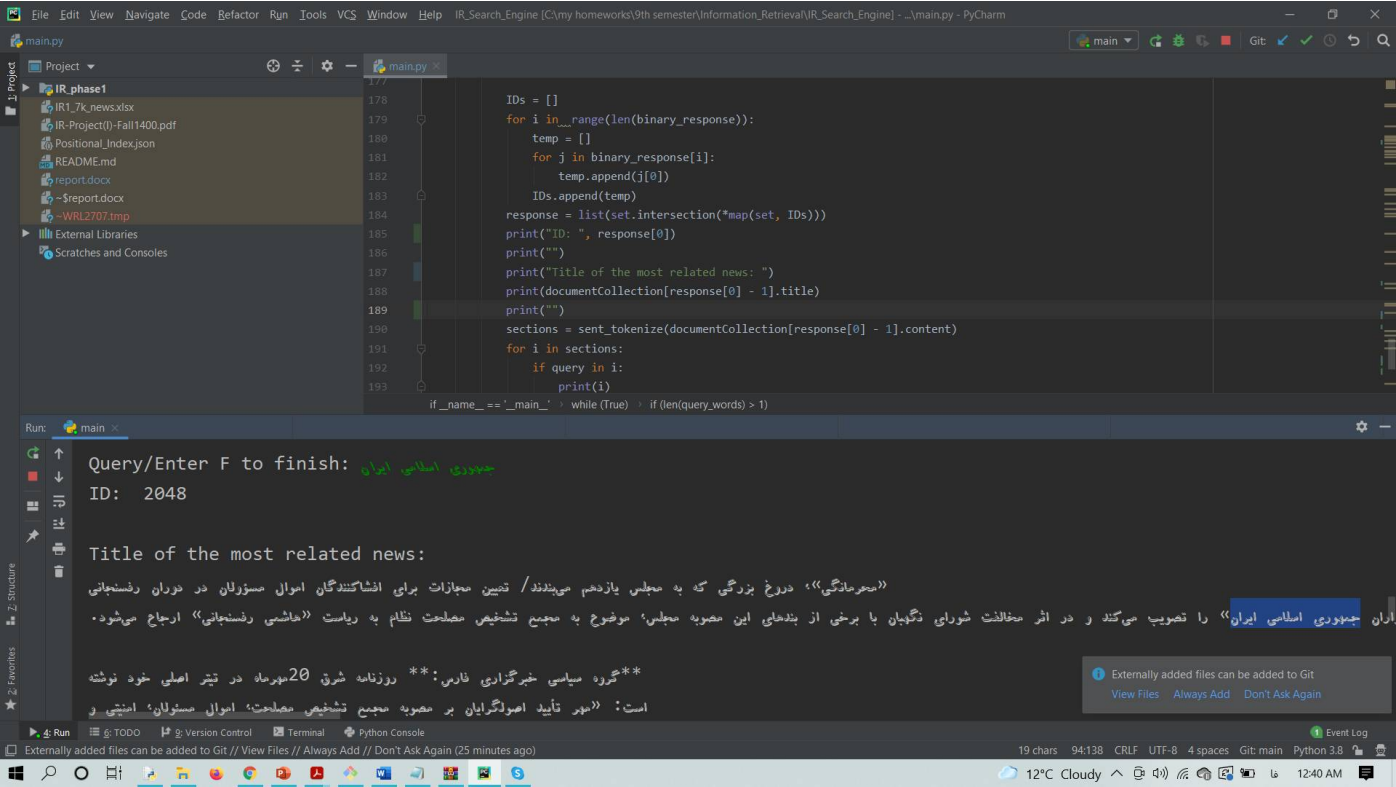
این مجموعه دوجلدی با تکیه بر پیمان‌نامه حقوق کودک سازمان ملل متحد بر آن است کودکان را بر حقوق فردی و اجتماعی خود آشنا کند.

