Assignment 4

Submitted by
Swati Verma
MT19073

Question 1.
NOTE:

1. **Proper analysis is written for the first query.**
2. **For the other 2 queries as the analysis is same because algorithm is same, only the screenshots are attached with proper labels.**
3. **All the necessary comments are written in the code you can refer it.**
4. **How the code is working i.e. methodology is written in Readme file**

1. Vector corresponding to the document 58044

```
print("vector corresponding to document 58044 is:\n",final_index['58044'])
```

```
vector corresponding to document 58044 is:
 {'xref': 0.31931414706795297, 'cantaloupesrvcscmuedu': 0.31931414706795297, 'compgraphics37261': 0, 'altgraphics519': 0, 'co
mpgraphicsanimation2614': 0, 'path': -6.0199979335502726e-05, 'cantaloupesrvcscmuedudasnewsharvardeduogicseuwmeduzaphodmpsohi
ostateedudarwinsuranetdtixdtnavymiloasyslipman': 0, 'lipmanoasysdtnavymil': 0, 'robert': 0, 'lipman': 0, 'newsgroups': -6.019
9979335502726e-05, 'compgraphicsaltgraphicscompgraphicsanimation': 0, 'subject': -6.0199979335502726e-05, 'call': 0.797947775
7398055, 'presentation': 1.468908083591528, 'navy': 0, 'scivizvr': 0, 'seminar': 0, 'messageid': -6.0199979335502726e-05, '32
850oasysdtnavymil': 0, 'date': -6.0199979335502726e-05, 'nineteen': 0, 'mar': 0, 'ninety-three': 0.6708156951239134, 'two hun
dred and one thousand and twenty-three': 0, 'gmt': 0.0563084828529976, 'articleid': 0, 'oasys32850': 0, 'expires': 0, 'thirt
y': 1.3935627792731005, 'apr': 0.020553979135905908, 'forty thousand': 0, 'replyto': 0.5585125258208924, 'followupto': 0.7962
470367771488, 'compgraphics': 0, 'distribution': 0, 'usa': 0, 'organization': 0.012569270568334682, 'carderock': 0, 'divisio
n': 0, 'nswc': 0, 'bethesda': 0, 'md': 0, 'line': 0.0010252458926111345, 'sixty-five': 0, 'scientific': 0, 'visualization':
0, 'virtual': 0, 'reality': 0, 'tuesday': 0, 'june': 1.3167228007603868, 'twenty-two': 0, 'one thousand, nine hundred and nin
ety-three': 0.44370530451490986, 'naval': 0, 'surface': 0, 'warfare': 0, 'center': 0.7837888453350919, 'formerly': 0, 'davi
d': 0.7550976223218626, 'taylor': 0, 'research': 0.7477907102928021, 'maryland': 0, 'sponsor': 0, 'ness': 0, 'engineering':
0, 'software': 0, 'system': 0, 'sponsoring': 0, 'oneday': 0, 'purpose': 0, 'present': 0, 'exchange': 0, 'information': 0.6784
598776515762, 'navyrelated': 0, 'program': 0.6699795986804411, 'development': 0.9990270907001494, 'application': 0, 'solicite
d': 0, 'aspect': 0, 'current': 0, 'work': 0.5983797049109253, 'worksinprogress': 0, 'proposed': 0, 'considered': 1.0196772415
713644, 'four': 1.4153816563001844, 'type': 0, 'available': 2.3296623942318773, 'one': 0.7678272374145099, 'regular': 0, 'two
thousand and thirty': 0, 'minute': 0.9295347383389095, 'length': 0, 'two': 1.141255414898942, 'short': 0, 'ten': 0.9913829735
394784, 'three': 1.679715685129124, 'video': 0, 'standalone': 0, 'videotape': 0, 'author': 1.8570610704597628, 'need': 0, 'at
```

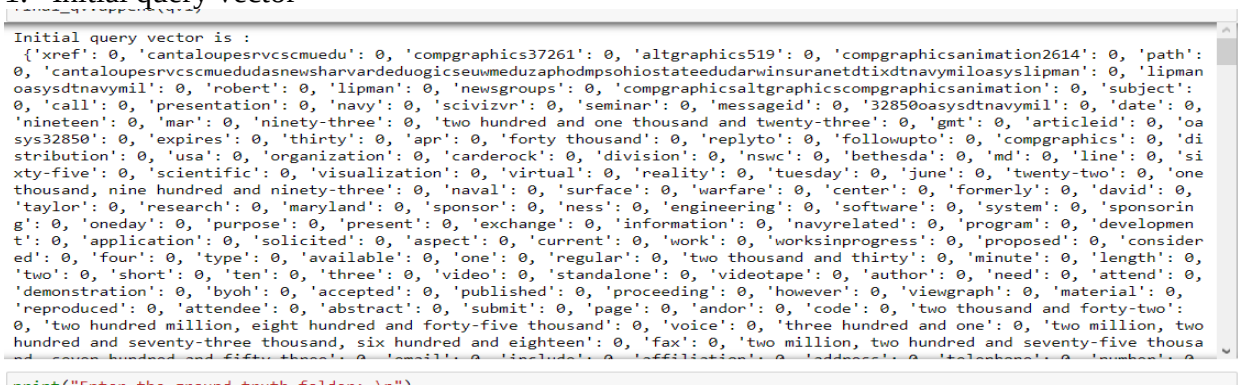
2. Query processing using cosine similarity
   Cosine similarity is taken between query and each of the documents and user is asked to enter the number of documents he or she wants to retrieve. Top K(here K=30) documents are then retrieved.

```
query=input("enter the query \n")
k=int(input("enter the value of k \n"))
ans=query_cosine(query,k)

print("Top ",k, " documents are: \n", ans)
```

```
enter the query
Marriott Hotel, Salt Lake City, Utah.  The business sessions, Karl Hess Institute,
enter the value of k
30
Top  30  documents are:
 [('176875', 0.2850513918879685), ('58044', 0.2849623062822999), ('60152', 0.2849623062822999), ('176874', 0.2362890981091281
3), ('61130', 0.15041991596444543), ('176873', 0.10201383208490092), ('59110', 0.10059366700110683), ('176929', 0.0963858684165
6063), ('61430', 0.08216981157321536), ('54763', 0.07720874523896099), ('60913', 0.06998218999118484), ('58068', 0.068115751990
03046), ('60186', 0.06811575199003046), ('59436', 0.06782503378775194), ('38530', 0.0659681131514981), ('38915', 0.062467930449
06809), ('53788', 0.06207718377153357), ('52557', 0.05999992170368692), ('59227', 0.05761009143341051), ('38993', 0.05484515073
295232), ('178810', 0.05404624741039915), ('52576', 0.052629108299636615), ('38357', 0.05085889804549226), ('61488', 0.04842620
478532875), ('38584', 0.047084024919294684), ('60819', 0.04627260106215433), ('54220', 0.04565468247108565), ('59154', 0.043451
62237723888), ('61159', 0.04332196465274718), ('59100', 0.04331656228683863)]
```

Question 2.
Query 1: Pretty good opinions on biochemistry machines
Relevant set 1: Documents inside folder sci.med

1. Initial query vector



```
Initial query vector is :
 {'xref': 0, 'cantaloupesrvcscmuedu': 0, 'compgraphics37261': 0, 'altgraphics519': 0, 'compgraphicsanimation2614': 0, 'path':
0, 'cantaloupesrvcscmuedudasnewsharvardeduogicseuwmeduzaphodmpsohiostateedudarwinsuranetdtixdtnavymiloasyslipman': 0, 'lipman
oasysdtnavymil': 0, 'robert': 0, 'lipman': 0, 'newsgroups': 0, 'compgraphicsaltgraphicscompgraphicsanimation': 0, 'subject':
0, 'call': 0, 'presentation': 0, 'navy': 0, 'scivizvr': 0, 'seminar': 0, 'messageid': 0, '32850oasysdtnavymil': 0, 'date': 0,
'nineteen': 0, 'mar': 0, 'ninety-three': 0, 'two hundred and one thousand and twenty-three': 0, 'gmt': 0, 'articleid': 0, 'oa
sys32850': 0, 'expires': 0, 'thirty': 0, 'apr': 0, 'forty thousand': 0, 'replyto': 0, 'followupto': 0, 'compgraphics': 0, 'di
stribution': 0, 'usa': 0, 'organization': 0, 'carderock': 0, 'division': 0, 'nswc': 0, 'bethesda': 0, 'md': 0, 'line': 0, 'si
xty-five': 0, 'scientific': 0, 'visualization': 0, 'virtual': 0, 'reality': 0, 'tuesday': 0, 'june': 0, 'twenty-two': 0, 'one
thousand, nine hundred and ninety-three': 0, 'naval': 0, 'surface': 0, 'warfare': 0, 'center': 0, 'formerly': 0, 'david': 0,
'taylor': 0, 'research': 0, 'maryland': 0, 'sponsor': 0, 'ness': 0, 'engineering': 0, 'software': 0, 'system': 0, 'sponsorin
g': 0, 'oneday': 0, 'purpose': 0, 'present': 0, 'exchange': 0, 'information': 0, 'navyrelated': 0, 'program': 0, 'developmen
t': 0, 'application': 0, 'solicited': 0, 'aspect': 0, 'current': 0, 'work': 0, 'worksinprogress': 0, 'proposed': 0, 'consider
ed': 0, 'four': 0, 'type': 0, 'available': 0, 'one': 0, 'regular': 0, 'two thousand and thirty': 0, 'minute': 0, 'length': 0,
'two': 0, 'short': 0, 'ten': 0, 'three': 0, 'video': 0, 'standalone': 0, 'videotape': 0, 'author': 0, 'need': 0, 'attend': 0,
'demonstration': 0, 'byoh': 0, 'accepted': 0, 'published': 0, 'proceeding': 0, 'however': 0, 'viewgraph': 0, 'material': 0,
'reproduced': 0, 'attendee': 0, 'abstract': 0, 'submit': 0, 'page': 0, 'andor': 0, 'code': 0, 'two thousand and forty-two':
0, 'two hundred million, eight hundred and forty-five thousand': 0, 'voice': 0, 'three hundred and one': 0, 'two million, two
hundred and seventy-three thousand, six hundred and eighteen': 0, 'fax': 0, 'two million, two hundred and seventy-five thousa
```

2. Result for initial query
   Cosine similarity is taken between query and each of the documents and user is asked to enter the number of documents he or she wants to retrieve. Top K(here K=100) documents are then retrieved.

```
enter the intial query
 Pretty good opinions on biochemistry machines
enter the value of k
100
top  100 documents after initial query are:
 [('38523', 0.16207741640316015), ('58082', 0.14645771525152543), ('38597', 0.115114041925314), ('38774', 0.10143287135219713),
('59504', 0.08971101479913714), ('59116', 0.08828470069490048), ('59393', 0.08641567280050193), ('38837', 0.08362804365250232),
('59400', 0.08357033547644269), ('59301', 0.08101416066013839), ('59060', 0.08008790202852739), ('59602', 0.07890467522083179),
('59488', 0.07842171397093384), ('58872', 0.07705692246646091), ('58987', 0.07643880456067242), ('58813', 0.07067821695592733),
('58801', 0.06625821253783812), ('59632', 0.06589275433632238), ('38959', 0.0656815848260574), ('59233', 0.064293011349751 32),
('59518', 0.06082384289872259), ('59207', 0.05580274954657562), ('37935', 0.054794101428834396), ('38845', 0.0547546600999183
7), ('53677', 0.05419334397418276), ('38649', 0.05418674652906282), ('59546', 0.052733167766448 53), ('53998', 0.052696359108984
12), ('38742', 0.05253948165201494), ('39052', 0.05247307012736698), ('53646', 0.05237616702312154), ('59165', 0.05231769113813
108), ('38416', 0.05116060325356 4245), ('60877', 0.05116060325356 4245), ('38240', 0.05090162932599169), ('59308', 0.05070658181
276308), ('179106', 0.04978991203753135), ('38908', 0.0497177425480131), ('38669', 0.04910868779220102), ('38538', 0.0485637020
7799187), ('59117', 0.04854907009730821), ('38304', 0.04843800700610383), ('176953', 0.04752709354435089), ('39033', 0.04722568
4492723435), ('178302', 0.0470703981554837), ('52653', 0.046738180321131514), ('53663', 0.04672607210166793), ('177004', 0.0465
7743140950671), ('38554', 0.0459562455415383), ('39001', 0.04571630207110923), ('53791', 0.04521698662272852), ('38438', 0.0449
47894894039206), ('176902', 0.04487395063266615), ('38730', 0.04464306036694627), ('54102', 0.04449963752090 2496), ('59183', 0.
04427628224745738), ('53827', 0.04418412456044122), ('59083', 0.04411641663912072), ('54098', 0.04399854898624299), ('39630',
0.04379801274699401), ('52565', 0.04377072323673542), ('62386', 0.04340722313337434), ('60942', 0.04316099853263107), ('38759',
0.04303269169007668), ('61454', 0.04292321408688534), ('59043', 0.04240849985089937 4), ('61329', 0.04185476821198743 5), ('17831
1', 0.04173556601631852), ('38395', 0.04159501735123928), ('59115', 0.0413152540343381), ('61309', 0.04053645157467716), ('5375
7', 0.039746660119488145), ('53995', 0.039662308504088375), ('38442', 0.0395678439220583), ('37922', 0.03954334907933634), ('59
063', 0.03906470341284782), ('38994', 0.03894347406512957), ('52552', 0.03882280215649338), ('53762', 0.0384431576311407), ('38
524', 0.03805452707606915 6), ('59238', 0.03776615334905368 5), ('38732', 0.03774735838851741), ('61183', 0.0376864225962759),
('38458', 0.03764614653965577), ('52578', 0.03756801196985535), ('59017', 0.037475202917518054), ('58953', 0.0370329265561007
8), ('38392', 0.036697629809315645), ('178804', 0.03657071728572572), ('38924', 0.03655282908580305), ('38955', 0.0364456739626
6398), ('59637', 0.03624772860888512), ('54726', 0.03621914595379859), ('38965', 0.035933942052888496), ('38549', 0.03586017819
677725), ('38478', 0.035764006965696836), ('53630', 0.035635871382556764), ('61068', 0.035477139797604175), ('59154', 0.0351182
69391289834), ('58821', 0.0351132446547962 8)]
```

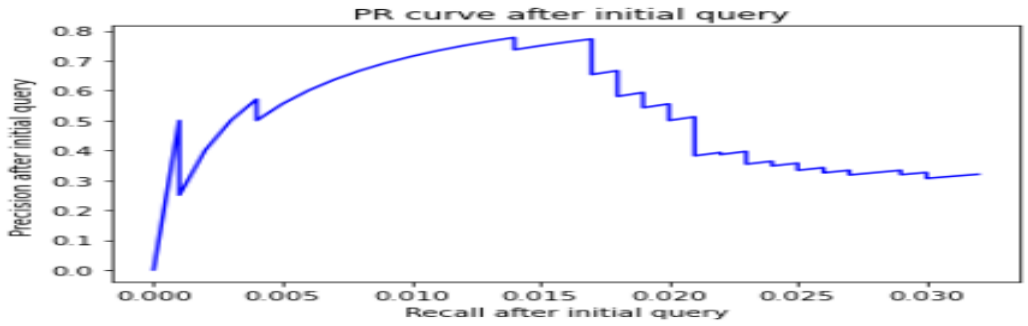
3. Precision, Recall and MAP for initial query
   All the files of folder ‘sci.med’ are taken as ground truth and precision and recall is calculated at each of the 100 points.
   As the top most retrieved document is not from the sci.med folder. Hence, the precision and recall values at first point are coming zero. After that whenever a relevant document comes, the precision value increases and as the number of retrieved relevant documents increases the recall also increases which makes the PR curve as sawtooth shape.

```
Enter the ground truth folder:

sci.med
precision after initial query are:  [0.0, 0.5, 0.3333333333333333, 0.25, 0.4, 0.5, 0.5714285714285714, 0.5, 0.5555555555555556,
0.6, 0.6363636363636364, 0.6666666666666666, 0.6923076923076923, 0.7142857142857143, 0.7333333333333333, 0.75, 0.76470588235294
11, 0.7777777777777778, 0.7368421052631579, 0.75, 0.7619047619047619, 0.7727272727272727, 0.7391304347826086, 0.708333333333333
4, 0.68, 0.6538461538461539, 0.6666666666666666, 0.6428571428571429, 0.6206896551724138, 0.6, 0.5806451612903226, 0.59375, 0.57
57575757575758, 0.5588235294117647, 0.5428571428571428, 0.5555555555555556, 0.5405405405405406, 0.5263157894736842, 0.512820512
8205128, 0.5, 0.5121951219512195, 0.5, 0.4883720930232558, 0.4772727272727273, 0.4666666666666667, 0.45652173913043476, 0.44680
851063829785, 0.4375, 0.42857142857142855, 0.42, 0.4117647058823529, 0.40384615384615385, 0.39622641509433965, 0.38888888888888
89, 0.38181818181818183, 0.39285714285714285, 0.38596491228070173, 0.3965517241379310 5, 0.3898305084745763, 0.3833333333333333
6, 0.3770491803278688, 0.3709677419354839, 0.36507936507936506, 0.359375, 0.35384615384615387, 0.36363636363636365, 0.358208955
2238806, 0.3529411764705882 6, 0.34782608695652173, 0.35714285714285715, 0.352112676056338, 0.3472222222222222, 0.34246575342465
75, 0.33783783783783783, 0.3333333333333333, 0.34210526315789475, 0.3376623376623376 6, 0.3333333333333333, 0.3291139240506329,
0.325, 0.3333333333333333, 0.32926829268292684, 0.3253012048192771, 0.32142857142857145, 0.3176470588235294, 0.3255813953488372
3, 0.3333333333333333, 0.32954545454545453, 0.3258426966292135, 0.32222222222222224, 0.31868131868131866, 0.32608695652173914,
0.3225806451612903, 0.3191489361702128, 0.3157894736842105, 0.3125, 0.3092783505154639 3, 0.30612244897959184, 0.313131313131313
15, 0.32]
recall after initial query are:  [0.0, 0.001, 0.001, 0.001, 0.001, 0.002, 0.003, 0.004, 0.004, 0.005, 0.006, 0.007, 0.008, 0.009, 0.0
1, 0.011, 0.012, 0.013, 0.014, 0.014, 0.015, 0.016, 0.017, 0.017, 0.017, 0.017, 0.017, 0.018, 0.018, 0.018, 0.018, 0.018, 0.01
9, 0.019, 0.019, 0.019, 0.02, 0.02, 0.02, 0.02, 0.02, 0.021, 0.021, 0.021, 0.021, 0.021, 0.021, 0.021, 0.021, 0.021, 0.021, 0.0
21, 0.021, 0.021, 0.021, 0.021, 0.021, 0.022, 0.022, 0.023, 0.023, 0.023, 0.023, 0.023, 0.023, 0.023, 0.023, 0.024, 0.024, 0.024, 0.02
4, 0.025, 0.025, 0.025, 0.025, 0.025, 0.025, 0.026, 0.026, 0.026, 0.026, 0.026, 0.027, 0.027, 0.027, 0.027, 0.027, 0.028, 0.02
9, 0.029, 0.029, 0.029, 0.029, 0.03, 0.03, 0.03, 0.03, 0.03, 0.03, 0.03, 0.031, 0.032]
MAP after initial query are:  0.539968246608691
```

4. Precision-Recall curve for initial query
   Saw tooth shape PR curve.

```
plt.plot(r, p,color="blue")
plt.xlabel("Recall after initial query")
plt.ylabel("Precision after initial query ")
plt.title(" PR curve after initial query")
```

```
Text(0.5, 1.0, ' PR curve after initial query')
```



5. Updated query vector

The user is asked to provide the set of relevant documents. Documents marked by users as relevant are taken as one rel_list and remaining (100- rel_list) are taken as non relevant documents and then using Rocchio's formula a new query vector is calculated with α= 1, β=0.7,andγ=0.25.

```
print("New query vector after 1st iteration is \n", new_query1)
```

```
New query vector after 1st iteration is
 {'xref': 0.044881377337884505, 'cantaloupesrvcscmuedu': 0.044881377337884505, 'compgraphics37261': 0.0, 'altgraphics519': 0.
0, 'compgraphicsanimation2614': 0.0, 'path': 0, 'cantaloupesrvcscmuedudasnewsharvardeduogicseuwmeduzaphodmpsohiostateedudarwi
nsuranetdtixdtnavymiloasyslipman': 0.0, 'lipmanoasysdtnavymil': 0.0, 'robert': 0.05220765738088672, 'lipman': 0.0, 'newsgroup
s': 0, 'compgraphicsaltgraphicscompgraphicsanimation': 0.0, 'subject': 0, 'call': 0.2053165619769735, 'presentation': 0, 'nav
y': 0.0, 'scivizvr': 0.0, 'seminar': 0.0, 'messageid': 0, '32850oasysdtnavymil': 0.0, 'date': 0, 'nineteen': 0.07639811998010
079, 'mar': 0, 'ninety-three': 0.034989556466506616, 'two hundred and one thousand and twenty-three': 0.0, 'gmt': 0.024277564
65987644, 'articleid': 0.051290825407408445, 'oasys32850': 0.0, 'expires': 0, 'thirty': 0.030222965355484317, 'apr': 0.008119
680315478474, 'forty thousand': 0.0, 'replyto': 0.1331121519873127, 'followupto': 0.04689010327687654, 'compgraphics': 0, 'di
stribution': 0.043544944801340316, 'usa': 0.02382073100286985, 'organization': 0.0035133600077517863, 'carderock': 0.0, 'divi
sion': 0.048386652634713395, 'nswc': 0.0, 'bethesda': 0, 'md': 0.33736842552780455, 'line': 0.0004496992272073126, 'sixty-fiv
e': 0, 'scientific': 0.10710571460324249, 'visualization': 0.0, 'virtual': 0, 'reality': 0, 'tuesday': 0, 'june': 0, 'twenty-
two': 0, 'one thousand, nine hundred and ninety-three': 0.037175212507603646, 'naval': 0.0, 'surface': 0.3385199899391944, 'w
arfare': 0.0, 'center': 0, 'formerly': 0.0, 'david': 0, 'taylor': 0.0, 'research': 0.04803989033943232, 'maryland': 0, 'spons
or': 0.0, 'ness': 0.0, 'engineering': 0, 'software': 0, 'system': 0.2005712774128295, 'sponsoring': 0.0, 'oneday': 0.0, 'purp
ose': 0, 'present': 0.0600369548484833, 'exchange': 0.0, 'information': 0.07880507007154663, 'navyrelated': 0.0, 'program':
0, 'development': 0, 'application': 0, 'solicited': 0.0, 'aspect': 0.08585915387024215, 'current': 0.15069115973606037, 'wor
k': 0.24147825237028397, 'worksinprogress': 0.0, 'proposed': 0.0794894612017972, 'considered': 0.13858440610202505, 'four':
0.06473325503236776, 'type': 0.26060298447296115, 'available': 0.03799730595838356, 'one': 0.1552199676877941, 'regular': 0,
```