به گزارش ایسنا، در کتاب «حقوق فردی من» و «حقوق اجتماعی من» ذیلی عنوان «کودکان و

نتیجه ی فاز دوم جمهوری اسلامی ایران



نتیجه ی فاز اول _ جمهوری اسلامی ایران



ت) یک پرسمان دشوار و کم تکرار تک کلمه ای (ژیمناستیک)

نتیجه

دوم

فاز

ی

File Edit View Navigate Code Refactor Run Tools VCS Window Help IR_Search_Engine [C:\my homeworks\9th semester\Information_Retrieval\IR_Search_Engine] - IR_phase_2.py - PyCharm

IR_phase_2.py

155
156
157 while(True):
158 k = 4 # number of results
159 query = input("Enter query / Enter F to finish: ")
160 normalized_query = normalizer.normalize(query)
161 temp_token_query = word_tokenize(normalized_query)
162 query_tokens = [token for token in temp_token_query if token not in stopwordslist]
163 for j in range(len(query_tokens)):
164 query_tokens[j] = lemmatizer.lemmatize(query_tokens[j])
165 query_tokens[j] = stemmer.stem(query_tokens[j])

Run: IR_phase_2

Enter query / Enter F to finish:
Title: هشدار حیات ژیمناستیک تهرانی در خصوص مسائل مالی مضطرب و اقدامات غیرعادی
Title: مدیر منابع داراییون ژیمناستیک مشخص شد
Title: جیروکود: برخی به دنبال طبع گرمی ژیمناستیک هستند / با پایکوت داراییون موافقتها بیشتر شد
Title: ثبت نام ۱۲ نامزد برای پست ریاست داراییون ژیمناستیک + اسامی
Content of first document:
به گزارش خبرنگاری فارس؛ رضا شاکرمی در خصوص انتشار خبری جینی بر وضعیت مالی مضطرب ژیمناستیک در یکی از مناطق جنوب شهر تهرانی اظهار داشت: حاشیه ها ما هم متوجه شدیم که این مسئله واقعتا درود و به پیگیری این موضوع [به صورت] [https://search.farsnews.ir/?q=on&on=] جوی پرداختیم. این مالی ژیمناستیک تحت اجازه فردی خاصی است و حاشیه ها خودی ژیمناستیکها دیگر را تشرین میدهد؛ وی عنوان کرد که مالی را اجازه داده است اما واقعتا این است که این فرد خودی مالی را اجازه کرده؛ پی چگونه ممکن است مالی که اجازه کرده را اجازه بدهد؟ بنابراین این اظهار تار او می تواند برای افراد از قانونی بود باشد. ضمن اینکه اجازه مالی که در اجازه است هم خلاف قانونی مصوب می شود.
وی افزود: این فرد که حاشیه ها هم از پیشگویی ژیمناستیک و هم تصدیق بعضی از ژیمناستیک تهرانی است؛ جوی کرده از تشرین مضطرب و پسری در مالی ژیمناستیک اطلاع نموده است؛ اما این اظهار تار در ژیمناستیک انجام شده که در چند روز

Externally added files can be added to Git // View Files // Always Add // Don't Ask Again (today 11:13 AM)

9 chars 223:47 CRLF UTF-8 4 spaces Git: main

13°C Mostly s... 2:59 PM

نتیجه ی فاز اول -

نتیجه

دوم

فاز

ی

File Edit View Navigate Code Refactor Run Tools VCS Window Help IR_Search_Engine [C:\my homeworks\9th semester\Information_Retrieval\IR_Search_Engine] - C:\my homeworks\9th semester\Information_Retrieval\IR_phase1\main.py