6. Result after 1st iteration according to the feedback provided by the user
   After that cosine similarity is calculated between the new query vector and all the documents and again top 100 new documents are retrieved.

```
Enter the p % of documents you want to mark as relevant

10
Enter the  10.0 documents you want to mark as relevant

58082,59393,59602,59518,59308,59183,59043,59115,59238,58953
docs after 1st iteration are:
 [('59183', 0.7073943843353744), ('59165', 0.4640926811771008), ('59518', 0.45181926685709317), ('58953', 0.41711859011599167),
('59602', 0.374166408623972), ('59504', 0.346282020814940 47), ('59238', 0.3039134851184073), ('59043', 0.2986162358406684), ('5
9034', 0.2840978106909882), ('59069', 0.25651374597400806), ('59332', 0.25645285196405165), ('59049', 0.2508156990566919), ('59
209', 0.24763330895978444), ('59333', 0.2420276068116615), ('59338', 0.2383473371257578), ('59115', 0.22696936889362343), ('593
18', 0.22637836487249768), ('59527', 0.22008236358046437), ('59322', 0.2186051073035 4843), ('59304', 0.2181203948800005), ('596
32', 0.21232573080982248), ('59488', 0.21207459882396443), ('59393', 0.20510298797222062), ('59592', 0.2041964881190417), ('580
82', 0.19834010003943744), ('59083', 0.1968637461158067), ('59308', 0.19660103124459152), ('59554', 0.19601855502876245), ('593
23', 0.19104882032787143), ('59063', 0.18994114688384406), ('59395', 0.18664424278009165), ('59044', 0.18528012250299608), ('59
499', 0.18492841925441647), ('59224', 0.1846895732489182), ('59470', 0.1788662306554469), ('59241', 0.173816573666747), ('5922
8', 0.173393811016756 6), ('59548', 0.17090494526274003), ('59459', 0.16732766924258055), ('59219', 0.16352493355157466), ('1790
58', 0.16008974614188 65), ('59207', 0.15957912003909805), ('59123', 0.15927681849618966), ('58578', 0.15823391565043934), ('592
37', 0.15768824405685183), ('59234', 0.1574449300862585), ('178293', 0.1556018159881379), ('59284', 0.15522568683693286), ('594
56', 0.1550523334722755), ('59225', 0.15154226188029585), ('39638', 0.1488761603825386), ('38403', 0.14874424314528456), ('3907
8', 0.1486950793152884), ('59085', 0.14774547923301318), ('59236', 0.14717159673268912), ('59023', 0.147170700510721), ('17857
1', 0.1462707784232549), ('176960', 0.14587209432397785), ('58109', 0.14456467394804987), ('178313', 0.14403998480534624), ('59
091', 0.14403273007378353), ('59189', 0.1417724723924200 8), ('58910', 0.14165189428656555), ('59206', 0.14141477168268715), ('1
78382', 0.14122746295472774), ('178786', 0.14089607349229666), ('59490', 0.14045947010291876), ('58155', 0.13951243359870355),
('59246', 0.13820619054708583), ('59637', 0.1377882868287704), ('179073', 0.13715240279276555), ('59118', 0.13596304130308062),
('59252', 0.13558032917110482), ('59286', 0.13513870398397032), ('178908', 0.1347625407056851), ('59341', 0.13375678686012019),
('58152', 0.13307294280679738), ('59179', 0.1321100757009633), ('179054', 0.13084933377988878), ('59101', 0.1308054389301554),
('178314', 0.13037065948552792), ('59255', 0.1292226498253089), ('58139', 0.12890714519137067), ('61049', 0.1282382234467681),
('176936', 0.1281371930946717), ('58872', 0.1272540811354577), ('59256', 0.1269787590152852), ('178547', 0.12612944518435637),
('59202', 0.1260892849436851), ('59093', 0.1256310596270757), ('61435', 0.12520982328611616), ('178914', 0.1250795937579 7614),
('178425', 0.12485015615599185), ('176944', 0.12181030433980922), ('58569', 0.12155728874639482), ('178672', 0.120945073668060
5), ('59125', 0.1196900842812425), ('38853', 0.1193454982871891), ('38376', 0.11926840769896864), ('178653', 0.1181320707527206
7)]
```

7. Star marked relevant documents
   The documents marked as relevant by the user is shown using '*' symbol.

```
print("docs after 1st iteration are: ",startlist1)
```

```
docs after 1st iteration are:  ['59183*', '59165', '59518*', '58953*', '59602*', '59504', '59238*', '59043*', '59034', '59069',
'59332', '59049', '59209', '59333', '59338', '59115*', '59318', '59527', '59322', '59304', '59632', '59488', '59393*', '59592',
'58082*', '59083', '59308*', '59554', '59323', '59063', '59395', '59044', '59499', '59224', '59470', '59241', '59228', '59548',
'59459', '59219', '179058', '59207', '59123', '58578', '59237', '59234', '178293', '59284', '59456', '59225', '39638', '38403',
'39078', '59085', '59236', '59023', '178571', '176960', '58109', '178313', '59091', '59189', '58910', '59206', '178382', '17878
6', '59490', '58155', '59246', '59637', '179073', '59118', '59252', '59286', '178908', '59341', '58152', '59179', '179054', '59
101', '178314', '59255', '58139', '61049', '176936', '58872', '59256', '178547', '59202', '59093', '61435', '178914', '178425',
'176944', '58569', '178672', '59125', '38853', '38376', '178653']
```

8. MAP, Precision and Recall after 1st iteration

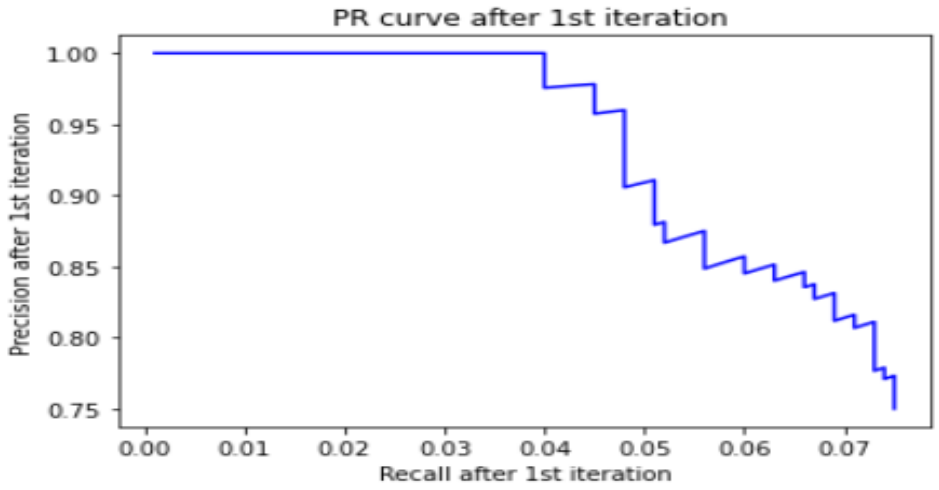
After the result obtained from 1st iteration again Map, precision and recall values are calculated for 100 points taking files of sci.med folder as ground truth. Now as the relevant document comes first so the value of precision becomes 1.0 for the starting documents. Also the value of recall increases as number of relevant retrieved document increases which makes the curve again saw tooth shape.

```
precision after 1st iteration are:  [1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0,
1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 0.97560975609756
1, 0.9761904761904762, 0.9767441860465116, 0.9772727272727273, 0.9777777777777777, 0.9782608695652174, 0.9574468085106383, 0.95
83333333333334, 0.9591836734693877, 0.96, 0.9411764705882353, 0.9230769230769231, 0.9056603773584906, 0.9074074074074074, 0.909
0909090909091, 0.9107142857142857, 0.8947368421052632, 0.8793103448275862, 0.8813559322033898, 0.8666666666666667, 0.8688524590
163934, 0.8709677419354839, 0.873015873015873, 0.875, 0.8615384615384616, 0.8484848484848485, 0.8507462686567164, 0.85294117647
05882, 0.855072463768116, 0.8571428571428571, 0.8450704225352113, 0.8472222222222222, 0.8493150684931506, 0.8513513513513513,
0.84, 0.8421052631578947, 0.8441558441558441, 0.8461538461538461, 0.8354430379746836, 0.8375, 0.8271604938271605, 0.82926829268
29268, 0.8313253012048193, 0.8214285714285714, 0.8117647058823529, 0.813953488372093, 0.8160919540229885, 0.8068181818181818,
0.8089887640449438, 0.8111111111111111, 0.8021978021978022, 0.7934782608695652, 0.7849462365591398, 0.776595744680851, 0.778947
3684210526, 0.7708333333333334, 0.7731958762886598, 0.7653061224489796, 0.7575757575757576, 0.75]
recall after 1st iteration are:  [0.001, 0.002, 0.003, 0.004, 0.005, 0.006, 0.007, 0.008, 0.009, 0.01, 0.011, 0.012, 0.013, 0.0
14, 0.015, 0.016, 0.017, 0.018, 0.019, 0.02, 0.021, 0.022, 0.023, 0.024, 0.025, 0.026, 0.027, 0.028, 0.029, 0.03, 0.031, 0.032,
0.033, 0.034, 0.035, 0.036, 0.037, 0.038, 0.039, 0.04, 0.04, 0.041, 0.042, 0.043, 0.044, 0.045, 0.045, 0.046, 0.047, 0.048, 0.0
48, 0.048, 0.048, 0.049, 0.05, 0.051, 0.051, 0.051, 0.052, 0.052, 0.053, 0.054, 0.055, 0.056, 0.056, 0.056, 0.057, 0.058, 0.05
9, 0.06, 0.06, 0.061, 0.062, 0.063, 0.063, 0.064, 0.065, 0.066, 0.066, 0.067, 0.067, 0.068, 0.069, 0.069, 0.069, 0.07, 0.071,
0.071, 0.072, 0.073, 0.073, 0.073, 0.073, 0.073, 0.074, 0.074, 0.075, 0.075, 0.075, 0.075]
MAP after 1st iteration are:  0.9420900822634712
```

9. Precision-Recall curve after iteration 1

```
import matplotlib.pyplot as plt

plt.plot(r1, p1,color="blue")
plt.xlabel("Recall after 1st iteration")
plt.ylabel("Precision after 1st iteration")
plt.title(" PR curve after 1st iteration")
```

```
Text(0.5, 1.0, ' PR curve after 1st iteration')
```



10. Updated query vector

Now as some new 'p' documents are marked as relevant, so the new query vector is again calculated on the basis of new set of relevant and non relevant documents using rocchio's algorithm with $\alpha = 1$, $\beta = 0.7$, and $\gamma = 0.25$.

```
print("New query vector after 2nd iteration is \n", new_query2)
```

```
New query vector after 2nd iteration is
 {'xref': 0.04528052002171944, 'cantaloupesrvcscmuedu': 0.04528052002171944, 'compgraphics37261': 0.0, 'altgraphics519': 0.0,
'compgraphicsanimation2614': 0.0, 'path': 0, 'cantaloupesrvcscmuedudasnewsharvardeduogicseuwmeduzaphodmpsohiostateedudarwinsu
ranetdtixdtnavymiloasyslipman': 0.0, 'lipmanoasysdtnavymil': 0.0, 'robert': 0.07892899321923826, 'lipman': 0.0, 'newsgroups':
0, 'compgraphicsaltgraphicscompgraphicsanimation': 0.0, 'subject': 0, 'call': 0.3238647835420398, 'presentation': 0, 'navy':
0, 'scivizvr': 0.0, 'seminar': 0.0, 'messageid': 0, '32850oasysdtnavymil': 0.0, 'date': 0, 'nineteen': 0.047172280016534666,
'mar': 0, 'ninety-three': 0.04819798111513432, 'two hundred and one thousand and twenty-three': 0.0, 'gmt': 0.054471824589958
535, 'articleid': 0.10956431268958783, 'oasys32850': 0.0, 'expires': 0.07710435012010143, 'thirty': 0.12475129702301879, 'ap
r': 0.01740160980431117, 'forty thousand': 0.0, 'replyto': 0.183378279311193, 'followupto': 0.14094678449617723, 'compgraphic
s': 0, 'distribution': 0.08398022422172904, 'usa': 0.020337007893297072, 'organization': 0.007049011262588228, 'carderock':
0.0, 'division': 0.019295873593980818, 'nswc': 0.0, 'bethesda': 0, 'md': 0.3590794397969428, 'line': 0.0008909299568674266,
'sixty-five': 0, 'scientific': 0.1727801077133365, 'visualization': 0, 'virtual': 0, 'reality': 0.03405304471468503, 'tuesda
y': 0, 'june': 0.07330458147406799, 'twenty-two': 0.11806057682445867, 'one thousand, nine hundred and ninety-three': 0.07465
285391559609, 'naval': 0, 'surface': 0.2747541493307064, 'warfare': 0.0, 'center': 0, 'formerly': 0, 'david': 0.0468015262111
445, 'taylor': 0.0, 'research': 0.05682035566227028, 'maryland': 0, 'sponsor': 0, 'ness': 0.0, 'engineering': 0, 'software':
0, 'system': 0.38750647903535906, 'sponsoring': 0, 'oneday': 0.0, 'purpose': 0, 'present': 0.17009991486887566, 'exchange':
0, 'information': 0.18366443786790804, 'navyrelated': 0.0, 'program': 0.04096479092700234, 'development': 0, 'application':
0, 'solicited': 0.0, 'aspect': 0.1442359999754427, 'current': 0.28076827008176086, 'work': 0.2494414810821059, 'worksinprogre
ss': 0.0, 'proposed': 0.04694872916494278, 'considered': 0.2392696494627325, 'four': 0.11177910505222823, 'type': 0.558988871
8005894, 'available': 0.2119765274399125
```

11. Result after 2nd iteration according to the feedback provided by the user

After that cosine similarity is calculated between the new query vector and all the documents and again top 100 new documents are retrieved.

```
Enter the  10.0 documents you want to mark as relevant

59165,59049,59322,59323,59228,58578,59490,59252,58139,59202
docs after 2nd iteration are:  [('59183', 0.6679373251727939), ('59165', 0.6095385605025277), ('59323', 0.4807562962180958),
('59333', 0.43279239536329217), ('58578', 0.4288256933902286), ('59322', 0.42016148092101424), ('59518', 0.400508157178036),
('59049', 0.3614123461996056), ('59034', 0.35603125020664883), ('59069', 0.3556072607256993), ('59304', 0.34970167561514226),
('59332', 0.3385013709311748), ('59504', 0.3143535574819597), ('59490', 0.3045945782453576), ('59318', 0.30458177736070957),
('59602', 0.29203955675877513), ('59338', 0.2918860875470126), ('58953', 0.28922427220984764), ('59241', 0.26666337644421473),
('59209', 0.26636875103294955), ('59488', 0.2637870136139695), ('59202', 0.26184829958201916), ('59228', 0.2616855455832438),
('59252', 0.2610286837412677), ('59395', 0.2598283369078862), ('59554', 0.25329326652642015), ('59527', 0.24858010128378738),
('59207', 0.24655834192512652), ('59499', 0.2416223655665585), ('59632', 0.23579121890120433), ('59237', 0.2242847953095546),
('58139', 0.2210216031863608), ('59044', 0.21955663284693808), ('59447', 0.21835132798227158), ('59548', 0.2055536621841373),
('59225', 0.20019763419934586), ('59470', 0.19794762790635972), ('59168', 0.19755418564104088), ('59179', 0.19580819025005747),
('59459', 0.19474051813042223), ('59296', 0.19454709096102002), ('59238', 0.19337764336408178), ('59043', 0.1903288041424598),
('59341', 0.18875924228231958), ('59284', 0.18846805943045983), ('59224', 0.18582563892448217), ('59123', 0.18481525243862587),
('59236', 0.18157919372578232), ('59101', 0.18067628142174455), ('59219', 0.17170546152566185), ('59023', 0.16960998355885937),
('59247', 0.1668477643494089), ('59330', 0.16472906759638822), ('59115', 0.16307945004089106), ('59234', 0.1630477567396802),
('59393', 0.16252031178242776), ('59246', 0.1621022887676193), ('59286', 0.16056551308012387), ('59206', 0.15991466361276885),
('58569', 0.15760129444203808), ('59085', 0.15674505846448175), ('59456', 0.1556408653965737), ('59118', 0.1540166350713795),
('58568', 0.15179138607882675), ('59091', 0.15112637956731334), ('58152', 0.15018020053242231), ('179058', 0.1482162308182960
2), ('58155', 0.14820211429159566), ('59216', 0.14771473403418242), ('58109', 0.1464793038436459), ('59255', 0.1462781364068963
8), ('59592', 0.14564514597149703), ('58910', 0.14536792778635257), ('59181', 0.1445253253910846), ('59203', 0.1441244213775994
3), ('59125', 0.1438611411125559), ('59283', 0.14341749943624119), ('59093', 0.14275417810038962), ('178293', 0.142453300392299
84), ('58984', 0.14161799904906267), ('59083', 0.1413546648864817), ('178786', 0.14080708568333067), ('39638', 0.13945481844167
23), ('39078', 0.13928999690867075), ('38403', 0.13911575502979237), ('59256', 0.13881253644323793), ('58577', 0.13849461904783
505), ('59189', 0.1372088928051047), ('59160', 0.13716821776013313), ('59432', 0.1366834234310456), ('59373', 0.135274692470526
95), ('59347', 0.1349162332121063), ('58882', 0.13482248269615524), ('59222', 0.13441574392255382), ('59122', 0.134317044431683
16), ('178382', 0.13361949299932738), ('58082', 0.133451805117361), ('178571', 0.1332686193104051), ('59637', 0.132839783688415
73), ('59226', 0.13190509874972606)]
```

12. Star marked relevant documents

    The documents marked as relevant in 2nd iteration by the user is shown using '*' symbol.

```
print("docs after 2nd iteration are: ",startlist2)
```

```
docs after 2nd iteration are:  ['59183', '59165*', '59323*', '59333', '58578*', '59322*', '59518', '59049*', '59034', '59069',
'59304', '59332', '59504', '59490*', '59318', '59602', '59338', '58953', '59241', '59209', '59488', '59202*', '59228*', '59252
*', '59395', '59554', '59527', '59207', '59499', '59632', '59237', '58139*', '59044', '59447', '59548', '59225', '59470', '5916
8', '59179', '59459', '59296', '59238', '59043', '59341', '59284', '59224', '59123', '59236', '59101', '59219', '59023', '5924
7', '59330', '59115', '59234', '59393', '59246', '59286', '59206', '58569', '59085', '59456', '59118', '58568', '59091', '5815
2', '179058', '58155', '59216', '58109', '59255', '59592', '58910', '59181', '59203', '59125', '59283', '59093', '178293', '589
84', '59083', '178786', '39638', '39078', '38403', '59256', '58577', '59189', '59160', '59432', '59373', '59347', '58882', '592
22', '59122', '178382', '58082', '178571', '59637', '59226']
```

13. MAP, Precision and Recall after 2nd iteration

    After the result obtained from 2nd iteration again Map, precision and recall values are calculated for 100 points taking files of sci.med folder as ground truth. Now as the relevant document comes first so the value of precision becomes 1.0 for the starting documents. Also the value of recall increases as number of relevant retrieved document increases which makes the curve again saw tooth shape.

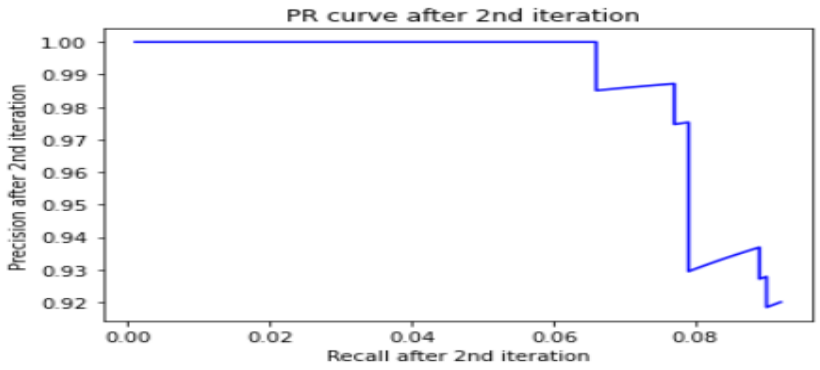
```
precision after 2nd iteration are:  [1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0,
1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.
0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 0.985074626865
6716, 0.9852941176470589, 0.9855072463768116, 0.9857142857142858, 0.9859154929577465, 0.9861111111111112, 0.9863013698630136,
0.9864864864864865, 0.9866666666666667, 0.9868421052631579, 0.987012987012987, 0.9871794871794872, 0.9746835443037974, 0.975,
0.9753086419753086, 0.9634146341463414, 0.9518072289156626, 0.9404761904761905, 0.9294117647058824, 0.9302325581395349, 0.93103
44827586207, 0.9318181818181818, 0.9325842696629213, 0.9333333333333333, 0.9340659340659341, 0.9347826086956522, 0.935483870967
7419, 0.9361702127659575, 0.9368421052631579, 0.9270833333333334, 0.9278350515463918, 0.9183673469387755, 0.9191919191919192,
0.92]


recall after 2nd iteration are:  [0.001, 0.002, 0.003, 0.004, 0.005, 0.006, 0.007, 0.008, 0.009, 0.01, 0.011, 0.012, 0.013, 0.0
14, 0.015, 0.016, 0.017, 0.018, 0.019, 0.02, 0.021, 0.022, 0.023, 0.024, 0.025, 0.026, 0.027, 0.028, 0.029, 0.03, 0.031, 0.032,
0.033, 0.034, 0.035, 0.036, 0.037, 0.038, 0.039, 0.04, 0.041, 0.042, 0.043, 0.044, 0.045, 0.046, 0.047, 0.048, 0.049, 0.05, 0.0
51, 0.052, 0.053, 0.054, 0.055, 0.056, 0.057, 0.058, 0.059, 0.06, 0.061, 0.062, 0.063, 0.064, 0.065, 0.066, 0.066, 0.067, 0.06
8, 0.069, 0.07, 0.071, 0.072, 0.073, 0.074, 0.075, 0.076, 0.077, 0.077, 0.078, 0.079, 0.079, 0.079, 0.079, 0.079, 0.08, 0.081,
0.082, 0.083, 0.084, 0.085, 0.086, 0.087, 0.088, 0.089, 0.089, 0.09, 0.09, 0.091, 0.092]


MAP after 2nd iteration is:  0.9880729839832986
```

14. Precision-Recall curve after iteration 2

    As now relevant documents come at higher position due to change in query vector the precision values are coming 1.0 for the starting dcouments.

```
plt.plot(r2, p2,color="blue")
plt.xlabel("Recall after 2nd iteration")
plt.ylabel("Precision after 2nd iteration")
plt.title(" PR curve after 2nd iteration")
```

```
Text(0.5, 1.0, ' PR curve after 2nd iteration')
```



15. Updated query vector

    Now as some new ‘p’ documents are marked as relevant, so the new query vector is again calculated on the basis of new set of relevant and non relevant documents using rocchio’s algorithm with α= 1, β= 0.7,andγ=0.25.

```
print("New query vector after 3rd iteration is \n", new_query3)
```

```
New query vector after 3rd iteration is
 {'xref': 0.06920721021982497, 'cantaloupesrvcscmuedu': 0.06920721021982497, 'compgraphics37261': 0.0, 'altgraphics519': 0.0,
'compgraphicsanimation2614': 0.0, 'path': 0, 'cantaloupesrvcscmuedudasnewsharvardeduogicseuwmeduzaphodmpsohiostateedudarwinsu
ranetdtixdtnavymiloasyslipman': 0.0, 'lipmanoasysdtnavymil': 0.0, 'robert': 0.11137020691314832, 'lipman': 0.0, 'newsgroups':
0, 'compgraphicsaltgraphicscompgraphicsanimation': 0.0, 'subject': 0, 'call': 0.5537260185959765, 'presentation': 0, 'navy':
0, 'scivizvr': 0.0, 'seminar': 0.0, 'messageid': 0, '32850oasysdtnavymil': 0.0, 'date': 0, 'nineteen': 0.06625538624672232,
'mar': 0.0, 'ninety-three': 0.06926727624312393, 'two hundred and one thousand and twenty-three': 0.0, 'gmt': 0.0796991618785
1468, 'articleid': 0.10330339544197253, 'oasys32850': 0.0, 'expires': 0.1472279096058054, 'thirty': 0.21461276342122432, 'ap
r': 0.02658291853178139, 'forty thousand': 0.0, 'replyto': 0.23096660685374026, 'followupto': 0.2306063823360151, 'compgraphi
cs': 0.0, 'distribution': 0.0872373376983957, 'usa': 0.007282988015746963, 'organization': 0.01066156797756472, 'carderock':
0.0, 'division': 0.119836018451739, 'nswc': 0.0, 'bethesda': 0, 'md': 0.7444740870051526, 'line': 0.0014121103128938571, 'six
ty-five': 0, 'scientific': 0.35935825135242494, 'visualization': 0, 'virtual': 0.0, 'reality': 0.012758324875928772, 'tuesda
y': 0, 'june': 0.055267282833514744, 'twenty-two': 0.13208116193781994, 'one thousand, nine hundred and ninety-three': 0.1428
589313903179, 'naval': 0.0, 'surface': 0.49257037928537134, 'warfare': 0.0, 'center': 0.03168220641939962, 'formerly': 0.1770
1581208884284, 'david': 0.4360932487528893, 'taylor': 0.0, 'research': 0.27395093831764405, 'maryland': 0.11436092823280669,
'sponsor': 0, 'ness': 0.0, 'engineering': 0, 'software': 0.14075005759548354, 'system': 0.7299070315239403, 'sponsoring': 0.
0, 'oneday': 0.0, 'purpose': 0.024455501286081635, 'present': 0.37273142953297833, 'exchange': 0, 'information': 0.5097347852
110518, 'navyrelated': 0.0, 'program': 0.22325876096225333, 'development': 0.22041408426091894, 'application': 0, 'solicite
d': 0.0, 'aspect': 0.1957770106065863, 'current': 0.3665590055705038, 'work': 0.5671362069735111, 'worksinprogress': 0.0, 'pr
```

16. Result after 3rd iteration according to the feedback provided by the user

After that cosine similarity is calculated between the new query vector and all the documents and again top 100 new documents are retrieved.

```
Enter the  10.0 documents you want to mark as relevant

59183,59241,59395,59499,59225,59238,59123,59101,59330,58569
docs after 3rd iteration are:  [('59183', 0.7304362234822233), ('59165', 0.597811593072471), ('59323', 0.4240815697568185), ('5
9333', 0.4112901177074424), ('59123', 0.4002030599859667), ('59322', 0.3864601979944167), ('59034', 0.37211991089146884), ('585
78', 0.3663312522355849), ('59069', 0.3615305605287441), ('59518', 0.36041748193452394), ('59049', 0.3602648224098391), ('5949
9', 0.35768019262954015), ('59332', 0.35694423931830094), ('59304', 0.3564262665874207), ('59395', 0.3371050236505117), ('5924
1', 0.32951038140301925), ('59318', 0.3191607985972803), ('59338', 0.30974841381183765), ('58569', 0.2990149087604488), ('5949
0', 0.2856028477967745), ('59225', 0.2839513915928892), ('59504', 0.28350438267213496), ('59252', 0.28243494109070794), ('5948
8', 0.2809486427391374), ('59209', 0.2808923379612597), ('59554', 0.2790991521665557), ('59228', 0.266424223104792), ('59602',
0.2658654210191151), ('59527', 0.2618484812759439), ('59632', 0.25976873004871326), ('59237', 0.2581877132340271), ('59330', 0.
24470891839467107), ('58953', 0.24370378221325018), ('59044', 0.24223992572941308), ('59101', 0.24211603502315218), ('59207',
0.2395245562721406), ('59284', 0.2353229358858084), ('59202', 0.22365812898519533), ('59286', 0.22161859782986376), ('59548',
0.2212237951187531), ('59236', 0.2166308794808842), ('58139', 0.21330283580138346), ('59470', 0.2130287349627144), ('59296', 0.
21283087696164144), ('59459', 0.20989500836687988), ('59238', 0.20975324212810803), ('59023', 0.19659777781858948), ('59447',
0.19442902218078456), ('59224', 0.1916779166949318), ('59283', 0.1901796002430219), ('59216', 0.18987271709002254), ('59341',
0.18968697197395948), ('59373', 0.18953673661960924), ('59043', 0.18716853876864994), ('59219', 0.18428746511730457), ('59247',
0.1820928258507293), ('59126', 0.18206458185823252), ('59179', 0.18166597953588443), ('59125', 0.18005235152776705), ('179058',
0.17819774011559011), ('178293', 0.17433134783108684), ('59206', 0.17315170115805098), ('59168', 0.1721039982922258), ('59246',
0.17099274861046337), ('59122', 0.17083067919558043), ('59347', 0.16896012669296684), ('58152', 0.1683349245544477), ('176960',
0.16548501044715044), ('59456', 0.16487464192528362), ('178786', 0.16393152893149907), ('58910', 0.16369870373228598), ('5943
5', 0.16361815239115318), ('58568', 0.16295015031605592), ('59234', 0.16171118094127304), ('59118', 0.16077508121640327), ('384
03', 0.15951930092121863), ('39638', 0.15949228831792944), ('39078', 0.15933360385201953), ('178571', 0.1584888205794827), ('58
155', 0.15772700422973296), ('59085', 0.15643286028429984), ('59091', 0.15601290396383016), ('178313', 0.15433169728592136),
('58984', 0.154123247191865), ('59181', 0.15386925354953992), ('178908', 0.15202105119208154), ('59285', 0.15193751569680888),
('59294', 0.1517168782303989), ('59093', 0.15080986888266915), ('58109', 0.15080543998088333), ('59572', 0.15055830970568107),
('178382', 0.15053054356311665), ('179073', 0.15044100695626617), ('58882', 0.14958662871278855), ('59203', 0.148257866824292
2), ('59256', 0.14808632688286003), ('59160', 0.14775068730846416), ('59637', 0.14759260196050228), ('178425', 0.14590080262773
14), ('58897', 0.1450559839739221)]
```

17. Star marked relevant documents

The documents marked as relevant in 3rd iteration by the user is shown using '*' symbol.

```
print("docs after 3rd iteration are: ",startlist3)
```

```
docs after 3rd iteration are:  ['59183*', '59165', '59323', '59333', '59123*', '59322', '59034', '58578', '59069', '59518', '59
049', '59499*', '59332', '59304', '59395*', '59241*', '59318', '59338', '58569*', '59490', '59225*', '59504', '59252', '59488',
'59209', '59554', '59228', '59602', '59527', '59632', '59237', '59330*', '58953', '59044', '59101*', '59207', '59284', '59202',
'59286', '59548', '59236', '58139', '59470', '59296', '59459', '59238*', '59023', '59447', '59224', '59283', '59216', '59341',
'59373', '59043', '59219', '59247', '59126', '59179', '59125', '179058', '178293', '59206', '59168', '59246', '59122', '59347',
'58152', '176960', '59456', '178786', '58910', '59435', '58568', '59234', '59118', '38403', '39638', '39078', '178571', '5815
5', '59085', '59091', '178313', '58984', '59181', '178908', '59285', '59294', '59093', '58109', '59572', '178382', '179073', '5
8882', '59203', '59256', '59160', '59637', '178425', '58897']
```

18. MAP, Precision and Recall after 3rd iteration

As now relevant documents come at higher position due to change in query vector the precision values are coming 1.0 for the starting dcouments.

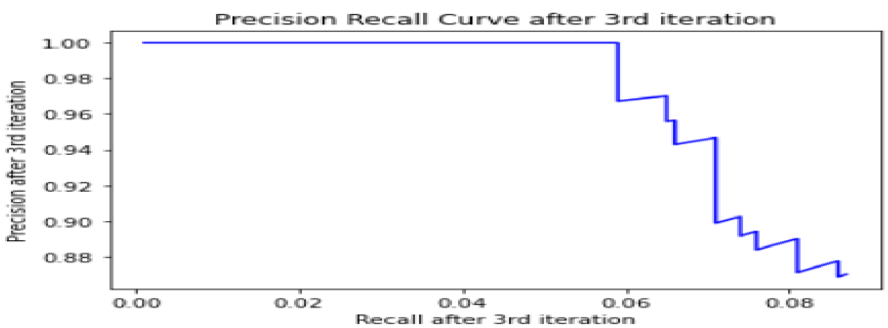
```
precision after 3rd iteration are:  [1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0,
1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.
0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 0.9833333333333333, 0.9672131147540983, 0.9677419
35483871, 0.9682539682539683, 0.96875, 0.9692307692307692, 0.9696969696969697, 0.9701492537313433, 0.9558823529411765, 0.956521
7391304348, 0.9428571428571428, 0.9436619718309859, 0.9444444444444444, 0.9452054794520548, 0.9459459459459459, 0.9466666666666
667, 0.9342105263157895, 0.922077922077922, 0.9102564102564102, 0.8987341772151899, 0.9, 0.9012345679012346, 0.902439024390243
9, 0.891566265060241, 0.8928571428571429, 0.8941176470588236, 0.8837209302325582, 0.8850574712643678, 0.8863636363636364, 0.887
6404494382022, 0.8888888888888888, 0.8901098901098901, 0.8804347826086957, 0.8709677419354839, 0.8723404255319149, 0.8736842105
263158, 0.875, 0.8762886597938144, 0.8775510204081632, 0.8686868686868687, 0.87]


recall after 3rd iteration are:  [0.001, 0.002, 0.003, 0.004, 0.005, 0.006, 0.007, 0.008, 0.009, 0.01, 0.011, 0.012, 0.013, 0.0
14, 0.015, 0.016, 0.017, 0.018, 0.019, 0.02, 0.021, 0.022, 0.023, 0.024, 0.025, 0.026, 0.027, 0.028, 0.029, 0.03, 0.031, 0.032,
0.033, 0.034, 0.035, 0.036, 0.037, 0.038, 0.039, 0.04, 0.041, 0.042, 0.043, 0.044, 0.045, 0.046, 0.047, 0.048, 0.049, 0.05, 0.0
51, 0.052, 0.053, 0.054, 0.055, 0.056, 0.057, 0.058, 0.059, 0.059, 0.059, 0.06, 0.061, 0.062, 0.063, 0.064, 0.065, 0.065, 0.06
6, 0.066, 0.067, 0.068, 0.069, 0.07, 0.071, 0.071, 0.071, 0.071, 0.071, 0.072, 0.073, 0.074, 0.074, 0.075, 0.076, 0.076, 0.077,
0.078, 0.079, 0.08, 0.081, 0.081, 0.081, 0.082, 0.083, 0.084, 0.085, 0.086, 0.086, 0.087]


MAP after 3rd iteration is:  0.97321657676322
```

19. Precision-Recall curve after iteration 3

    As now relevant documents come at higher position due to change in query vector the precision values are coming 1.0 for the starting documents.

```python
plt.plot(r3, p3,color="blue")
plt.xlabel("Recall after 3rd iteration")
plt.ylabel("Precision after 3rd iteration")
plt.title("Precision Recall Curve after 3rd iteration")
```

```
Text(0.5, 1.0, 'Precision Recall Curve after 3rd iteration')
```



20. Updated query vector

    Now as some new ‘p’ documents are marked as relevant, so the new query vector is again calculated on the basis of new set of relevant and non relevant documents using rocchio’s algorithm with α= 1, β= 0.7,andγ=0.25.

```python
print("New query vector after 4th iteration is \n", new_query4)
```

New query vector after 4th iteration is
 {'xref': 0.1194571733636765, 'cantaloupesrvcscmuedu': 0.1194571733636765, 'compgraphics37261': 0.0, 'altgraphics519': 0.0,
'compgraphicsanimation2614': 0.0, 'path': 0, 'cantaloupesrvcscmuedudasnewsharvardeduogicseuwmeduzaphodmpsohiostateedudarwinsu
ranetdtixdtnavymiloasyslipman': 0.0, 'lipmanoasysdtnavymil': 0.0, 'robert': 0.14789677740406817, 'lipman': 0.0, 'newsgroups':
0, 'compgraphicsaltgraphicscompgraphicsanimation': 0.0, 'subject': 0, 'call': 0.7221492759954898, 'presentation': 0, 'navy':
0, 'scivizvr': 0.0, 'seminar': 0.0, 'messageid': 0, '32850oasysdtnavymil': 0.0, 'date': 0, 'nineteen': 0.12867929729803024,
'mar': 0.0, 'ninety-three': 0.10562272395070511, 'two hundred and one thousand and twenty-three': 0.0, 'gmt': 0.1032246654190
5487, 'articleid': 0.16127619393528622, 'oasys32850': 0.0, 'expires': 0.12390535966022118, 'thirty': 0.19466348547415385, 'ap
r': 0.03432813221514691, 'forty thousand': 0.0, 'replyto': 0.3561175057317639, 'followupto': 0.27062827286876123, 'compgraphi
cs': 0.0, 'distribution': 0.11716616552378661, 'usa': 0, 'organization': 0.013873358150206808, 'carderock': 0.0, 'division':
0.15739143301342529, 'nswc': 0.0, 'bethesda': 0, 'md': 1.0088005436722338, 'line': 0.0018098947480576354, 'sixty-five': 0, 's
cientific': 0.39047390532053033, 'visualization': 0, 'virtual': 0.0, 'reality': 0, 'tuesday': 0, 'june': 0.03794198282350965
5, 'twenty-two': 0.09623427164137875, 'one thousand, nine hundred and ninety-three': 0.1628546804829353, 'naval': 0.0, 'surfa
ce': 0.8365522074334222, 'warfare': 0.0, 'center': 0, 'formerly': 0.16826949800333985, 'david': 0.3208810546730957, 'taylor':
0.0, 'research': 0.2382698899208109, 'maryland': 0.07959219336566573, 'sponsor': 0, 'ness': 0.0, 'engineering': 0, 'softwar
e': 0.0465234604965854, 'system': 0.8484830663587316, 'sponsoring': 0, 'oneday': 0.0, 'purpose': 0, 'present': 0.377316513676
52597, 'exchange': 0, 'information': 0.4838154001987306, 'navyrelated': 0.0, 'program': 0.12423673926402179, 'development':
0.15295337767937123, 'application': 0, 'solicited': 0.0, 'aspect': 0.24463795011228825, 'current': 0.44090207617878724, 'wor
k': 0.7096463474634847, 'worksinprogress': 0.0, 'proposed': 0.07249689638115353, 'considered': 0.6172059118543464, 'four': 0.

21. Result after 4th iteration according to the feedback provided by the user

After that cosine similarity is calculated between the new query vector and all the documents and again top 100 new documents are retrieved.

```
Enter the  10.0 documents you want to mark as relevant

59632,59207,59219,58139,59341,58910,59118,58155,59294,58882
docs after 4th iteration are:  [('59183', 0.7678861371430928), ('59165', 0.5959402497037753), ('59518', 0.40475376529672585),
('59323', 0.3773381041082927), ('59333', 0.37542837447971555), ('59034', 0.3570386818658101), ('59322', 0.3502671362231863),
('59069', 0.3421312289992769), ('59332', 0.3415146169607874), ('59049', 0.340998056892826), ('59123', 0.33386949377414843), ('5
9304', 0.3260622446249408), ('58578', 0.3226040311028339), ('59499', 0.3213614311125904), ('59504', 0.3128651521033456), ('5960
2', 0.3065658454653925), ('58953', 0.3054251558064399), ('59395', 0.3041293437677281), ('59318', 0.3038193257980682), ('59241',
0.2976191244679461), ('59338', 0.2940665737534391), ('59209', 0.28232428422052974), ('59488', 0.2618432443601789), ('59554', 0.
25955211962237346), ('58569', 0.2579057584841937), ('59527', 0.254240624317299), ('59225', 0.2525840320699595), ('59238', 0.252
0944965293195), ('59490', 0.2517703125983039), ('59252', 0.25155614708905993), ('59632', 0.2464625566336047), ('59228', 0.24521
58749601986 7), ('59237', 0.23679068979941403), ('59043', 0.23241504230558355), ('59044', 0.2302716851714681), ('59207', 0.22039
627656603064), ('59330', 0.21795979709788552), ('59101', 0.21589016642275702), ('59548', 0.21096368903065407), ('59470', 0.2086
898038795494), ('59284', 0.2057738389443764), ('59459', 0.20195355681297553), ('59236', 0.200951186829698), ('59202', 0.2002183
4761486715), ('58139', 0.19523985430643387), ('59296', 0.19278921431663235), ('59224', 0.1912822991405874), ('59286', 0.1882194
2386595722), ('59023', 0.18589558550789406), ('59219', 0.18228059158238463), ('59341', 0.1799542315777642), ('59179', 0.1709625
8799728786), ('59216', 0.16797028952915288), ('59115', 0.16707033457513387), ('59247', 0.1664022734181119), ('59447', 0.1662237
0788003052), ('59246', 0.16579498060840703), ('59373', 0.1636583298061507), ('59456', 0.16326832463053947), ('179058', 0.162942
6300316265), ('59206', 0.1621573114088053), ('59393', 0.161152707675196), ('178293', 0.1604913058159162), ('59283', 0.160111356
8060542), ('59234', 0.1589121120650012), ('59063', 0.15720725615179043), ('59592', 0.15701392821347532), ('58152', 0.1557118409
004822), ('58910', 0.15564538398833228), ('59118', 0.15541357791941413), ('59126', 0.15481134814232977), ('59125', 0.1541748161
301494), ('178786', 0.15326246398982712), ('59085', 0.15284129047221784), ('59091', 0.15249912876114352), ('59168', 0.152353429
44917638), ('176960', 0.15103195410213566), ('58155', 0.15058540798949188), ('58109', 0.15011992602072957), ('59347', 0.1488205
044689222), ('178571', 0.14638780021996955), ('58568', 0.1461327064343014), ('59083', 0.14611496512188218), ('39638', 0.1458063
5733724637), ('38403', 0.14570187641365823), ('39078', 0.1456127210444248), ('59122', 0.1453031448177057), ('59435', 0.14507252
5127379), ('59093', 0.1433725183222504), ('59181', 0.14285838581033067), ('59572', 0.1427161475945208), ('59637', 0.14207094549
309), ('58984', 0.14194738579509558), ('178382', 0.14125638896105286), ('178313', 0.14124251554474085), ('59256', 0.14110214480
69807), ('178908', 0.13912359293795418), ('59189', 0.13836091408473772), ('59255', 0.13835414510790658), ('179073', 0.137974829
82175346)]
```

22. Star marked relevant documents

The documents marked as relevant in 4th iteration by the user is shown using '*' symbol.

```
print("docs after 4th iteration are: ",startlist4)
```

```
docs after 4th iteration are:  ['59183*', '59165', '59518*', '59323', '59333', '59034', '59322', '59069', '59332', '59049', '59
123', '59304', '58578', '59499', '59504', '59602*', '58953*', '59395', '59318', '59241', '59338', '59209', '59488', '59554', '5
8569', '59527', '59225', '59238*', '59490', '59252', '59632', '59228', '59237', '59043*', '59044', '59207', '59330', '59101',
'59548', '59470', '59284', '59459', '59236', '59202', '58139', '59296', '59224', '59286', '59023', '59219', '59341', '59179',
'59216', '59115*', '59247', '59447', '59246', '59373', '59456', '179058', '59206', '59393*', '178293', '59283', '59234', '5906
3', '59592', '58152', '58910', '59118', '59126', '59125', '178786', '59085', '59091', '59168', '176960', '58155', '58109', '593
47', '178571', '58568', '59083', '39638', '38403', '39078', '59122', '59435', '59093', '59181', '59572', '59637', '58984', '178
382', '178313', '59256', '178908', '59189', '59255', '179073']
```

23. Final result with all the marked relevant documents by user (docs are marked with '*')

```
print("Relevant documents after 4 iterations are: \n ",final_star)
```

```
Relevant documents after 4 iterations are:
  ['59183*', '59165*', '59518*', '59323*', '59333', '59034', '59322*', '59069', '59332', '59049*', '59123*', '59304', '58578*',
'59499*', '59504', '59602*', '58953*', '59395*', '59318', '59241*', '59338', '59209', '59488', '59554', '58569*', '59527', '592
25*', '59238*', '59490*', '59252*', '59632', '59228*', '59237', '59043*', '59044', '59207', '59330*', '59101*', '59548', '5947
0', '59284', '59459', '59236', '59202*', '58139*', '59296', '59224', '59286', '59023', '59219', '59341', '59179', '59216', '591
15*', '59247', '59447', '59246', '59373', '59456', '179058', '59206', '59393*', '178293', '59283', '59234', '59063', '59592',
'58152', '58910', '59118', '59126', '59125', '178786', '59085', '59091', '59168', '176960', '58155', '58109', '59347', '17857
1', '58568', '59083', '39638', '38403', '39078', '59122', '59435', '59093', '59181', '59572', '59637', '58984', '178382', '1783
13', '59256', '178908', '59189', '59255', '179073']
```

24. MAP, Precision and Recall after 4th iteration

As now relevant documents comes at higher position due to change in query vector the precision values are coming 1.0 for the starting documents.