main.py

if __name__ == '__main__': while (True): if (len(query.words) == 1)
Run: IR_phase_2 main (1)
Query/Enter F to finish:
Number of repetitions in documents: 8
Title of the most related news:
جیروکود: برخی به دنبال طبع گرمی ژیمناستیک هستند / با پایکوت داراییون موافقتها بیشتر شد.
رضا جیروکود در گفتارگو با خبرنگار ورزشی سوختاری فارس؛ درباره عدم اتمام تیم ملی ژیمناستیک خبری به سببها جینی ژایی اظهار داشت: واقعا بابت این موضوع متاسفم. داراییون از حاشیه مالی درخواست اتمام تیم ملی ژیمناستیک را به شورای بزرگ وزارت ورزش داده بود و بارها این مسئله را پیگیری کرده اما متاسفانه این شورا هیچ پاسخی نداد. وی افزود: در مورد مالی این رقابتها جینی اسدکیمی بعضی تهیه تاریخ ژیمناستیک کشور را کسب کرده بود و ناشی از این برای کسب قیام خوب داشت. حاشیه برخی به دنبال این هستند که به هر نحو ممکن ژیمناستیک ایران را طبع کنند تا از این طریق حواشی را به گرمی بکشند. مورس تیم ملی ژیمناستیک خبری گفت: از آنای جینی جواسق میگویم مدای اصلی ژیمناستیک کشور را بکشند. شاید اگر فای ریاست داراییون در اختیار کسی بود که برایش فوای نمیگردد ریسی کدام داراییون است؛ اینگونه اقدامات جینی ژیمناستیک کشور انجام نمیشد اما حاشا جینی جیو نباشد است. جینی که هر چه بیشتر ژیمناستیک را پایکوت کرده؛ موافقتها بیشتر شد. جیروکود در مورد صدور پیشگیری ایران در رنگت جینی تصریح کرد: چنین اتفاق تاریخی اتفاق داد که ژیمناستیک ایران به کتی حصول شد است؛ آن هم در شرایطی که بعضی تلاش کردند جیو پای جیرو داراییون بگذارند. مورس تیم ملی ژیمناستیک خبری خاطرنشان کرد: شاید که برخی از سببها [جینی شورای اسامی] [https://www.farsnews.ir/special] جینی شورای اسامی) قصد دارند به اقداماتی که جینی داراییون شده؛ درود کنند. از حلق ژیمناستیک از هیواد اسدکیمی و جیوگوری از اتمام طبعشان به سببها جینی تا شایع دومین داراییون.
رضا جیروکود در گفتارگو با خبرنگار ورزشی سوختاری فارس؛ درباره عدم اتمام تیم ملی ژیمناستیک خبری به سببها جینی ژایی اظهار داشت: واقعا بابت این موضوع متاسفم. داراییون از حاشیه مالی درخواست اتمام تیم ملی ژیمناستیک را به شورای بزرگ وزارت ورزش داده بود و بارها این مسئله را پیگیری کرده اما متاسفانه این شورا هیچ پاسخی نداد.
وی افزود: در مورد مالی این رقابتها جینی اسدکیمی بعضی تهیه تاریخ ژیمناستیک کشور را کسب کرده بود و ناشی از این برای کسب قیام خوب داشت. جینی به مدای ریسی که با وزنگای داده حلقه دنیا زاینت میگذرد. سببها جینی هم که حقیقت جیرو جینی مالی دنیا بود. از طرف دیگر کامیابی کسبجیو ورزشکاران است و باید به ژایی سببها تا برای کسب آرای لازم با سایر ورزشکاران رفایی میگذرد. متاسفانه برخی به دنبال این هستند که به هر نحو ممکن ژیمناستیک ایران را طبع کنند تا از این طریق حواشی را به گرمی بکشند.