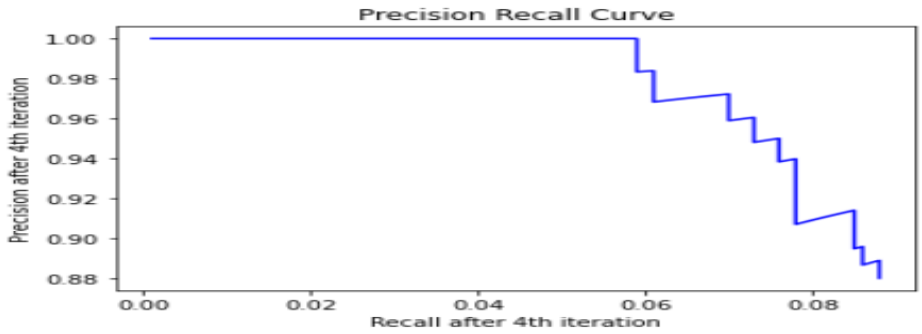
```
precision after 4th iteration are:  [1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0,
1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.
0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 0.9833333333333333, 0.9836065573770492, 0.9838709
677419355, 0.9682539682539683, 0.96875, 0.9692307692307692, 0.9696969696969697, 0.9701492537313433, 0.9705882352941176, 0.97101
44927536232, 0.9714285714285714, 0.97183098591549, 0.9722222222222222, 0.958904109589041, 0.9594594594594594, 0.96, 0.96052631
57894737, 0.948051948051948, 0.9487179487179487, 0.9493670886075949, 0.95, 0.9382716049382716, 0.9390243902439024, 0.9397590361
445783, 0.9285714285714286, 0.9176470588235294, 0.9069767441860465, 0.9080459770114943, 0.9090909090909091, 0.9101123595505618,
0.9111111111111111, 0.9120879120879121, 0.9130434782608695, 0.9139784946236559, 0.9042553191489362, 0.8947368421052632, 0.89583
33333333334, 0.8865979381443299, 0.887755102040816, 0.8888888888888888, 0.88]


recall after 4th iteration are:  [0.001, 0.002, 0.003, 0.004, 0.005, 0.006, 0.007, 0.008, 0.009, 0.01, 0.011, 0.012, 0.013, 0.0
14, 0.015, 0.016, 0.017, 0.018, 0.019, 0.02, 0.021, 0.022, 0.023, 0.024, 0.025, 0.026, 0.027, 0.028, 0.029, 0.03, 0.031, 0.032,
0.033, 0.034, 0.035, 0.036, 0.037, 0.038, 0.039, 0.04, 0.041, 0.042, 0.043, 0.044, 0.045, 0.046, 0.047, 0.048, 0.049, 0.05, 0.0
51, 0.052, 0.053, 0.054, 0.055, 0.056, 0.057, 0.058, 0.059, 0.059, 0.06, 0.061, 0.061, 0.062, 0.063, 0.064, 0.065, 0.066, 0.06
7, 0.068, 0.069, 0.07, 0.07, 0.071, 0.072, 0.073, 0.073, 0.074, 0.075, 0.076, 0.076, 0.077, 0.078, 0.078, 0.078, 0.078, 0.079,
0.08, 0.081, 0.082, 0.083, 0.084, 0.085, 0.085, 0.085, 0.086, 0.086, 0.087, 0.088, 0.088]


MAP after 4th iteration is:  0.9813544412540296
```

25. Precision-Recall curve after iteration 4

    As now relevant documents comes at higher position due to change in query vector the precision values are coming 1.0 for the starting documents.

```
plt.plot(r4, p4,color="blue")
plt.xlabel("Recall after 4th iteration")
plt.ylabel("Precision after 4th iteration")
plt.title("Precision Recall Curve")
```

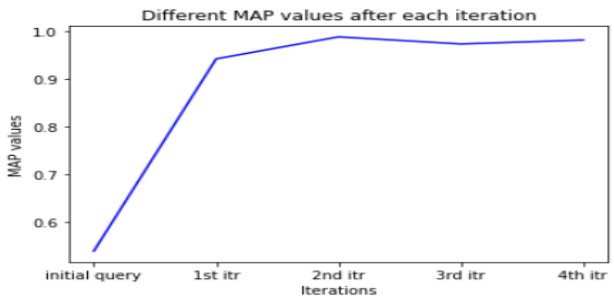
```
Text(0.5, 1.0, 'Precision Recall Curve')
```



26. MAP at the end of 4 iterations

    As the precision is increasing after each feedback provided iteration, so the mean average precision values are also increasing after each iteration. It means that the results are becoming more accurate after each iteration.

```
[0.539968246608691, 0.9420900822634712, 0.9880729839832986, 0.97321657676322, 0.9813544412540296]
Text(0.5, 1.0, 'Different MAP values after each iteration')
```

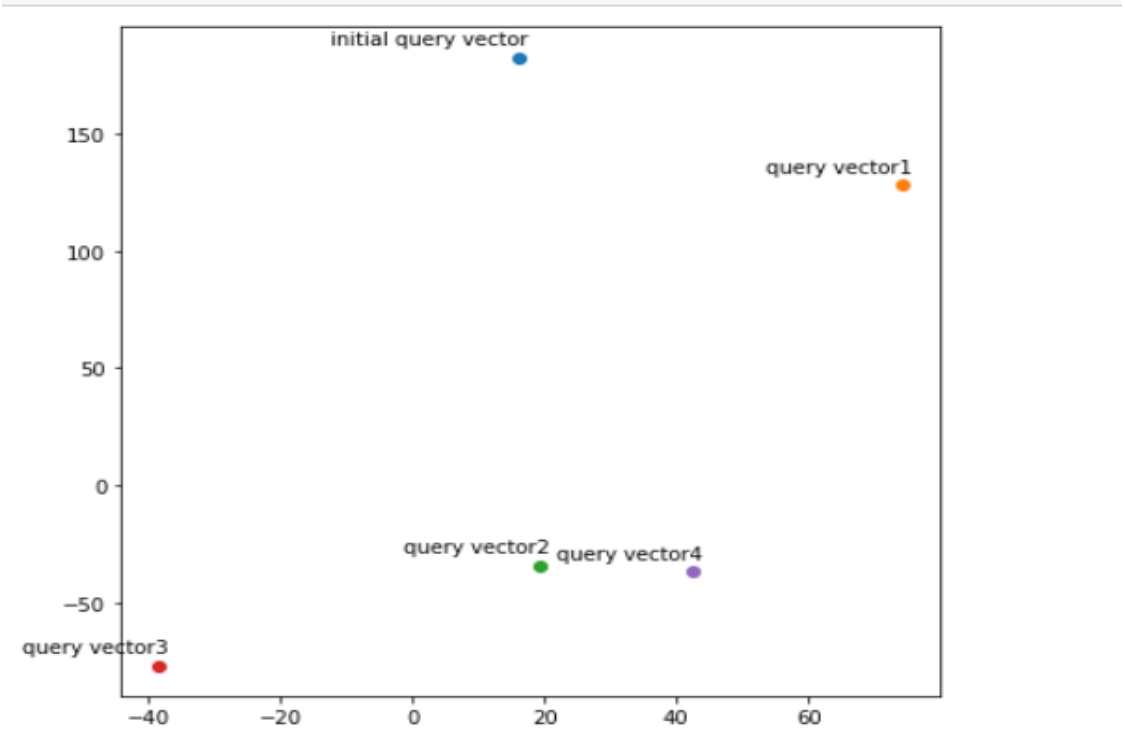


27. T-SNE plot for all the query vectors

In Rocchio's algorithm each time a new representation of information need is computed on the basis of feedback given by the user. Here the user marks the document as relevant or non-relevant. This increases the recall of the overall system.

The key concept in Rocchio's algorithm is centroid. **This algorithm moves the new query towards the relevant documents and away from the non-relevant documents**. Negative weights are given 0 importance in this algorithm.

As it is not feasible for us to analyze multidimensional data, we have used t-SNE which is a dimensionality reduction technique used for proper visualization of the given data.
So here we cannot visualize our query vectors as they have dimension of size vocabulary. Hence t-SNE is used below where each point corresponds to a query vector. We can see that initial query vector belongs to the higher value in the graph. Query1 comes closer to the centroid of relevant document and again all the other query vectors are coming closer to each other and at lower position which shows that the centroid of the relevant document must belong to lower position in the graph.
As the query vector 4 is the most refined query, so we can say that the centroid of the relevant documents has to be some where near purple point which belongs to query vector4.



Query 2: Scientific tools for preserving rights and body
Relevant set 2: Documents inside folder talk.politics.misc

1. Initial query vector

```
Initial query vector is :
 {'xref': 0, 'cantaloupesrvcscmuedu': 0, 'compgraphics37261': 0, 'altgraphics519': 0, 'compgraphicsanimation2614': 0, 'path':
0, 'cantaloupesrvcscmuedudasnewsharvardeduogicseuwmeduzaphodmpsohiostateedudarwinsuranetdtixdtnavymiloasyslipman': 0, 'lipman
oasysdtnavymil': 0, 'robert': 0, 'lipman': 0, 'newsgroups': 0, 'compgraphicsaltgraphicscompgraphicsanimation': 0, 'subject':
0, 'call': 0, 'presentation': 0, 'navy': 0, 'scivizvr': 0, 'seminar': 0, 'messageid': 0, '32850oasysdtnavymil': 0, 'date': 0,
'nineteen': 0, 'mar': 0, 'ninety-three': 0, 'two hundred and one thousand and twenty-three': 0, 'gmt': 0, 'articleid': 0, 'oa
sys32850': 0, 'expires': 0, 'thirty': 0, 'apr': 0, 'forty thousand': 0, 'replyto': 0, 'followupto': 0, 'compgraphics': 0, 'di
stribution': 0, 'usa': 0, 'organization': 0, 'carderock': 0, 'division': 0, 'nswc': 0, 'bethesda': 0, 'md': 0, 'line': 0, 'si
xty-five': 0, 'scientific': 1.0437724975152622, 'visualization': 0, 'virtual': 0, 'reality': 0, 'tuesday': 0, 'june': 0, 'twe
nty-two': 0, 'one thousand, nine hundred and ninety-three': 0, 'naval': 0, 'surface': 0, 'warfare': 0, 'center': 0, 'formerl
y': 0, 'david': 0, 'taylor': 0, 'research': 0, 'maryland': 0, 'sponsor': 0, 'ness': 0, 'engineering': 0, 'software': 0, 'syst
em': 0, 'sponsoring': 0, 'oneday': 0, 'purpose': 0, 'present': 0, 'exchange': 0, 'information': 0, 'navyrelated': 0, 'progra
m': 0, 'development': 0, 'application': 0, 'solicited': 0, 'aspect': 0, 'current': 0, 'work': 0, 'worksinprogress': 0, 'propo
sed': 0, 'considered': 0, 'four': 0, 'type': 0, 'available': 0, 'one': 0, 'regular': 0, 'two thousand and thirty': 0, 'minut
e': 0, 'length': 0, 'two': 0, 'short': 0, 'ten': 0, 'three': 0, 'video': 0, 'standalone': 0, 'videotape': 0, 'author': 0, 'ne
ed': 0, 'attend': 0, 'demonstration': 0, 'byoh': 0, 'accepted': 0, 'published': 0, 'proceeding': 0, 'however': 0, 'viewgrap
h': 0, 'material': 0, 'reproduced': 0, 'attendee': 0, 'abstract': 0, 'submit': 0, 'page': 0, 'andor': 0, 'code': 0, 'two thou
sand and forty-two': 0, 'two hundred million, eight hundred and forty-five thousand': 0, 'voice': 0, 'three hundred and one':
0, 'two million, two hundred and seventy-three thousand, six hundred and eighteen': 0, 'fax': 0, 'two million, two hundred an
```

2. Result for initial query
   Cosine similarity is taken between query and each of the documents and user is asked to enter the number of documents he or she wants to retrieve. Top N(here N=100) documents are then retrieved.

```
enter the intial query
 Scientific tools for preserving rights and body
enter the value of k
100
top  100 documents after initial query are:
 [('61335', 0.10505332863961168), ('59434', 0.10502672361792531), ('61385', 0.07873147052336395), ('39655', 0.073192316388264
9), ('59137', 0.07048361788635653), ('60809', 0.06971994080254776), ('60797', 0.0673659599027222), ('60229', 0.0659221631817606
1), ('37920', 0.06566813507018805), ('58053', 0.06566813507018805), ('61079', 0.06421541336541815), ('60820', 0.063890051222154
44), ('60248', 0.06334560916612887), ('39736', 0.06254202044924066), ('178540', 0.06228185611626066), ('60819', 0.0584874530427
495), ('58569', 0.05660239760488963), ('61009', 0.05629679152604827), ('58900', 0.053567388125561895), ('59291', 0.050042009246
02019), ('61456', 0.04922890045140002), ('38409', 0.04788220183043207), ('58898', 0.04763127835002974), ('58997', 0.04733223550
562661), ('59548', 0.047283387242569835), ('37919', 0.04726957566534258), ('58052', 0.04726957566534258), ('58719', 0.047017022
77466562), ('58108', 0.04429092452549571), ('179106', 0.04413823482626761), ('58983', 0.044034314992416154), ('178546', 0.04373
303061611921), ('38375', 0.043704747610912496), ('38484', 0.04353453894377965), ('38852', 0.043529419711644407), ('38764', 0.043
21499080446276), ('58775', 0.04307746443337585), ('61145', 0.04300450149532338), ('58139', 0.04265035643907839), ('60872', 0.04
26118438979152), ('58941', 0.04183205782648998), ('60938', 0.04137589471512744), ('59273', 0.04133393015686746), ('39055', 0.04
104169810954415), ('59372', 0.04101635340334315), ('59635', 0.04086444526930199), ('59371', 0.04075469926728846), ('179045', 0.
040528177286516O1), ('179047', 0.040399180933298084), ('58958', 0.040227257245460626), ('58880', 0.04002708369276278), ('3860
9', 0.03980188489689949), ('38311', 0.03942415548929469), ('59069', 0.03938210297862666), ('59871', 0.03868975442891348), ('372
61', 0.038634978107355), ('38912', 0.03837417671903167), ('58961', 0.03792779407879595), ('178817', 0.03780352127492619), ('390
21', 0.037759746661951726), ('59233', 0.03772433265186211), ('58910', 0.03770373322216642), ('61568', 0.037522937058861734),
('60196', 0.03719828351411357), ('58921', 0.03696872306005461), ('59023', 0.03690243360347496), ('59589', 0.03672901109495136),
('61331', 0.03665165390127265), ('58981', 0.036624709788802945), ('178879', 0.03657028207824159), ('58113', 0.03627297717350820
6), ('58882', 0.036268716799420946), ('38272', 0.03593631836423884), ('58947', 0.03548457174784546), ('38223', 0.03547025010951
4746), ('59596', 0.03529412048886632), ('61257', 0.035200153930083425), ('61241', 0.03509644502896389), ('178802', 0.0349694863
699555), ('53556', 0.034831505630741895), ('39073', 0.034639839896188), ('38897', 0.03462456162375003), ('176974', 0.0345017613
5017141), ('59587', 0.03437686693263479), ('61118', 0.03434929356018378), ('39057', 0.034028280978115326), ('58939', 0.03392967
5362389645), ('59338', 0.03354477798156893), ('178340', 0.03346448456068945), ('61273', 0.03319677056999851 4), ('58938', 0.0331
0548070641074), ('38769', 0.033050558873821516), ('178945', 0.0330409062629923), ('39643', 0.032954792897657265), ('58867', 0.0
32787095772592804), ('176850', 0.03277921707083707), ('59269', 0.03265733227737051), ('58792', 0.032596582925197216), ('62408',
0.032554678818562194), ('53685', 0.03245067900905791)]
```

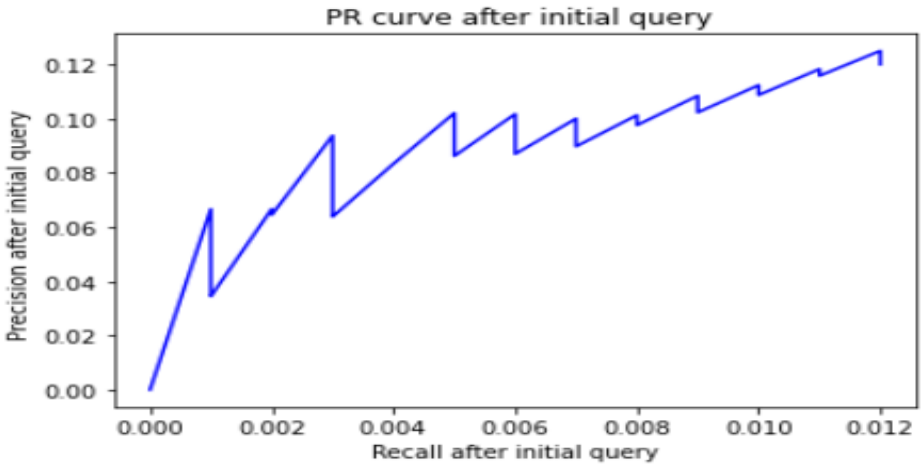
3. Precision, Recall and MAP for initial query

```
Enter the ground truth folder:

talk.politics.misc
precision after initial query are:  [0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.06666666666666667,
0.0625, 0.058823529411764705, 0.05555555555555555, 0.05263157894736842, 0.05, 0.047619047619047616, 0.045454545454545456, 0.043
478260869565216, 0.041666666666666664, 0.04, 0.038461538461538464, 0.037037037037037035, 0.03571428571428571, 0.034482758620689
655, 0.06666666666666667, 0.0645161290322580 6, 0.09375, 0.09090909090909091, 0.08823529411764706, 0.08571428571428572, 0.083333
33333333333, 0.08108108108108109, 0.07894736842105263, 0.07692307692307693, 0.075, 0.07317073170731707, 0.07142857142857142, 0.
06976744186046512, 0.06818181818181818, 0.06666666666666667, 0.06521739130434782, 0.06382978723404255, 0.08333333333333333, 0.1
0204081632653061, 0.1, 0.09803921568627451, 0.09615384615384616, 0.09433962264150944, 0.09259259259259259, 0.09090909090909091,
0.08928571428571429, 0.08771929824561403, 0.08620689655172414, 0.1016949152542373, 0.1, 0.09836065573770492, 0.096774193548387
1, 0.09523809523809523, 0.09375, 0.0923076923076923 1, 0.09090909090909091, 0.08955223880597014, 0.08823529411764706, 0.08695652
173913043, 0.1, 0.09859154929577464, 0.09722222222222222, 0.0958904109589041, 0.0945945945945946, 0.09333333333333334, 0.092105
26315789473, 0.09090909090909091, 0.08974358974358974, 0.10126582278481013, 0.1, 0.09876543209876543, 0.0975609756097561, 0.108
43373493975904, 0.10714285714285714, 0.10588235294117647, 0.10465116279069768, 0.10344827586206896, 0.10227272727272728, 0.1123
595505617977 5, 0.1111111111111111, 0.10989010989010989, 0.10869565217391304, 0.11827956989247312, 0.11702127659574468, 0.115789
47368421053, 0.125, 0.12371134020618557, 0.12244897959183673, 0.12121212121212122, 0.12]
recall after initial query are:  [0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.001, 0.001, 0.001, 0.
001, 0.001, 0.001, 0.001, 0.001, 0.001, 0.001, 0.001, 0.001, 0.001, 0.001, 0.001, 0.001, 0.002, 0.002, 0.003, 0.003, 0.003, 0.003, 0.0
03, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003, 0.004, 0.005, 0.005, 0.005, 0.005, 0.005, 0.00
5, 0.005, 0.005, 0.005, 0.005, 0.006, 0.006, 0.006, 0.006, 0.006, 0.006, 0.006, 0.006, 0.006, 0.006, 0.006, 0.006, 0.007, 0.007, 0.00
7, 0.007, 0.007, 0.007, 0.007, 0.007, 0.007, 0.008, 0.008, 0.008, 0.008, 0.009, 0.009, 0.009, 0.009, 0.009, 0.009, 0.01, 0.01,
0.01, 0.01, 0.011, 0.011, 0.011, 0.012, 0.012, 0.012, 0.012, 0.012]
MAP after initial query are:  0.09829092303552289
```

4. Precision-Recall curve for initial query

```
plt.plot(r, p,color="blue")
plt.xlabel("Recall after initial query")
plt.ylabel("Precision after initial query ")
plt.title(" PR curve after initial query")
```

```
Text(0.5, 1.0, ' PR curve after initial query')
```



5. Updated query vector

```
print("New query vector after 1st iteration is \n", new_query1)
```

```
New query vector after 1st iteration is
 {'xref': 0.1686811689945925, 'cantaloupesrvcscmuedu': 0.1686811689945925, 'compgraphics37261': 0, 'altgraphics519': 0, 'comp
graphicsanimation2614': 0, 'path': 0, 'cantaloupesrvcscmuedudasnewsharvardeduogicseuwmeduzaphodmpsohiostateedudarwinsuranetdt
ixdtnavymiloasyslipman': 0, 'lipmanoasysdtnavymil': 0, 'robert': 0, 'lipman': 0, 'newsgroups': 0, 'compgraphicsaltgraphicscom
pgraphicsanimation': 0, 'subject': 0, 'call': 0, 'presentation': 0, 'navy': 0, 'scivizvr': 0, 'seminar': 0, 'messageid': 0,
'32850oasysdtnavymil': 0, 'date': 0, 'nineteen': 0, 'mar': 0, 'ninety-three': 0.06251493081957452, 'two hundred and one thous
and and twenty-three': 0, 'gmt': 0.02229698101510546, 'articleid': 0, 'oasys32850': 0, 'expires': 0, 'thirty': 0, 'apr': 0.00
975616891130798, 'forty thousand': 0, 'replyto': 0.015480510226557338, 'followupto': 0, 'compgraphics': 0, 'distribution': 0.
02548006237810159, 'usa': 0.03770569070566423, 'organization': 0.0042580658094939754, 'carderock': 0, 'division': 0.055219533
777731686, 'nswc': 0, 'bethesda': 0, 'md': 0, 'line': 0.0004272866969537387, 'sixty-five': 0, 'scientific': 1.223367671481455
9, 'visualization': 0, 'virtual': 0, 'reality': 0, 'tuesday': 0, 'june': 0, 'twenty-two': 0.03694723561424461, 'one thousand,
nine hundred and ninety-three': 0.021446200823421586, 'naval': 0, 'surface': 0, 'warfare': 0, 'center': 0, 'formerly': 0, 'da
vid': 0, 'taylor': 0, 'research': 0, 'maryland': 0, 'sponsor': 0, 'ness': 0, 'engineering': 0, 'software': 0, 'system': 0.019
86947516869729, 'sponsoring': 0, 'oneday': 0, 'purpose': 0, 'present': 0, 'exchange': 0, 'information': 0, 'navyrelated': 0,
'program': 0, 'development': 0, 'application': 0, 'solicited': 0, 'aspect': 0, 'current': 0, 'work': 0.01103656702432955, 'wo
rksinprogress': 0, 'proposed': 0, 'considered': 0, 'four': 0, 'type': 0, 'available': 0, 'one': 0.005946150732540248, 'regula
r': 0, 'two thousand and thirty': 0, 'minute': 0, 'length': 0, 'two': 0, 'short': 0, 'ten': 0.16192427913290164, 'three': 0.0
495567791395242, 'video': 0, 'standalone': 0, 'videotape': 0, 'author': 0, 'need': 0, 'attend': 0, 'demonstration': 0, 'byo
h': 0, 'accepted': 0, 'published': 0, 'proceeding': 0, 'however': 0.018638819436515394, 'viewgraph': 0, 'material': 0, 'repro
```

6. Result after 1st iteration according to the feedback provided by the user

```
Enter the p % of documents you want to mark as relevant

8
Enter the  8.0 documents you want to mark as relevant

178540,179106,178546,179045,178817,178879,178802,176850
docs after 1st iteration are:
 [('178546', 0.5505751212905082), ('178540', 0.5384032998690667), ('178802', 0.42601425785980046), ('178817', 0.406293762207711
13), ('178340', 0.3070965044986697), ('178879', 0.2744295569300923), ('179106', 0.27345751136314655), ('176850', 0.271150541238
49863), ('178695', 0.26986936720667154), ('179045', 0.2522265637824333), ('178587', 0.23696713427799537), ('178431', 0.18902961
38547921), ('178413', 0.1736166597141219), ('178721', 0.1699855022751588 8), ('178432', 0.16852503368481372), ('178492', 0.16728
401644993243), ('178329', 0.15777200884955836), ('178475', 0.1565462482238 9201), ('178419', 0.14963798529174563), ('178434', 0.
14794030769388314), ('178553', 0.1367601522400018), ('178416', 0.1366558806048883), ('178453', 0.13533485406334161), ('178517',
0.12865564188341663), ('178349', 0.12589359212135456), ('178446', 0.1222135634991021 3), ('178315', 0.11680530510599899), ('1790
91', 0.11637790478938011), ('178322', 0.11440798511917649), ('178944', 0.11110931976460663), ('178552', 0.1098326723169233),
('178321', 0.1088654769685187), ('179008', 0.10765680706936519), ('176907', 0.10723319185253583), ('176865', 0.1065070020275648
8), ('178352', 0.09932295602156661), ('178673', 0.09603659638231338), ('179071', 0.09576650791361277), ('178809', 0.09373546144
229172), ('179114', 0.09334413852672134), ('178535', 0.09100996459421135), ('176902', 0.09097285549277168), ('177004', 0.090723
72326219387), ('178449', 0.0896424115587 2715), ('178430', 0.08882274336628138), ('178978', 0.08843120830298318), ('178724', 0.0
882357948062912), ('178474', 0.08749734973243342), ('178698', 0.0869447457207 2575), ('178382', 0.0841377581747485), ('178865',
0.08387825261853979), ('179058', 0.08284911869458769), ('176896', 0.08267860186731307), ('178528', 0.08133283215509725), ('6133
5', 0.08128584788431775), ('59434', 0.08126526203736019), ('178908', 0.07977051718050224), ('178758', 0.0796193856745872), ('17
9041', 0.07896609246930668), ('59183', 0.07859731927943964), ('179009', 0.07820566016485 94), ('179033', 0.07798050115033828),
('178603', 0.07795382607475591), ('178911', 0.07779511056993764), ('178913', 0.07770969533688549), ('178518', 0.077675922764056
72), ('176898', 0.07707642712900147), ('178435', 0.07697171166426883), ('178810', 0.07688940453540256), ('179069', 0.0761841924
3232544), ('59548', 0.07602660068649578), ('178886', 0.07589624904389879), ('178487', 0.07571363334602886), ('178776', 0.075582
94398925365), ('178501', 0.07556869137893367), ('176901', 0.07542545827408241), ('179054', 0.07537571874427748), ('178300', 0.0
751934608582413), ('176970', 0.07508670208613578), ('178779', 0.07375033207259721), ('179073', 0.0735887002612508), ('176953',
0.07344479525287184), ('179015', 0.0730605859435604), ('179113', 0.07300671432543906), ('178313', 0.07273589430194044), ('17696
0', 0.07264023429174803), ('178711', 0.0724788547045 3977), ('178861', 0.07234024775246833), ('179029', 0.07229706813924795),
('178888', 0.0719995357994116), ('178571', 0.07197200340693996), ('178628', 0.07194714248638791), ('58910', 0.0717932064168687
3), ('178617', 0.07094064944028738), ('176946', 0.07081358514061765), ('178898', 0.07070188556010355), ('176903', 0.07065919355
938985), ('58569', 0.0705160692076 5986), ('176895', 0.07009612288402335), ('178877', 0.07003214153721798)]
```

7. Star marked relevant documents

```
print("docs after 1st iteration are: ",startlist1)
```

```
docs after 1st iteration are:  ['178546*', '178540*', '178802*', '178817*', '178340', '178879*', '179106*', '176850*', '17869
5', '179045*', '178587', '178431', '178413', '178721', '178432', '178492', '178329', '178475', '178419', '178434', '178553', '1
78416', '178453', '178517', '178349', '178446', '178315', '179091', '178322', '178944', '178552', '178321', '179008', '176907',
'176865', '178352', '178673', '179071', '178809', '179114', '178535', '176902', '177004', '178449', '178430', '178978', '17872
4', '178474', '178698', '178382', '178865', '179058', '176896', '178528', '61335', '59434', '178908', '178758', '179041', '5918
3', '179009', '179033', '178603', '178911', '178913', '178518', '176898', '178435', '178810', '179069', '59548', '178886', '178
487', '178776', '178501', '176901', '179054', '178300', '176970', '178779', '179073', '176953', '179015', '179113', '178313',
'176960', '178711', '178861', '179029', '178888', '178571', '178628', '58910', '178617', '176946', '178898', '176903', '58569',
'176895', '178877']
```

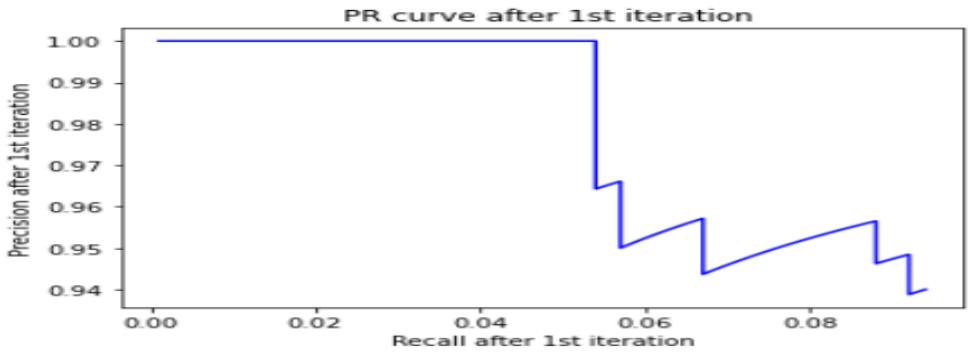
8. MAP, Precision and Recall after 1st iteration

```
precision after 1st iteration are:  [1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0,
1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.
0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 0.9818181818181818, 0.9642857142857143, 0.9649122807017544, 0.965517241379
3104, 0.9661016949152542, 0.95, 0.9508196721311475, 0.9516129032258065, 0.9523809523809523, 0.953125, 0.9538461538461539, 0.954
5454545454546, 0.9552238805970149, 0.9558823529411765, 0.9565217391304348, 0.9571428571428572, 0.9436619718309859, 0.9444444444
444444, 0.9452054794520548, 0.9459459459459459, 0.9466666666666667, 0.9473684210526315, 0.948051948051948, 0.9487179487179487,
0.9493670886075949, 0.95, 0.9506172839506173, 0.9512195121951219, 0.9518072289156626, 0.9523809523809523, 0.9529411764705882,
0.9534883720930233, 0.9540229885057471, 0.9545454545454546, 0.9550561797752809, 0.9555555555555556, 0.9560439560439561, 0.95652
17391304348, 0.946236559139785, 0.9468085106382979, 0.9473684210526315, 0.9479166666666666, 0.9484536082474226, 0.9387755102040
817, 0.9393939393939394, 0.94]
recall after 1st iteration are:  [0.001, 0.002, 0.003, 0.004, 0.005, 0.006, 0.007, 0.008, 0.009, 0.01, 0.011, 0.012, 0.013, 0.0
14, 0.015, 0.016, 0.017, 0.018, 0.019, 0.02, 0.021, 0.022, 0.023, 0.024, 0.025, 0.026, 0.027, 0.028, 0.029, 0.03, 0.031, 0.032,
0.033, 0.034, 0.035, 0.036, 0.037, 0.038, 0.039, 0.04, 0.041, 0.042, 0.043, 0.044, 0.045, 0.046, 0.047, 0.048, 0.049, 0.05, 0.0
51, 0.052, 0.053, 0.054, 0.054, 0.054, 0.055, 0.056, 0.057, 0.057, 0.058, 0.059, 0.06, 0.061, 0.062, 0.063, 0.064, 0.065, 0.06
6, 0.067, 0.067, 0.068, 0.069, 0.07, 0.071, 0.072, 0.073, 0.074, 0.075, 0.076, 0.077, 0.078, 0.079, 0.08, 0.081, 0.082, 0.083,
0.084, 0.085, 0.086, 0.087, 0.088, 0.088, 0.089, 0.09, 0.091, 0.092, 0.092, 0.093, 0.094]
MAP after 1st iteration are:  0.9795483156535947
```

9. Precision-Recall curve after iteration1

```
import matplotlib.pyplot as plt

plt.plot(r1, p1,color="blue")
plt.xlabel("Recall after 1st iteration")
plt.ylabel("Precision after 1st iteration")
plt.title(" PR curve after 1st iteration")
```

```
Text(0.5, 1.0, ' PR curve after 1st iteration')
```



10. Updated query vector

```
print("New query vector after 2nd iteration is \n", new_query2)
```

```
New query vector after 2nd iteration is
 {'xref': 0.27454901418319355, 'cantaloupesrvcscmuedu': 0.27454901418319355, 'compgraphics37261': 0.0, 'altgraphics519': 0.0,
'compgraphicsanimation2614': 0.0, 'path': 0, 'cantaloupesrvcscmuedudasnewsharvardeduogicseuwmeduzaphodmpsohiostateedudarwinsu
ranetdtixdtnavymiloasyslipman': 0.0, 'lipmanoasysdtnavymil': 0.0, 'robert': 0, 'lipman': 0.0, 'newsgroups': 0, 'compgraphicsa
ltgraphicscompgraphicsanimation': 0.0, 'subject': 0, 'call': 0, 'presentation': 0, 'navy': 0, 'scivizvr': 0.0, 'seminar': 0,
'messageid': 0, '32850oasysdtnavymil': 0.0, 'date': 0, 'nineteen': 0.04882504068990304, 'mar': 0.0, 'ninety-three': 0.0970803
0196849898, 'two hundred and one thousand and twenty-three': 0.0, 'gmt': 0.039931938972474317, 'articleid': 0, 'oasys32850':
0.0, 'expires': 0, 'thirty': 0, 'apr': 0.01900545952246563, 'forty thousand': 0.0, 'replyto': 0.04440348031371069, 'followupt
o': 0, 'compgraphics': 0.0, 'distribution': 0.12610087799025124, 'usa': 0.1355859386070562, 'organization': 0.008191201301842
58, 'carderock': 0.0, 'division': 0.2017846369449672, 'nswc': 0.0, 'bethesda': 0.0, 'md': 0, 'line': 0.0008538113679333666,
'sixty-five': 0, 'scientific': 1.1698919784405541, 'visualization': 0, 'virtual': 0.0, 'reality': 0, 'tuesday': 0, 'june': 0.
11521324506653384, 'twenty-two': 0.023243677765992017, 'one thousand, nine hundred and ninety-three': 0.05693520265296782, 'n
aval': 0.0, 'surface': 0, 'warfare': 0.0, 'center': 0, 'formerly': 0.0, 'david': 0.13393803377640484, 'taylor': 0.0, 'researc
h': 0, 'maryland': 0.0, 'sponsor': 0.0, 'ness': 0.0, 'engineering': 0, 'software': 0, 'system': 0.24505359030514626, 'sponsor
ing': 0, 'oneday': 0.0, 'purpose': 0, 'present': 0, 'exchange': 0, 'information': 0, 'navyrelated': 0.0, 'program': 0.0955547
1203932941, 'development': 0, 'application': 0, 'solicited': 0.0, 'aspect': 0, 'current': 0, 'work': 0, 'worksinprogress': 0.
0, 'proposed': 0, 'considered': 0.0882485987474942, 'four': 0.04345745743618437, 'type': 0, 'available': 0, 'one': 0.06559079
007975516, 'regular': 0, 'two thousand and thirty': 0, 'minute': 0, 'length': 0.0, 'two': 0.05073342122247762, 'short': 0, 't
en': 0.20654333358958737, 'three': 0.00013203535569648178, 'video': 0.0, 'standalone': 0.0, 'videotape': 0, 'author': 0, 'nee
d': 0.04806343776947647, 'attend': 0, 'demonstration': 0, 'byob': 0.0, 'accepted': 0, 'published': 0, 'proceeding': 0.0, 'how
```

11. Result after 2nd iteration according to the feedback provided by the user

```
Enter the  8.0 documents you want to mark as relevant

178340,178587,178475,178352,179071,178535,178449,178474
docs after 2nd iteration are:  [('178540', 0.5973423732074875), ('178546', 0.5889326678441816), ('178340', 0.5340903709353829),
('178587', 0.47647602853309035), ('178695', 0.43657747157133736), ('178475', 0.37161193263986436), ('178802', 0.340128907642309
6), ('178453', 0.33370431571039444), ('178449', 0.33185298817637554), ('178474', 0.3120040075318525), ('178431', 0.298391143305
9686), ('178413', 0.2959494375025263), ('178817', 0.2887987376904445), ('178416', 0.2856684590355181), ('179071', 0.28369387927
22084), ('178432', 0.27477661501205525), ('178446', 0.2631914232038628), ('178329', 0.25402743332027217), ('178419', 0.24936970
75736318), ('178492', 0.2485702508024964 5), ('178535', 0.2324954881179283), ('178434', 0.22100426091270972), ('178879', 0.20849
10407620833), ('178352', 0.1991605074454411), ('178349', 0.19875695371397153), ('179106', 0.1813399197250004), ('176850', 0.173
15982099492297), ('178944', 0.17223531706647835), ('179045', 0.16309220118256626), ('176865', 0.15740799252560658), ('176907',
0.15725220838604057), ('179091', 0.1517200245410143), ('179008', 0.1502712822486863), ('178673', 0.14633703269743148), ('17887
0', 0.14627974708495567), ('178645', 0.14337785212771592), ('178430', 0.13428383062140115), ('178528', 0.13404208593688507),
('178435', 0.13223372203619613), ('178382', 0.1315690402628703), ('179041', 0.12937746056180 59), ('178519', 0.1221153383453983
8), ('178978', 0.12159005213749945), ('178724', 0.1195125225439 8004), ('179081', 0.11872473261266456), ('179058', 0.11843232823
354836), ('178518', 0.11832764749868177), ('178501', 0.11538118128852037), ('177004', 0.11387374587301916), ('178721', 0.113051
04126017905), ('178646', 0.112758919510 9228), ('178487', 0.11205472684152636), ('178914', 0.10946059735369487), ('178300', 0.10
500955125503962), ('178649', 0.10340790646474869), ('178509', 0.10249653234734131), ('178422', 0.10123121880587771), ('178786',
0.10097343774458006), ('178317', 0.10076482721466767), ('178425', 0.10049710853565355), ('176953', 0.0994683576765967), ('17903
3', 0.0990721025044 3796), ('179029', 0.0987105590574 7707), ('178672', 0.09855717713304604), ('176951', 0.09718642302009392),
('178894', 0.0971445869162985), ('178314', 0.0967276928705598), ('178776', 0.0966287169415144), ('176970', 0.0962596277181718
4), ('176960', 0.09608359629769171), ('178865', 0.0959120417747 8659), ('178691', 0.09549471620497332), ('176902', 0.09451943478
19097), ('178433', 0.0945165129219708), ('178571', 0.09447467515484426), ('178799', 0.0940563968638292), ('178913', 0.093976932
1885809), ('178313', 0.0937935111274 4391), ('179028', 0.0936313519371 4978), ('178908', 0.09263528832082818), ('178529', 0.09227
897300803627), ('176874', 0.09207901936553778), ('178512', 0.0916864564786 8795), ('59183', 0.0916856648873 3089), ('178315', 0.0
9072949902489644), ('178357', 0.0900824029042 5407), ('178553', 0.0898687310379 8136), ('178873', 0.0892806215812 3273), ('17905
4', 0.08826598661525233), ('176929', 0.0879938603290345), ('178552', 0.08797909887378055), ('178513', 0.08740794902151368), ('1
79073', 0.08688232875515481), ('38403', 0.0867864393616 1261), ('176873', 0.0867825990050064), ('39638', 0.0866499202748 9002),
('39078', 0.08646829716795058), ('178293', 0.0862828862051102), ('178417', 0.0839041596489 7861), ('61560', 0.0838336044463368
5)]
```

12. Star marked relevant documents

```
print("docs after 2nd iteration are: ",startlist2)
```

```
docs after 2nd iteration are:  ['178540', '178546', '178340*', '178587*', '178695', '178475*', '178802', '178453', '178449*',
'178474*', '178431', '178413', '178817', '178416', '179071*', '178432', '178446', '178329', '178419', '178492', '178535*', '178
434', '178879', '178352*', '178349', '179106', '176850', '178944', '179045', '176865', '176907', '179091', '179008', '178673',
'178870', '178645', '178430', '178528', '178435', '178382', '179041', '178519', '178978', '178724', '179081', '179058', '17851
8', '178501', '177004', '178721', '178646', '178487', '178914', '178300', '178649', '178509', '178422', '178786', '178317', '17
8425', '176953', '179033', '179029', '178672', '176951', '178894', '178314', '178776', '176970', '176960', '178865', '178691',
'176902', '178433', '178571', '178799', '178913', '178313', '179028', '178908', '178529', '176874', '178512', '59183', '17831
5', '178357', '178553', '178873', '179054', '176929', '178552', '178513', '179073', '38403', '176873', '39638', '39078', '17829
3', '178417', '61560']
```