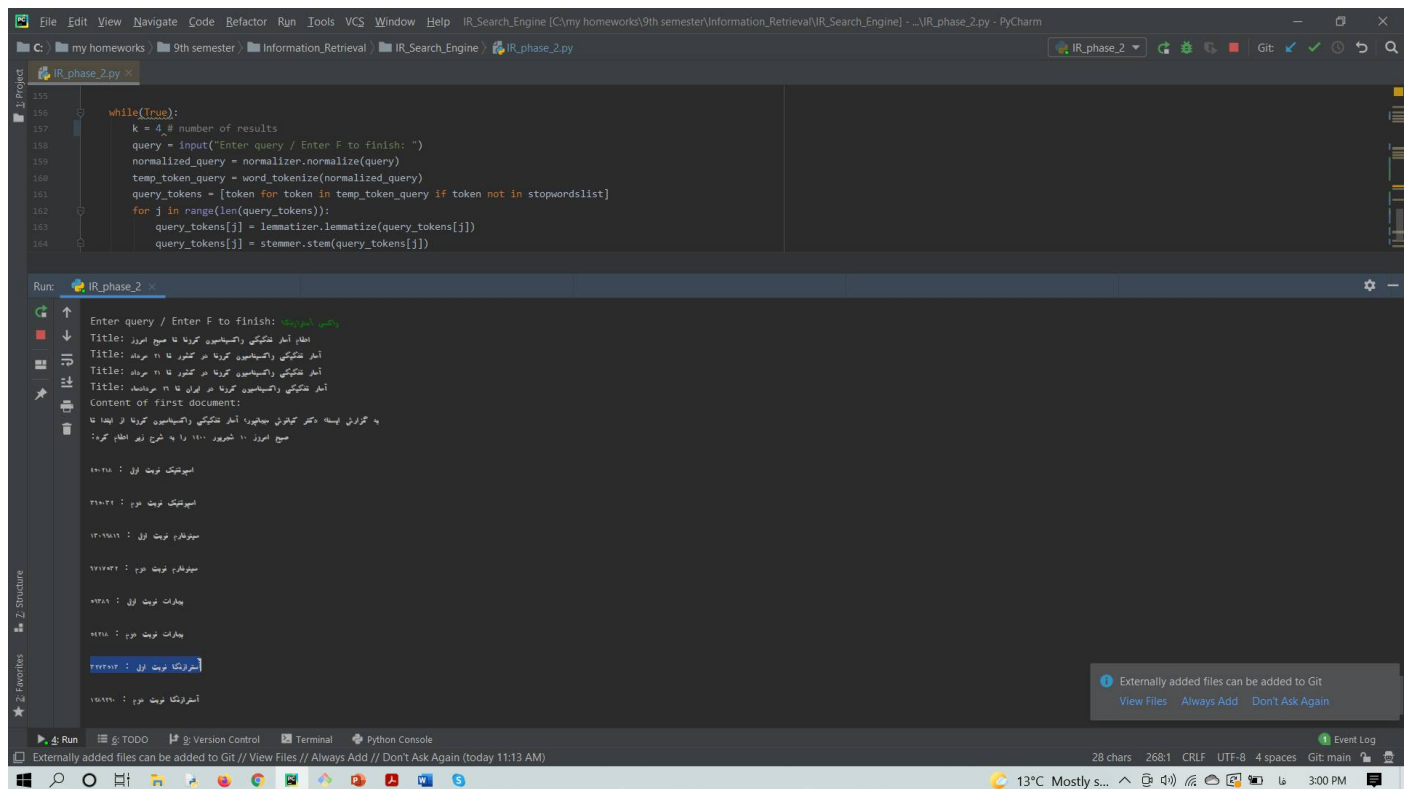
Externally added files can be added to Git // View Files // Always Add // Don't Ask Again (today 11:13 AM)