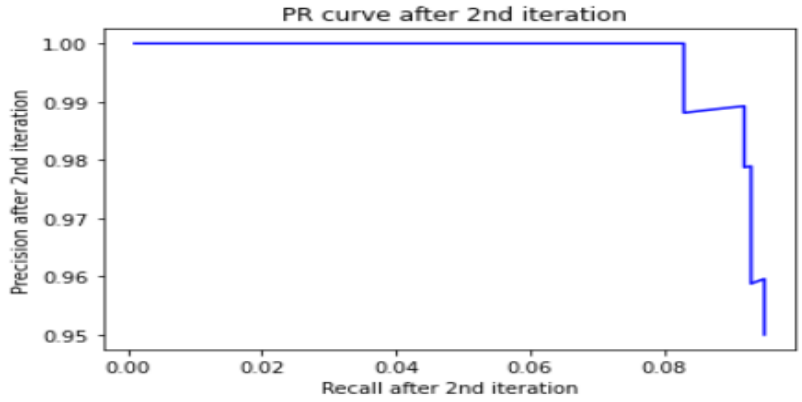
13. MAP, Precision and Recall after 2nd iteration

```
precision after 2nd iteration are:  [1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0,
1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.
0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0,
1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 0.9880952380952381, 0.9882352941176471, 0.988372093023255
8, 0.9885057471264368, 0.9886363636363636, 0.9887640449438202, 0.9888888888888889, 0.989010989010989, 0.9891304347826086, 0.989
247311827957, 0.9787234042553191, 0.9789473684210527, 0.96875, 0.9587628865979382, 0.9591836734693877, 0.9595959595959596, 0.9
5]


recall after 2nd iteration are:  [0.001, 0.002, 0.003, 0.004, 0.005, 0.006, 0.007, 0.008, 0.009, 0.01, 0.011, 0.012, 0.013, 0.0
14, 0.015, 0.016, 0.017, 0.018, 0.019, 0.02, 0.021, 0.022, 0.023, 0.024, 0.025, 0.026, 0.027, 0.028, 0.029, 0.03, 0.031, 0.032,
0.033, 0.034, 0.035, 0.036, 0.037, 0.038, 0.039, 0.04, 0.041, 0.042, 0.043, 0.044, 0.045, 0.046, 0.047, 0.048, 0.049, 0.05, 0.0
51, 0.052, 0.053, 0.054, 0.055, 0.056, 0.057, 0.058, 0.059, 0.06, 0.061, 0.062, 0.063, 0.064, 0.065, 0.066, 0.067, 0.068, 0.06
9, 0.07, 0.071, 0.072, 0.073, 0.074, 0.075, 0.076, 0.077, 0.078, 0.079, 0.08, 0.081, 0.082, 0.083, 0.083, 0.084, 0.085, 0.086,
0.087, 0.088, 0.089, 0.09, 0.091, 0.092, 0.092, 0.093, 0.093, 0.093, 0.094, 0.095, 0.095]


MAP after 2nd iteration is:  0.9978580859878353
```

14. Precision-Recall curve after iteration 2

```
plt.plot(r2, p2,color="blue")
plt.xlabel("Recall after 2nd iteration")
plt.ylabel("Precision after 2nd iteration")
plt.title(" PR curve after 2nd iteration")
```

```
Text(0.5, 1.0, ' PR curve after 2nd iteration')
```



15. Updated query vector

```
print("New query vector after 3rd iteration is \n", new_query3)
```

```
New query vector after 3rd iteration is
 {'xref': 0.383095089501423, 'cantaloupesrvcscmuedu': 0.383095089501423, 'compgraphics37261': 0.0, 'altgraphics519': 0.0, 'co
mpgraphicsanimation2614': 0.0, 'path': 0, 'cantaloupesrvcscmuedudasnewsharvardeduogicseuwmeduzaphodmpsohiostateedudarwinsuran
etdtixdtnavymiloasyslipman': 0.0, 'lipmanoasysdtnavymil': 0.0, 'robert': 0, 'lipman': 0.0, 'newsgroups': 0, 'compgraphicsaltg
raphicscompgraphicsanimation': 0.0, 'subject': 0, 'call': 0, 'presentation': 0, 'navy': 0, 'scivizvr': 0.0, 'seminar': 0, 'me
ssageid': 0, '32850oasysdtnavymil': 0.0, 'date': 0, 'nineteen': 0.04427258934352513, 'mar': 0.0, 'ninety-three': 0.0873514133
8967298, 'two hundred and one thousand and twenty-three': 0.0, 'gmt': 0.06221677290737249, 'articleid': 0, 'oasys32850': 0.0,
'expires': 0, 'thirty': 0, 'apr': 0.028415914517949926, 'forty thousand': 0.0, 'replyto': 0.15368522808128543, 'followupto':
0.04489240146858717, 'compgraphics': 0.0, 'distribution': 0.21810379337950359, 'usa': 0.21274300302373592, 'organization': 0.
011478558311036568, 'carderock': 0.0, 'division': 0.16735507511786174, 'nswc': 0.0, 'bethesda': 0, 'md': 0.1359767417488841,
'line': 0.001389390653918247, 'sixty-five': 0, 'scientific': 1.2470907309562345, 'visualization': 0.0, 'virtual': 0.0, 'reali
ty': 0, 'tuesday': 0, 'june': 0.0956121909583166, 'twenty-two': 0.07080485760529753, 'one thousand, nine hundred and ninety-t
hree': 0.11582242626075129, 'naval': 0, 'surface': 0, 'warfare': 0.0, 'center': 0.04164523975787416, 'formerly': 0.0, 'davi
d': 0.2486102541272511, 'taylor': 0.0, 'research': 0, 'maryland': 0, 'sponsor': 0.0, 'ness': 0.0, 'engineering': 0, 'softwar
e': 0, 'system': 0.3266242169312463, 'sponsoring': 0, 'oneday': 0.0, 'purpose': 0.02294030038393527, 'present': 0.02932061632
240178, 'exchange': 0.11194759159826873, 'information': 0, 'navyrelated': 0.0, 'program': 0.03968636613962491, 'development':
0, 'application': 0, 'solicited': 0, 'aspect': 0, 'current': 0.044539683099432265, 'work': 0.1203020666637748, 'worksinprogre
ss': 0.0, 'proposed': 0, 'considered': 0.17766702005732457, 'four': 0.09564681636906834, 'type': 0, 'available': 0, 'one': 0.
18615172143895475, 'regular': 0.05050062207170254, 'two thousand and thirty': 0.0, 'minute': 0, 'length': 0, 'two': 0.0283547
02952788982, 'short': 0, 'ten': 0.2780761165551217, 'three': 0, 'video': 0.0, 'standalone': 0, 'videotape': 0.0, 'author': 0,
```

16. Result after 3rd iteration according to the feedback provided by the user

```
Enter the  8.0 documents you want to mark as relevant

178446,176850,179091,178914,178509,178317,178776,178908
docs after 3rd iteration are:  [('178546', 0.4797630935926448), ('178540', 0.4758087870260328), ('178340', 0.4266054308137571),
('178587', 0.3973167055717983), ('178908', 0.3961426435145264), ('178776', 0.3649026923894124), ('178914', 0.3561720700186404),
('178695', 0.34879553843029854), ('178446', 0.3257566978044476), ('178475', 0.3098242177000677), ('178509', 0.309554556421828
2), ('178317', 0.3041103597278675), ('178474', 0.3023801469824379), ('178449', 0.29519461308431355), ('178453', 0.2933086383546
242), ('178432', 0.28795890253671985), ('179071', 0.27576515135130986), ('178413', 0.26722810251113854), ('179091', 0.256104918
51800515), ('176850', 0.2558353773665004), ('178419', 0.2554788027757497), ('178802', 0.2493112931616015), ('178330', 0.2464996
994489531), ('178724', 0.24601906290476366), ('178431', 0.24440487333261757), ('178352', 0.24198756086031975), ('178416', 0.241
6364848388065), ('178527', 0.2385983167702339), ('178492', 0.23410787860762694), ('178489', 0.22427034098853513), ('178786', 0.
22333123339324504), ('178434', 0.22187380122992742), ('178817', 0.21709620733914076), ('179073', 0.2141964761994844), ('17832
9', 0.21052555377405485), ('179058', 0.20547053886162858), ('178512', 0.2050742060938545), ('178349', 0.1959536098818184), ('17
8513', 0.18643844980875734), ('178535', 0.18148741520108896), ('178879', 0.18114362727230157), ('178672', 0.18111786585141904),
('179034', 0.18100985397239655), ('179106', 0.1803546280950591), ('178382', 0.17713002119308752), ('178433', 0.1752713840263474
8), ('178571', 0.17362221245651907), ('178425', 0.17307153592249847), ('176951', 0.1712476701195341), ('178944', 0.169000193809
43973), ('178865', 0.16846378932004133), ('178357', 0.16769528989374427), ('178547', 0.16467521227006127), ('178314', 0.1637622
4994661137), ('179029', 0.16365211024384482), ('179054', 0.16241960680952006), ('178965', 0.162118920428346), ('178313', 0.1607
861186094245), ('178295', 0.16032159968628804), ('178550', 0.15999068256139287), ('178929', 0.15936246602208654), ('176960', 0.
15702540806399648), ('178688', 0.154030916359234), ('178403', 0.15319358230948776), ('178293', 0.15291242067613775), ('178645',
0.15286110598185385), ('176970', 0.1526999054993391), ('179008', 0.1526958339779663), ('177004', 0.14952335153353427), ('17686
5', 0.14952084006105862), ('178402', 0.148729135795469), ('178528', 0.14844984364709798), ('178529', 0.14671480244157706), ('17
8435', 0.1439657443195967), ('178538', 0.14326892956859613), ('176907', 0.1420044837703199), ('179045', 0.1399936729540355),
('176944', 0.13988823558720487), ('178487', 0.13650430690678864), ('178673', 0.1364642237308892), ('59183', 0.1362244895481054
3), ('178430', 0.13468242595996863), ('178873', 0.1340708824640847), ('178870', 0.13375022638679634), ('178894', 0.133574741379
72575), ('178762', 0.13305602896230795), ('176953', 0.1329539291110896), ('178503', 0.1328775238303745), ('178876', 0.131769852
13292128), ('178406', 0.1307057795785956), ('178799', 0.12966567459044417), ('178405', 0.12837187158737365), ('178422', 0.12823
257750192005), ('178913', 0.12725863989918965), ('179088', 0.12672372947787422), ('178773', 0.12657186231577147), ('179067', 0.
12613655572354324), ('178950', 0.12584291098656458), ('176936', 0.1252082199328637), ('178641', 0.1249189072774262)]
```

17. Star marked relevant documents

```
print("docs after 3rd iteration are: ",startlist3)
```

```
docs after 3rd iteration are:  ['178546', '178540', '178340', '178587', '178908*', '178776*', '178914*', '178695', '178446*',
'178475', '178509*', '178317*', '178474', '178449', '178453', '178432', '179071', '178413', '179091*', '176850*', '178419', '17
8802', '178330', '178724', '178431', '178352', '178416', '178527', '178492', '178489', '178786', '178434', '178817', '179073',
'178329', '179058', '178512', '178349', '178513', '178535', '178879', '178672', '179034', '179106', '178382', '178433', '17857
1', '178425', '176951', '178944', '178865', '178357', '178547', '178314', '179029', '179054', '178965', '178313', '178295', '17
8550', '178929', '176960', '178688', '178403', '178293', '178645', '176970', '179008', '177004', '176865', '178402', '178528',
'178529', '178435', '178538', '176907', '179045', '176944', '178487', '178673', '59183', '178430', '178873', '178870', '17889
4', '178762', '176953', '178503', '178876', '178406', '178799', '178405', '178422', '178913', '179088', '178773', '179067', '17
8950', '176936', '178641']
```

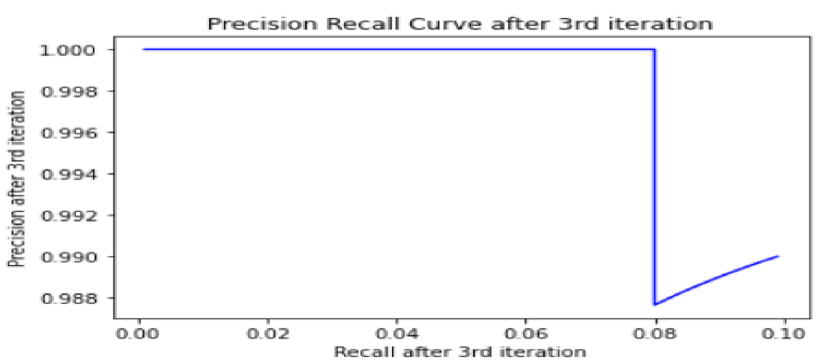
18. MAP, Precision and Recall after 3rd iteration

```
precision after 3rd iteration are:  [1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0,
1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.
0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0,
1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 0.9876543209876543, 0.9878048780487805, 0.9879518072289156, 0.9880952380
952381, 0.9882352941176471, 0.9883720930232558, 0.9885057471264368, 0.9886363636363636, 0.9887640449438202, 0.9888888888888889,
0.989010989010989, 0.9891304347826086, 0.989247311827957, 0.9893617021276596, 0.9894736842105263, 0.9895833333333334, 0.9896907
216494846, 0.9897959183673469, 0.98989898989899, 0.99]


recall after 3rd iteration are:  [0.001, 0.002, 0.003, 0.004, 0.005, 0.006, 0.007, 0.008, 0.009, 0.01, 0.011, 0.012, 0.013, 0.0
14, 0.015, 0.016, 0.017, 0.018, 0.019, 0.02, 0.021, 0.022, 0.023, 0.024, 0.025, 0.026, 0.027, 0.028, 0.029, 0.03, 0.031, 0.032,
0.033, 0.034, 0.035, 0.036, 0.037, 0.038, 0.039, 0.04, 0.041, 0.042, 0.043, 0.044, 0.045, 0.046, 0.047, 0.048, 0.049, 0.05, 0.0
51, 0.052, 0.053, 0.054, 0.055, 0.056, 0.057, 0.058, 0.059, 0.06, 0.061, 0.062, 0.063, 0.064, 0.065, 0.066, 0.067, 0.068, 0.06
9, 0.07, 0.071, 0.072, 0.073, 0.074, 0.075, 0.076, 0.077, 0.078, 0.079, 0.08, 0.08, 0.081, 0.082, 0.083, 0.084, 0.085, 0.086,
0.087, 0.088, 0.089, 0.09, 0.091, 0.092, 0.093, 0.094, 0.095, 0.096, 0.097, 0.098, 0.099]


MAP after 3rd iteration is:  0.9978833074779618
```

19. Precision-Recall curve after iteration 3

```python
plt.plot(r3, p3,color="blue")
plt.xlabel("Recall after 3rd iteration")
plt.ylabel("Precision after 3rd iteration")
plt.title("Precision Recall Curve after 3rd iteration")
```

```
Text(0.5, 1.0, 'Precision Recall Curve after 3rd iteration')
```



20. Updated query vector

```python
print("New query vector after 4th iteration is \n", new_query4)
```

```
New query vector after 4th iteration is
 {'xref': 0.5216546211755526, 'cantaloupesrvcscmuedu': 0.5216546211755526, 'compgraphics37261': 0.0, 'altgraphics519': 0.0,
'compgraphicsanimation2614': 0.0, 'path': 0, 'cantaloupesrvcscmuedudasnewsharvardeduogicseuwmeduzaphodmpsohiostateedudarwinsu
ranetdtixdtnavymiloasyslipman': 0.0, 'lipmanoasysdtnavymil': 0.0, 'robert': 0, 'lipman': 0.0, 'newsgroups': 0, 'compgraphicsa
ltgraphicscompgraphicsanimation': 0.0, 'subject': 0, 'call': 0, 'presentation': 0, 'navy': 0, 'scivizvr': 0.0, 'seminar': 0,
'messageid': 0, '32850oasysdtnavymil': 0.0, 'date': 0, 'nineteen': 0.027007502969208157, 'mar': 0.0, 'ninety-three': 0.165017
97325754122, 'two hundred and one thousand and twenty-three': 0.0, 'gmt': 0.08646781268156123, 'articleid': 0.0, 'oasys3285
0': 0.0, 'expires': 0.0, 'thirty': 0, 'apr': 0.037626168417954624, 'forty thousand': 0.0, 'replyto': 0.17172808402907075, 'fo
llowupto': 0.02421066025359629, 'compgraphics': 0.0, 'distribution': 0.24108452512295897, 'usa': 0.23887793834923976, 'organi
zation': 0.01565059872810939, 'carderock': 0.0, 'division': 0.23258868136025754, 'nswc': 0.0, 'bethesda': 0, 'md': 0.11073893
797440984, 'line': 0.001786423169802242, 'sixty-five': 0, 'scientific': 1.619170264980785, 'visualization': 0.0, 'virtual':
0.0, 'reality': 0, 'tuesday': 0, 'june': 0.09133711692987379, 'twenty-two': 0.12157848866095448, 'one thousand, nine hundred
and ninety-three': 0.13074569822032245, 'naval': 0, 'surface': 0, 'warfare': 0, 'center': 0.02637662588771003, 'formerly': 0.
0, 'david': 0.20026274641029715, 'taylor': 0.0, 'research': 0, 'maryland': 0, 'sponsor': 0, 'ness': 0.0, 'engineering': 0, 's
oftware': 0, 'system': 0.3100457591847312, 'sponsoring': 0, 'oneday': 0.0, 'purpose': 0, 'present': 0, 'exchange': 0.10706970
089681911, 'information': 0, 'navyrelated': 0.0, 'program': 0, 'development': 0, 'application': 0, 'solicited': 0.0, 'aspec
t': 0, 'current': 0, 'work': 0.08425234495851167, 'worksinprogress': 0.0, 'proposed': 0, 'considered': 0.08095485514761493,
'four': 0.0291062543570106, 'type': 0, 'available': 0.03098204378527826, 'one': 0.133934693226166, 'regular': 0.0283076218558
24836, 'two thousand and thirty': 0.0, 'minute': 0, 'length': 0, 'two': 0, 'short': 0, 'ten': 0.4370384431109674, 'three': 0.
```

21. Result after 4th iteration according to the feedback provided by the user

```
Enter the  8.0 documents you want to mark as relevant

178449,178382,187944,179029,178295,179045,178673,178950
docs after 4th iteration are:  [('178546', 0.5742988752114335), ('178540', 0.5660283657371689), ('178340', 0.4252017908369999
4), ('178587', 0.38243859248301093), ('178802', 0.3536155330358039), ('178695', 0.35223109191672364), ('178817', 0.322530775936
68517), ('176850', 0.2965949108911706), ('178475', 0.29075125862856094), ('178446', 0.2871341694219442), ('178908', 0.278740900
6614901), ('178776', 0.27597572465441056), ('178432', 0.27117680293063745), ('178914', 0.25718256531784334), ('178474', 0.25403
30031635768), ('178413', 0.25006003308042457), ('178453', 0.2484720648749891), ('178431', 0.24577174698554743), ('178879', 0.23
989511681687803), ('178509', 0.23479879784518093), ('178419', 0.23479128967536558), ('178449', 0.23472416965211484), ('179071',
0.2317782404992172), ('179091', 0.2307645693031834), ('179106', 0.23053106219691488), ('178317', 0.23023185031312496), ('17835
2', 0.2200950285871807), ('178492', 0.22003084875384366), ('178416', 0.21299473797221863), ('178434', 0.2111992611128461), ('17
8329', 0.20151372893988553), ('179045', 0.1950397380812022), ('178724', 0.18729285738892057), ('178330', 0.1815467034670769),
('178349', 0.17768110351796565), ('178527', 0.17598380280383108), ('178535', 0.1632594108086029), ('178512', 0.160223535686122
4), ('178489', 0.1598946734981414), ('178786', 0.1571417560257544), ('178944', 0.14902068841871452), ('178513', 0.1474088079623
3857), ('176865', 0.13911703914957177), ('179073', 0.13823254404989124), ('179008', 0.13774749242106737), ('179058', 0.13737491
465996296), ('178721', 0.13681846787663862), ('178357', 0.13594184055735953), ('176951', 0.13569809165259472), ('176907', 0.135
02761202719263), ('178382', 0.13059656715909526), ('178673', 0.1295268548783674), ('178865', 0.1276402274249001), ('179029', 0.
12697234719000103), ('178965', 0.12302069481324765), ('178487', 0.122931353320774), ('177004', 0.12282730418167034), ('178645',
0.1223554028813977), ('178672', 0.12088119069553084), ('178435', 0.11978750273177903), ('178433', 0.1180591252089360l), ('17842
5', 0.1175118984141245), ('178430', 0.11699767455875502), ('178688', 0.11692330860571315), ('176970', 0.11678510163355793), ('1
78528', 0.11531628242176098), ('178518', 0.1149361557702396), ('178552', 0.11336354085779268), ('178553', 0.11218744743325426),
('179034', 0.11136053930478621), ('178315', 0.11062330382381164), ('178571', 0.1105889078294487), ('178870', 0.11043887263629),
('178295', 0.11010512697647495), ('178529', 0.1100728106034062), ('178313', 0.10839451881009456), ('59183', 0.1073745850755129
4), ('179041', 0.10690554259777803), ('179054', 0.10685394312708565), ('178929', 0.1065555267880302), ('176960', 0.106256706818
71768), ('178517', 0.10617998449205288), ('178978', 0.10586961917902381), ('178322', 0.10539163848920963), ('178314', 0.1043993
0780824287), ('179081', 0.10398262576413383), ('176953', 0.10389630258410072), ('178894', 0.103893533511982), ('178913', 0.1035
7479620719659), ('178402', 0.1033617677221564), ('179033', 0.10266616877976643), ('178321', 0.10255841480707346), ('178501', 0.
1010289417019437), ('179114', 0.10083563706351897), ('178300', 0.10059473897295423), ('176948', 0.09998076872301881), ('17876
2', 0.09976566567321708), ('176902', 0.09974617131445121), ('178293', 0.09924424032158405), ('178403', 0.0992103140521560l)]
```

22. Star marked relevant documents

```
print("docs after 4th iteration are: ",startlist4)
```

```
docs after 4th iteration are:  ['178546*', '178540*', '178340', '178587', '178802*', '178695', '178817*', '176850*', '178475',
'178446', '178908', '178776', '178432', '178914', '178474', '178413', '178453', '178431', '178879*', '178509', '178419', '17844
9', '179071', '179091', '179106*', '178317', '178352', '178492', '178416', '178434', '178329', '179045*', '178724', '178330',
'178349', '178527', '178535', '178512', '178489', '178786', '178944', '178513', '176865', '179073', '179008', '179058', '17872
1', '178357', '176951', '176907', '178382', '178673', '178865', '179029', '178965', '178487', '177004', '178645', '178672', '17
8435', '178433', '178425', '178430', '178688', '176970', '178528', '178518', '178552', '178553', '179034', '178315', '178571',
'178870', '178295', '178529', '178313', '59183', '179041', '179054', '178929', '176960', '178517', '178978', '178322', '17831
4', '179081', '176953', '178894', '178913', '178402', '179033', '178321', '178501', '179114', '178300', '176948', '178762', '17
6902', '178293', '178403']
```

23. Final result with all the marked relevant documents by user (docs are marked with '*')

```
print("Relevant documents after 4 iterations are: \n ",final_star)
```

```
Relevant documents after 4 iterations are:
  ['178546*', '178540*', '178340*', '178587*', '178802*', '178695', '178817*', '176850*', '178475*', '178446*', '178908*', '178
776*', '178432', '178914*', '178474*', '178413', '178453', '178431', '178879*', '178509*', '178419', '178449*', '179071*', '179
091*', '179106*', '178317*', '178352*', '178492', '178416', '178434', '178329', '179045*', '178724', '178330', '178349', '17852
7', '178535*', '178512', '178489', '178786', '178944', '178513', '176865', '179073', '179008', '179058', '178721', '178357', '1
76951', '176907', '178382', '178673', '178865', '179029', '178965', '178487', '177004', '178645', '178672', '178435', '178433',
'178425', '178430', '178688', '176970', '178528', '178518', '178552', '178553', '179034', '178315', '178571', '178870', '17829
5', '178529', '178313', '59183', '179041', '179054', '178929', '176960', '178517', '178978', '178322', '178314', '179081', '176
953', '178894', '178913', '178402', '179033', '178321', '178501', '179114', '178300', '176948', '178762', '176902', '178293',
'178403']
```

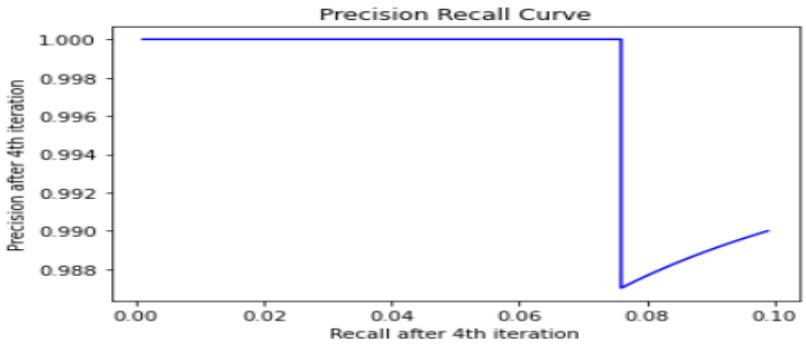
24. MAP, Precision and Recall after 4th iteration

```
precision after 4th iteration are:  [1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0,
1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.
0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0,
1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 0.987012987012987, 0.9871794871794872, 0.9873417721518988, 0.9875, 0.9876543209876543, 0.987
8048780487805, 0.9879518072289156, 0.9880952380952381, 0.9882352941176471, 0.9883720930232558, 0.9885057471264368, 0.9886363636
363636, 0.9887640449438202, 0.9888888888888889, 0.989010989010989, 0.9891304347826086, 0.989247311827957, 0.9893617021276596,
0.9894736842105263, 0.9895833333333334, 0.9896907216494846, 0.9897959183673469, 0.98989898989899, 0.99]


recall after 4th iteration are:  [0.001, 0.002, 0.003, 0.004, 0.005, 0.006, 0.007, 0.008, 0.009, 0.01, 0.011, 0.012, 0.013, 0.0
14, 0.015, 0.016, 0.017, 0.018, 0.019, 0.02, 0.021, 0.022, 0.023, 0.024, 0.025, 0.026, 0.027, 0.028, 0.029, 0.03, 0.031, 0.032,
0.033, 0.034, 0.035, 0.036, 0.037, 0.038, 0.039, 0.04, 0.041, 0.042, 0.043, 0.044, 0.045, 0.046, 0.047, 0.048, 0.049, 0.05, 0.0
51, 0.052, 0.053, 0.054, 0.055, 0.056, 0.057, 0.058, 0.059, 0.06, 0.061, 0.062, 0.063, 0.064, 0.065, 0.066, 0.067, 0.068, 0.06
9, 0.07, 0.071, 0.072, 0.073, 0.074, 0.075, 0.076, 0.076, 0.077, 0.078, 0.079, 0.08, 0.081, 0.082, 0.083, 0.084, 0.085, 0.086,
0.087, 0.088, 0.089, 0.09, 0.091, 0.092, 0.093, 0.094, 0.095, 0.096, 0.097, 0.098, 0.099]


MAP after 4th iteration is:  0.997374980006437
```

25. Precision-Recall curve after iteration 4

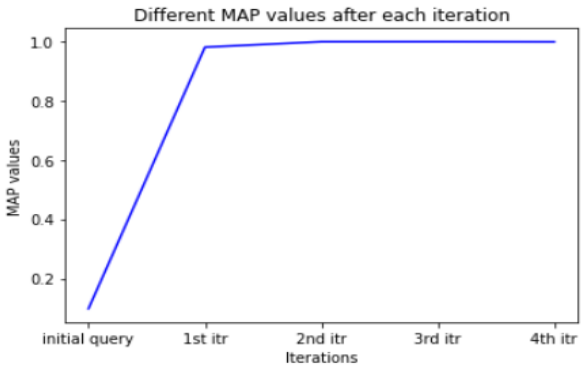
```
plt.plot(r4, p4,color="blue")
plt.xlabel("Recall after 4th iteration")
plt.ylabel("Precision after 4th iteration")
plt.title("Precision Recall Curve")
```

```
Text(0.5, 1.0, 'Precision Recall Curve')
```



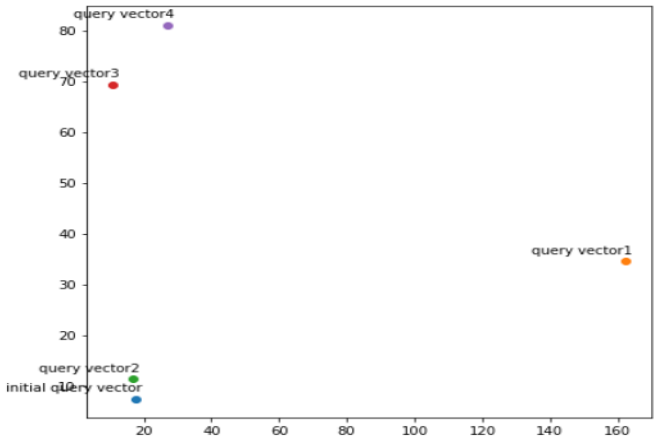
26. MAP at the end of 4 iterations

```
[0.09829092303552289, 0.9795483156535947, 0.9978580859878353, 0.9978833074779618, 0.997374980006437]

Text(0.5, 1.0, 'Different MAP values after each iteration')
```



27. T-SNE plot for all the query vectors

```
plt.figure(figsize=(7,7))
for i in range(len(x)):
    plt.scatter(x[i],y[i])
    plt.annotate(labels[i],xy=(x[i], y[i]),xytext=(5, 5),textcoords='offset points',ha='right',va='bottom')
plt.show()
```



Query 3: Frequently asked questions on State-of-the-art visualisation tools
Relevant set 3: Documents inside folder sci.med

1. Initial query vector


```
Initial query vector is :
 {'xref': 0, 'cantaloupesrvcscmuedu': 0, 'compgraphics37261': 0, 'altgraphics519': 0, 'compgraphicsanimation2614': 0, 'path':
0, 'cantaloupesrvcscmuedudasnewsharvardeduogicseuwmeduzaphodmpsohiostateedudarwinsuranetdtixdtnavymiloasyslipman': 0, 'lipman
oasysdtnavymil': 0, 'robert': 0, 'lipman': 0, 'newsgroups': 0, 'compgraphicsaltgraphicscompgraphicsanimation': 0, 'subject':
0, 'call': 0, 'presentation': 0, 'navy': 0, 'scivizvr': 0, 'seminar': 0, 'messageid': 0, '32850oasysdtnavymil': 0, 'date': 0,
'nineteen': 0, 'mar': 0, 'ninety-three': 0, 'two hundred and one thousand and twenty-three': 0, 'gmt': 0, 'articleid': 0, 'oa
sys32850': 0, 'expires': 0, 'thirty': 0, 'apr': 0, 'forty thousand': 0, 'replyto': 0, 'followupto': 0, 'compgraphics': 0, 'di
stribution': 0, 'usa': 0, 'organization': 0, 'carderock': 0, 'division': 0, 'nswc': 0, 'bethesda': 0, 'md': 0, 'line': 0, 'si
xty-five': 0, 'scientific': 0, 'visualization': 0, 'virtual': 0, 'reality': 0, 'tuesday': 0, 'june': 0, 'twenty-two': 0, 'one
thousand, nine hundred and ninety-three': 0, 'naval': 0, 'surface': 0, 'warfare': 0, 'center': 0, 'formerly': 0, 'david': 0,
'taylor': 0, 'research': 0, 'maryland': 0, 'sponsor': 0, 'ness': 0, 'engineering': 0, 'software': 0, 'system': 0, 'sponsorin
g': 0, 'oneday': 0, 'purpose': 0, 'present': 0, 'exchange': 0, 'information': 0, 'navyrelated': 0, 'program': 0, 'developmen
t': 0, 'application': 0, 'solicited': 0, 'aspect': 0, 'current': 0, 'work': 0, 'worksinprogress': 0, 'proposed': 0, 'consider
ed': 0, 'four': 0, 'type': 0, 'available': 0, 'one': 0, 'regular': 0, 'two thousand and thirty': 0, 'minute': 0, 'length': 0,
'two': 0, 'short': 0, 'ten': 0, 'three': 0, 'video': 0, 'standalone': 0, 'videotape': 0, 'author': 0, 'need': 0, 'attend': 0,
'demonstration': 0, 'byoh': 0, 'accepted': 0, 'published': 0, 'proceeding': 0, 'however': 0, 'viewgraph': 0, 'material': 0,
'reproduced': 0, 'attendee': 0, 'abstract': 0, 'submit': 0, 'page': 0, 'andor': 0, 'code': 0, 'two thousand and forty-two':
0, 'two hundred million, eight hundred and forty-five thousand': 0, 'voice': 0, 'three hundred and one': 0, 'two million, two
hundred and seventy-three thousand, six hundred and eighteen': 0, 'fax': 0, 'two million, two hundred and seventy-five thousa
```


2. Result for initial query
   Cosine similarity is taken between query and each of the documents and user is asked to enter the number of documents he or she wants to retrieve. Top N(here N=120) documents are then retrieved.


```
enter the intial query
 Frequently asked questions on state-of-the-art visualisation tools
enter the value of k
120
top  120 documents after initial query are:
 [('38962', 0.13076521098532082), ('37919', 0.11093199873580913), ('58052', 0.11093199873580913), ('37920', 0.1032140206506015
5), ('58053', 0.10321402065060155), ('61146', 0.08505313474006876), ('178540', 0.08005457736877879), ('38692', 0.07582754592338
617), ('38226', 0.06665390432897685), ('39655', 0.0636203882489307), ('61335', 0.06192241023886999), ('59434', 0.06190672823156
369), ('178340', 0.06142282739369062), ('38817', 0.06120212160785284), ('60809', 0.06060212220939798), ('38733', 0.059357575123
82941), ('178546', 0.059142043643053324), ('38823', 0.05887415755526643), ('60797', 0.05855598968938008), ('38400', 0.057747406
423487885), ('38236', 0.0560706997054576), ('60820', 0.05553465260527318), ('178713', 0.05542982385475616), ('38827', 0.0549018
1404242343), ('39736', 0.05436291429480169), ('59063', 0.05176385753869066), ('58569', 0.05055244723653025), ('61079', 0.048557
92358402392), ('38822', 0.04821072649642609), ('62373', 0.04791188560661775), ('38497', 0.047052621724421435), ('60819', 0.0461
3087082129971), ('38820', 0.04451276562907359), ('178450', 0.04450128006804642), ('178587', 0.04405967288393872), ('59871', 0.0
43940625218986895), ('58139', 0.043596687462078496), ('59324', 0.04348967357358362), ('59061', 0.042943701417032264), ('61456',
0.04279085448220878), ('178451', 0.0422661492205452), ('61385', 0.0422603669838703), ('38824', 0.042215177176169155), ('59310',
0.042127547125936235), ('38409', 0.04162027410781832), ('59100', 0.04009355856131211), ('59064', 0.03986224003135725), ('6100
9', 0.03956035235010687), ('38636', 0.039528704304211355), ('38272', 0.03924843516267776), ('58577', 0.037875017163949365), ('1
78695', 0.03784353579410378), ('38484', 0.03784118889131817), ('38963', 0.03735115226274614), ('39495', 0.0372904480287342),
('59009', 0.0368804830456377), ('60959', 0.036060866874021325), ('39055', 0.03567435622987396), ('60215', 0.03545078077593478),
('38434', 0.03528053401761101), ('179047', 0.03511586601906359), ('178302', 0.034972625896893504), ('58958', 0.0349664261282506
6), ('58844', 0.03444412099624873), ('61241', 0.03441888196660251), ('38311', 0.03426835223126349), ('53908', 0.033731132381452
74), ('38848', 0.03358393329566373), ('38490', 0.03347117475718761), ('59185', 0.032947554052949767), ('178908', 0.032818876198
435334), ('38778', 0.032659162602864965), ('61568', 0.03261577091318458), ('61006', 0.03251883694736533), ('178449', 0.03222985
7303614426), ('62406', 0.031730435316695685), ('59079', 0.03167907117673656), ('54136', 0.031591291820386794), ('59904', 0.0311
7850994592930 4), ('38223', 0.030831529791777654), ('59178', 0.03075420215268413), ('59026', 0.030710969268976453), ('178707',
0.03056190783802463), ('178416', 0.030364560428436292), ('53556', 0.030276318893466672), ('61145', 0.03021971920031214), ('3905
7', 0.02957813817214429), ('39053', 0.029395427055677943), ('61397', 0.02914181703593627), ('38957', 0.02848333536950221), ('62
126', 0.028415766487732256), ('54508', 0.028352996887746944), ('62408', 0.028297250421333), ('58838', 0.028230989879860592),
('53685', 0.02820685147531005), ('58880', 0.028127456139311967), ('58868', 0.027691687596036157), ('59098', 0.0276859859835210
6), ('54525', 0.02760053110324693), ('59087', 0.027514758913502323), ('39074', 0.027499377553648705), ('59432', 0.0271574845538
1521), ('38916', 0.02697407154315091), ('38912', 0.02696594088222141), ('39003', 0.02680329153284366), ('61000', 0.026798685817
423926), ('39065', 0.02672806691682489), ('39021', 0.02653417436598341 4), ('178453', 0.026532002464471096), ('39671', 0.0264088
3922666215), ('61074', 0.02635037899673 6087), ('59637', 0.02626151773517498), ('38853', 0.026169638135523907), ('38376', 0.0261
6853845126336), ('54086', 0.026137191180300275), ('54047', 0.02597247631986 7273), ('39014', 0.02591590548133329 2), ('62388', 0.
02585961014437 92), ('38711', 0.025810833149127778), ('58077', 0.025708203708514614)]
```

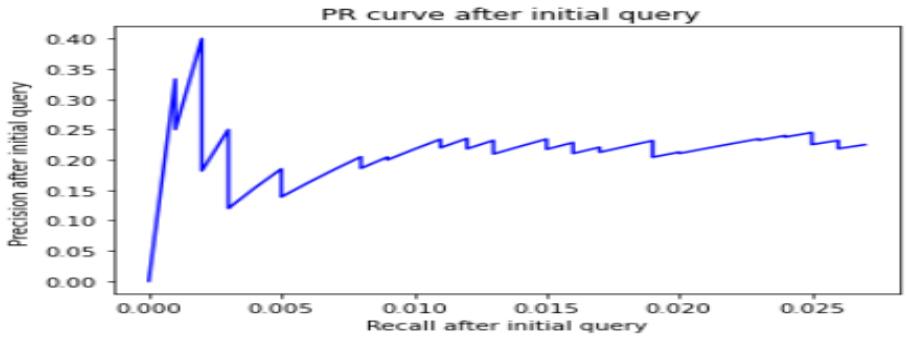
3. Precision, Recall and MAP for initial query

```
Enter the ground truth folder:

sci.med
precision after initial query are:  [0.0, 0.0, 0.3333333333333333, 0.25, 0.4, 0.3333333333333333, 0.2857142857142857, 0.25, 0.2
222222222222222, 0.2, 0.18181818181818182, 0.25, 0.23076923076923078, 0.21428571428571427, 0.2, 0.1875, 0.17647058823529413, 0.
16666666666666666, 0.15789473684210525, 0.15, 0.14285714285714285, 0.13636363636363635, 0.13043478260869565, 0.125, 0.12, 0.153
84615384615385, 0.14518518518518517, 0.1785714285714285, 0.1724137931034483, 0.16666666666666666, 0.16129032258064516, 0.1562
5, 0.15151515151515152, 0.14705882352941177, 0.14285714285714285, 0.1388888888888889, 0.16216216216216217, 0.18421052631578946,
0.20512820512820512, 0.2, 0.1951219512195122, 0.19047619047619047, 0.18604651162790697, 0.20454545454545456, 0.2, 0.21739130434
782608, 0.23404255319148937, 0.2291666666666666, 0.22448979591836735, 0.22, 0.23529411764705882, 0.23076923076923078, 0.226415
09433962265, 0.2222222222222222, 0.21818181818181817, 0.23214285714285715, 0.22807017543859648, 0.22413793103448276, 0.22033898
305084745, 0.21666666666666667, 0.21311475409836064, 0.20967741935483872, 0.2222222222222222, 0.234375, 0.23076923076923078, 0.
22727272727272727, 0.22388059701492538, 0.22058823529411764, 0.21739130434782608, 0.22857142857142856, 0.22535211267605634, 0.2
222222222222222, 0.2191780821917808, 0.21621621621621623, 0.21333333333333335, 0.21052631578947367, 0.22077922077922077, 0.2179
4871794871795, 0.21518987341772153, 0.2125, 0.2222222222222222, 0.23170731707317074, 0.2289156626506024, 0.2261904761904762, 0.
2235294117647059, 0.22093023255813954, 0.21839080459770116, 0.2159090909090909, 0.21348314606741572, 0.2111111111111111, 0.2087
912087912088, 0.20652173913043478, 0.20430107526881722, 0.2127659574468085, 0.21052631578947367, 0.21875, 0.2268041237113402,
0.23469387755102042, 0.23232323232323232, 0.24, 0.2376237623762376, 0.24509803921568626, 0.24271844660194175, 0.240384615384615
4, 0.23809523809523808, 0.2358490566037736, 0.2336448598130841, 0.23148148148148148, 0.22935779816513763, 0.22727272727272727,
0.22522522522522523, 0.23214285714285715, 0.23008849557522124, 0.22807017543859648, 0.22608695652173913, 0.22413793103448276,
0.2222222222222222, 0.22033898305084745, 0.2184873949579832, 0.225]
recall after initial query are:  [0.0, 0.0, 0.001, 0.001, 0.002, 0.002, 0.002, 0.002, 0.002, 0.002, 0.002, 0.003, 0.003, 0.003,
0.003, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003, 0.003, 0.004, 0.005, 0.005, 0.005, 0.005, 0.005, 0.005,
0.005, 0.005, 0.005, 0.005, 0.006, 0.007, 0.008, 0.008, 0.008, 0.008, 0.008, 0.009, 0.009, 0.01, 0.011, 0.011, 0.011, 0.011, 0.
012, 0.012, 0.012, 0.012, 0.012, 0.013, 0.013, 0.013, 0.013, 0.013, 0.013, 0.013, 0.014, 0.015, 0.015, 0.015, 0.015, 0.015, 0.0
15, 0.016, 0.016, 0.016, 0.016, 0.016, 0.016, 0.016, 0.017, 0.017, 0.017, 0.017, 0.018, 0.019, 0.019, 0.019, 0.019, 0.019, 0.01
9, 0.019, 0.019, 0.019, 0.019, 0.019, 0.019, 0.019, 0.02, 0.02, 0.021, 0.022, 0.023, 0.023, 0.024, 0.024, 0.024, 0.025, 0.025, 0.025, 0.025,
0.025, 0.025, 0.025, 0.025, 0.025, 0.025, 0.026, 0.026, 0.026, 0.026, 0.026, 0.026, 0.026, 0.026, 0.026, 0.027]
MAP after initial query are:  0.22934867106598114
```

4. Precision-Recall curve for initial query

```
plt.plot(r, p,color="blue")
plt.xlabel("Recall after initial query")
plt.ylabel("Precision after initial query ")
plt.title(" PR curve after initial query")
```

```
Text(0.5, 1.0, ' PR curve after initial query')
```