266:26 CRLF UTF-8 4 spaces Git: main

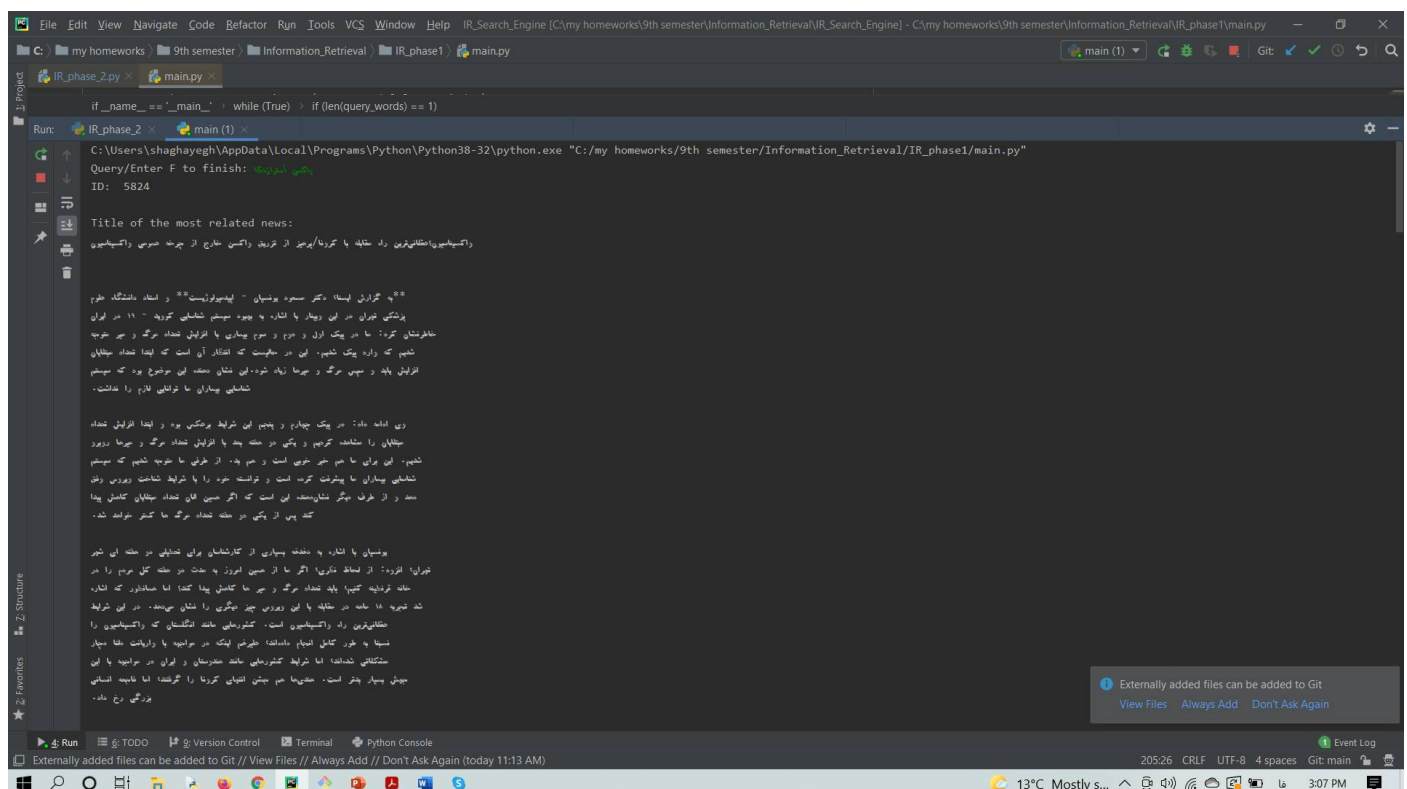
13°C Mostly s... 3:08 PM

نتیجہ ی

نتیجه	ی	فاز	دوم-
-------	---	-----	------



نتیجه ی فاز اول –



سیستمی که در این فاز پیاده سازی شده است، سند های بسیار مطابقی را به عنوان نتایج بازیابی کرده است. برای بررسی این موضوع متن سندی که بیشترین امتیاز را دارد چاپ شده است. هم چنین عنوان ۴ سند مرتبط تر را نشان می دهد ($k = 4$).

۲. تمامی پرسمان های مطرح شده در سیستم بازیابی اطلاعات فاز اول و دوم در گزارش آمده است. برای پرسمان "ژیمناستیک"، سندی که به عنوان بهترین سند در فاز یک برگردانده شده، در فاز دوم، سومین مرتبط ترین سند است. برای پرسمان "واکسن آسترازنکا"، سند بازیابی شده از فاز اول در هیچ کدام از نتایج فاز دوم نمایش داده نشده است.

سیستم فاز اول نتوانسته سند مناسبی را برای پرسمان "واکسن آسترازنکا" بیابد. زیرا دنبال سندی بوده است که این دو کلمه را دقیقا کنار هم داشته باشد. اما سیستم فاز دوم سندی را برگردانده که درباره ی نوبت های تزریق واکسن صحبت کرده و به نظر می رسد بهتر می تواند کاربر را به هدفش برساند. این قضیه درباره ی پرسمان های تک کلمه ای آنچنان مشهود نیست. اما برای پرسمان های چند کلمه ای، شاخص مکانی سندی را بهتر می داند که دقیقا حاوی همان ترکیب از کلمات باشد. اما سیستم فاز دوم سندی که کلمات کوثری را با تکرار بیشتری داشته باشد مرتبط تر می داند (استفاده از tf-idf). این قضیه باعث شده برای کوثری های طولانی تر سیستم فاز دوم بهتر عمل کند مثال آن در کوثری واکسن آسترازنکا و دانشگاه امیرکبیر مشهود است. سیستم فاز اول، برای کوثری دانشگاه امیرکبیر سندی را برگردانده که عموما راجع به دانشگاه ها صحبت می کند و فقط یک مثال از دانشگاه امیرکبیر زده است در حالی که سیستم دوم اتفاقی که در دانشگاه امیرکبیر افتاده است را مطرح می کند که به وضوح سند مرتبط تری است.