5. Updated query vector

```
print("New query vector after 1st iteration is \n", new_query1)
```

```
New query vector after 1st iteration is
 {'xref': 0.04641881582376725, 'cantaloupesrvcscmuedu': 0.04641881582376725, 'compgraphics37261': 0.0, 'altgraphics519': 0.0,
'compgraphicsanimation2614': 0.0, 'path': 0, 'cantaloupesrvcscmuedudasnewsharvardeduogicseuwmeduzaphodmpsohiostateedudarwinsu
ranetdtixdtnavymiloasyslipman': 0.0, 'lipmanoasysdtnavymil': 0.0, 'robert': 0, 'lipman': 0.0, 'newsgroups': 0, 'compgraphicsa
ltgraphicscompgraphicsanimation': 0.0, 'subject': 0, 'call': 0.2008103441123436, 'presentation': 0, 'navy': 0.0, 'scivizvr':
0.0, 'seminar': 0.0, 'messageid': 0, '32850oasysdtnavymil': 0.0, 'date': 0, 'nineteen': 0.02601464125578079, 'mar': 0, 'ninet
y-three': 0.08586886564028003, 'two hundred and one thousand and twenty-three': 0.0, 'gmt': 0.019317094933789265, 'articlei
d': 0, 'oasys32850': 0.0, 'expires': 0.0663051456971498, 'thirty': 0, 'apr': 0.009888641048300988, 'forty thousand': 0.0, 're
plyto': 0.02508134953917896, 'followupto': 0.11430391167690865, 'compgraphics': 0, 'distribution': 0.0016143308942439434, 'us
a': 0.022161471048880086, 'organization': 0.0037962989229972377, 'carderock': 0.0, 'division': 0, 'nswc': 0.0, 'bethesda': 0.
0, 'md': 0, 'line': 0.0004190376202763033, 'sixty-five': 0, 'scientific': 0.18285173120684317, 'visualization': 0.41647748953
93735, 'virtual': 0, 'reality': 0, 'tuesday': 0.0, 'june': 0, 'twenty-two': 0.10037662116406999, 'one thousand, nine hundred
and ninety-three': 0.01755618825803461, 'naval': 0, 'surface': 0, 'warfare': 0, 'center': 0, 'formerly': 0, 'david': 0, 'tayl
or': 0.0, 'research': 0, 'maryland': 0, 'sponsor': 0.0, 'ness': 0.0, 'engineering': 0, 'software': 0, 'system': 0.06803736797
731097, 'sponsoring': 0.0, 'oneday': 0.0, 'purpose': 0.019090706581050283, 'present': 0, 'exchange': 0, 'information': 0, 'na
vyrelated': 0.0, 'program': 0, 'development': 0, 'application': 0, 'solicited': 0.0, 'aspect': 0, 'current': 0, 'work': 0.015
7067445030488, 'worksinprogress': 0.0, 'proposed': 0, 'considered': 0.029776080184222155, 'four': 0, 'type': 0.06506669693323
777, 'available': 0.05363694886834501, 'one': 0.02436041375889876, 'regular': 0, 'two thousand and thirty': 0.0, 'minute': 0,
'length': 0.0, 'two': 0.08419917454608669, 'short': 0.08288326850589832, 'ten': 0.010270591051951505, 'three': 0.057298781058
```

6. Result after 1st iteration according to the feedback provided by the user

```
Enter the p % of documents you want to mark as relevant

10
Enter the  12.0 documents you want to mark as relevant

58052,58053,59434,59063,59310,59100,58577,59009,58958,53908,58838,58880
docs after 1st iteration are:
 [('37920', 0.38315756771665926), ('58053', 0.38315756771665926), ('37919', 0.3733355940010124), ('58052', 0.3733355940010124),
('59434', 0.32930743070504265), ('61335', 0.32344157150208214), ('58577', 0.2891398674740878), ('59009', 0.26645503028533946),
('58838', 0.25988023903771484), ('58880', 0.2223859650292101), ('59063', 0.21599111761546877), ('58958', 0.2032508565593242),
('59100', 0.19953560862583586), ('59310', 0.1979435520896827), ('59064', 0.14199962769329252), ('58578', 0.14036366134588524),
('61385', 0.13563659484445817), ('58569', 0.1301205187126554), ('59079', 0.12713872422731146), ('59098', 0.12554063168673396),
('59178', 0.12326147113820354), ('53908', 0.12325089806077934), ('59043', 0.12094733640475523), ('61146', 0.11380490601652109),
('59026', 0.11173177082583817), ('38962', 0.10722632215640937), ('58568', 0.10559658134849136), ('58885', 0.0997300032293047),
('59238', 0.09880582283461589), ('38376', 0.09546949185017922), ('38853', 0.0954175587575038), ('38733', 0.09357995032007879),
('58139', 0.09346583765738749), ('38778', 0.08950115475719621), ('38403', 0.08644484947288884), ('39638', 0.08633741189510578),
('178571', 0.08569470410263828), ('39078', 0.0854964596055852), ('59432', 0.0851344329845478), ('38692', 0.08505367635152554),
('54086', 0.08482678788956906), ('38375', 0.08447327850994812), ('38852', 0.08411902915287381), ('178540', 0.0823693412632762),
('178908', 0.08113597829416992), ('59183', 0.07949069515076329), ('179073', 0.0782433822639824), ('38848', 0.0781667304680535
2), ('38226', 0.07706438087803014), ('38377', 0.07695553455302978), ('59873', 0.07660632181579023), ('38851', 0.076517188957968
58), ('58109', 0.07640503120975536), ('59256', 0.07624716743885383), ('59871', 0.07568657270503688), ('178547', 0.0751883066999
113), ('178546', 0.0750789911279739), ('179058', 0.0749277100876793), ('59872', 0.07442056263632721), ('178293', 0.07422239159
538493), ('58761', 0.07335651368824396), ('59234', 0.07290172021934274), ('176960', 0.07250538462230387), ('59459', 0.072319921
58705375), ('179034', 0.07216404958231631), ('58849', 0.07131655188342603), ('59637', 0.07113439092125211), ('59165', 0.0709240
6683353857), ('60809', 0.07042575204972368), ('60820', 0.06984964346118683), ('61241', 0.06970390468163953), ('178314', 0.06933
335167483869), ('59189', 0.06920302052511135), ('179054', 0.06906996805961152), ('59347', 0.06895915287007821), ('59087', 0.068
77053442361457), ('58984', 0.0687644372559951), ('58916', 0.06744556987306534), ('176944', 0.06740531894489336), ('178340', 0.0
6737108545560185), ('59125', 0.0661242537692248), ('59123', 0.06549966333090923), ('58897', 0.06547721364070849), ('38506', 0.0
6513515200979406), ('60797', 0.06513347956945931), ('59286', 0.06502870873561234), ('59848', 0.06414596418256104), ('61418', 0.
06407280238091624), ('60215', 0.06403869923097104), ('178550', 0.06399053348990143), ('61079', 0.06377244358977349), ('59028',
0.06366957454450863), ('178587', 0.06357713254502513), ('38897', 0.06342615258802849), ('178786', 0.0634253246405319), ('6240
6', 0.06309386693451109), ('38636', 0.06306492152481481), ('53468', 0.0630602106212074), ('178713', 0.06304712446674526), ('592
84', 0.06282822236001505), ('39736', 0.06275857366719975), ('38665', 0.06249515279994551), ('59444', 0.06172153326426129 5), ('5
8781', 0.061370547311072964), ('59162', 0.06127707887361888), ('61456', 0.060774581349658), ('178382', 0.06073555055629337),
('58907', 0.06050273398138676), ('59301', 0.06006254376189866), ('59126', 0.05946242381743862), ('58152', 0.05942064542766401),
('39490', 0.05930293351383553), ('58850', 0.059169448975844315), ('39655', 0.0588465885270352), ('38400', 0.05876549960070804
5), ('59061', 0.05862353517427172), ('59510', 0.058423289460632945), ('58107', 0.058235187373486284), ('178313', 0.057902218145
30203), ('176936', 0.05783066437237118)]
```

7. Star marked relevant documents

```
print("docs after 1st iteration are: ",startlist1)
```

```
docs after 1st iteration are:  ['37920', '58053*', '37919', '58052*', '59434*', '61335', '58577*', '59009*', '58838*', '58880
*', '59063*', '58958*', '59100*', '59310*', '59064', '58578', '61385', '58569', '59079', '59098', '59178', '53908*', '59043',
'61146', '59026', '38962', '58568', '58885', '59238', '38376', '38853', '38733', '58139', '38778', '38403', '39638', '178571',
'39078', '59432', '38692', '54086', '38375', '38852', '178540', '178908', '59183', '179073', '38848', '38226', '38377', '5987
3', '38851', '58109', '59256', '59871', '178547', '178546', '179058', '59872', '178293', '58761', '59234', '176960', '59459',
'179034', '58849', '59637', '59165', '60809', '60820', '61241', '178314', '59189', '179054', '59347', '59087', '58984', '5891
6', '176944', '178340', '59125', '59123', '58897', '38506', '60797', '59286', '59848', '61418', '60215', '178550', '61079', '59
028', '178587', '38897', '178786', '62406', '38636', '53468', '178713', '59284', '39736', '38665', '59444', '58781', '59162',
'61456', '178382', '58907', '59301', '59126', '58152', '39490', '58850', '39655', '38400', '59061', '59510', '58107', '178313',
'176936']
```

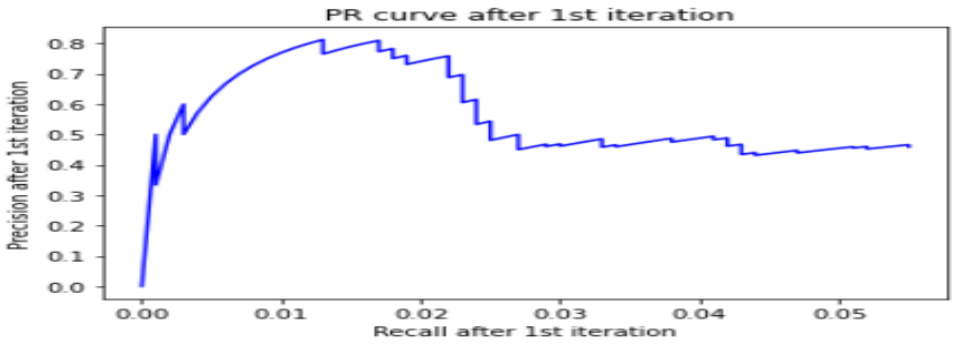
8. MAP, Precision and Recall after 1st iteration

```
precision after 1st iteration are:  [0.0, 0.5, 0.3333333333333333, 0.5, 0.6, 0.5, 0.5714285714285714, 0.625, 0.666666666666666
6, 0.7, 0.7272727272727273, 0.75, 0.7692307692307693, 0.7857142857142857, 0.8, 0.8125, 0.7647058823529411, 0.7777777777777778,
0.7894736842105263, 0.8, 0.8095238095238095, 0.7727272727272727, 0.782608695652174, 0.75, 0.76, 0.7307692307692307, 0.740740740
7407407, 0.75, 0.7586206896551724, 0.7333333333333333, 0.7096774193548387, 0.6875, 0.696969696969697, 0.6764705882352942, 0.657
1428571428571, 0.6388888888888888, 0.6216216216216216, 0.6052631578947368, 0.6153846153846154, 0.6, 0.5853658536585366, 0.57142
85714285714, 0.5581395348837209, 0.5454545454545454, 0.5333333333333333, 0.5434782608695652, 0.5319148936170213, 0.520833333333
3334, 0.5102040816326531, 0.5, 0.49019607843137253, 0.4807692307692308, 0.49056603773584906, 0.5, 0.4909090909090909, 0.4821428
5714285715, 0.47368421052631576, 0.46551724137931033, 0.4576271186440678, 0.45, 0.45901639344262296, 0.46774193548387094, 0.460
3174603174603, 0.46875, 0.46153846153846156, 0.4696969696969697, 0.47761194029850745, 0.4852941176470588, 0.4782608695652174,
0.4714285714285714, 0.4647887323943662, 0.4583333333333333, 0.4657534246575342, 0.4594594594594595, 0.4666666666666667, 0.47368
421052631576, 0.4805194805194805, 0.48717948717948717, 0.4810126582278481, 0.475, 0.48148148148148145, 0.4878048780487805, 0.49
39759036144578, 0.4880952380952381, 0.4823529411764706, 0.4883720930232558, 0.4827586206896552, 0.4772727272727273, 0.471910112
35955055, 0.4666666666666667, 0.46153846153846156, 0.4673913043478261, 0.46236559139784944, 0.4574468085106383, 0.4526315789473
6844, 0.4479166666666667, 0.44329896907216493, 0.4387755102040816, 0.43434343434343436, 0.44, 0.43564356435643564, 0.4313725490
1960786, 0.4368932038834951, 0.4423076923076923, 0.44761904761904764, 0.44339622641509435, 0.4392523364485981, 0.444444444444444
44, 0.44954128440366975, 0.45454545454545453, 0.4594594594594595, 0.45535714285714285, 0.46017699115044247, 0.4561403508771929
6, 0.45217391304347826, 0.45689655172413796, 0.46153846153846156, 0.4661016949152542, 0.46218487394957986, 0.4583333333333333]
recall after 1st iteration are:  [0.0, 0.001, 0.001, 0.001, 0.002, 0.003, 0.003, 0.004, 0.005, 0.006, 0.007, 0.008, 0.009, 0.01, 0.01
1, 0.012, 0.013, 0.013, 0.014, 0.015, 0.016, 0.017, 0.017, 0.018, 0.018, 0.019, 0.019, 0.02, 0.021, 0.022, 0.022, 0.022, 0.022,
0.023, 0.023, 0.023, 0.023, 0.023, 0.023, 0.024, 0.024, 0.024, 0.024, 0.024, 0.024, 0.024, 0.025, 0.025, 0.025, 0.025, 0.025,
0.025, 0.025, 0.026, 0.027, 0.027, 0.027, 0.027, 0.027, 0.027, 0.027, 0.028, 0.029, 0.029, 0.03, 0.03, 0.031, 0.032, 0.033, 0.0
33, 0.033, 0.033, 0.033, 0.034, 0.034, 0.035, 0.036, 0.037, 0.038, 0.038, 0.038, 0.039, 0.04, 0.041, 0.041, 0.041, 0.042, 0.04
2, 0.042, 0.042, 0.042, 0.042, 0.043, 0.043, 0.043, 0.043, 0.043, 0.043, 0.043, 0.043, 0.044, 0.044, 0.044, 0.045, 0.046, 0.04
7, 0.047, 0.047, 0.048, 0.049, 0.05, 0.051, 0.051, 0.052, 0.052, 0.052, 0.053, 0.054, 0.055, 0.055, 0.055]
MAP after 1st iteration are:  0.5756985745719785
```

9. Precision-Recall curve after iteration1

```
import matplotlib.pyplot as plt

plt.plot(r1, p1,color="blue")
plt.xlabel("Recall after 1st iteration")
plt.ylabel("Precision after 1st iteration")
plt.title(" PR curve after 1st iteration")
```

```
Text(0.5, 1.0, ' PR curve after 1st iteration')
```



10. Updated query vector

```
print("New query vector after 2nd iteration is \n", new_query2)
```

```
New query vector after 2nd iteration is
 {'xref': 0.1257484242359618, 'cantaloupesrvcscmuedu': 0.1257484242359618, 'compgraphics37261': 0.0, 'altgraphics519': 0.0,
'compgraphicsanimation2614': 0.0, 'path': 0, 'cantaloupesrvcscmuedudasnewsharvardeduogicseuwmeduzaphodmpsohiostateedudarwinsu
ranetdtixdtnavymiloasyslipman': 0.0, 'lipmanoasysdtnavymil': 0.0, 'robert': 0.010939186276509423, 'lipman': 0.0, 'newsgroup
s': 0, 'compgraphicsaltgraphicscompgraphicsanimation': 0.0, 'subject': 0, 'call': 0.3182762317773609, 'presentation': 0, 'nav
y': 0, 'scivizvr': 0.0, 'seminar': 0.0, 'messageid': 0, '32850oasysdtnavymil': 0.0, 'date': 0, 'nineteen': 0, 'mar': 0.172176
00607981995, 'ninety-three': 0.15043229956207765, 'two hundred and one thousand and twenty-three': 0.0, 'gmt': 0.044505130494
93109, 'articleid': 0.043800333578108704, 'oasys32850': 0.0, 'expires': 0.19111965882260135, 'thirty': 0, 'apr': 0.0190188887
6516991, 'forty thousand': 0.0, 'replyto': 0.07688198639055402, 'followupto': 0.20595526330594505, 'compgraphics': 0, 'distri
bution': 0, 'usa': 0, 'organization': 0.007535571034893798, 'carderock': 0.0, 'division': 0, 'nswc': 0.0, 'bethesda': 0, 'm
d': 0.11291552639915796, 'line': 0.0009676153965457789, 'sixty-five': 0, 'scientific': 0.36060227850676607, 'visualization':
0.2987082821439065, 'virtual': 0, 'reality': 0.033341840550569844, 'tuesday': 0, 'june': 0.039930690477398695, 'twenty-two':
0.1408695393327785, 'one thousand, nine hundred and ninety-three': 0.0332872469790292, 'naval': 0, 'surface': 0.0468583875574
5085, 'warfare': 0, 'center': 0, 'formerly': 0, 'david': 0.006970642868426344, 'taylor': 0, 'research': 0, 'maryland': 0, 'sp
onsor': 0, 'ness': 0.0, 'engineering': 0, 'software': 0, 'system': 0.3573946331146667, 'sponsoring': 0, 'oneday': 0.0, 'purpo
se': 0, 'present': 0.06350181570399316, 'exchange': 0, 'information': 0.3419264861526957, 'navyrelated': 0.0, 'program': 0.02
2940308301213186, 'development': 0.0070497253763939505, 'application': 0, 'solicited': 0.0, 'aspect': 0.039064876902744274,
'current': 0.06272401622737182, 'work': 0.2602131202353888, 'worksinprogress': 0.0, 'proposed': 0.037394091491411374, 'consid
ered': 0.08741321687133699, 'four': 0.2561406706446597, 'type': 0.39835996957656983, 'available': 0.20181881359255036, 'one':
```

11. Result after 2nd iteration according to the feedback provided by the user

```
Enter the  12.0 documents you want to mark as relevant

59064,58578,59043,59432,58139,59183,59871,59189,59459,58849,58897,58152
docs after 2nd iteration are:  [('59183', 0.5008814925300872), ('58578', 0.4890096719180433), ('58152', 0.37691776200689014),
('59871', 0.34813427245948897), ('59165', 0.319492727614625), ('59459', 0.3113440317028118), ('58897', 0.29823311916563355),
('59432', 0.28161101861647253), ('58577', 0.27754893332164954), ('58849', 0.27683557675278125), ('58850', 0.2692998442667209),
('59189', 0.2649828978261223), ('59234', 0.25750136173400706), ('58826', 0.23706884432462927), ('58139', 0.2260541770020511),
('58907', 0.21415385922778787), ('37920', 0.2135014473218872), ('58053', 0.2135014473218872), ('58838', 0.21278489078930793),
('59034', 0.2109849900180376), ('59009', 0.20820587809310306), ('37919', 0.20465825877725002), ('58052', 0.20465825877725002),
('59043', 0.20263061226265314), ('59069', 0.2006914970373562), ('59434', 0.19939698824929353), ('59064', 0.19926176792627417),
('59255', 0.19759739453418249), ('58984', 0.1973593858993986), ('61335', 0.19585837918519877), ('59063', 0.1956149216673861 2),
('59100', 0.19481103655972712), ('59049', 0.19365295729562623), ('59554', 0.18535720346439127), ('58569', 0.184079609502345 73),
('59338', 0.18308620538723402), ('58874', 0.18107555160174654), ('58568', 0.1766254080736332), ('58880', 0.1757666454166503),
('59333', 0.17420724871422139), ('59323', 0.17413822290479897), ('59238', 0.17111630068595562), ('59499', 0.1705093926132536),
('59332', 0.16917776062986292), ('59488', 0.16818762433121856), ('59632', 0.16656658444450958), ('59256', 0.16577679773910314),
('59209', 0.16432551535795337), ('59322', 0.16032735129660639), ('59178', 0.16000637137722273), ('59527', 0.15913929983581854),
('59079', 0.15753760629582808), ('58109', 0.15629767822459348), ('59123', 0.15571855218334565), ('59284', 0.1541691704886206),
('59304', 0.1532014787999353), ('59044', 0.15273507403588688), ('59207', 0.15101558204429633), ('58781', 0.14983074113684847),
('59395', 0.14864775520034795), ('179058', 0.14740501692326902), ('59518', 0.1453149393593674 6), ('59241', 0.1436325192938245
3), ('38403', 0.14280227336968523), ('59202', 0.14274008065415572), ('59848', 0.14266204572409316), ('39638', 0.142486114720676
18), ('59286', 0.14204376804017177), ('39078', 0.14196110122371294), ('59160', 0.1419323770301285), ('61435', 0.1415199196415011
2), ('59380', 0.14101763015724084), ('59548', 0.14052936592304732), ('59905', 0.13963850138949835), ('178571', 0.13760844091031
954), ('178293', 0.13724252787551172), ('59228', 0.13706810794801083), ('59318', 0.13705768897052764), ('59470', 0.134675530741
97242), ('59637', 0.13391071027004595), ('59873', 0.1338002291206712), ('178786', 0.13352431612588397), ('59872', 0.13272165327
279425), ('176960', 0.1321558405201349), ('179073', 0.13093883190616543), ('61146', 0.13041457498514597), ('59456', 0.130246884
44254412), ('38778', 0.12915109664432195), ('59490', 0.12909914271652448), ('59224', 0.12729974796777258), ('61049', 0.12722642
52600432), ('59310', 0.12707265423889533), ('59874', 0.12658543156139385), ('58953', 0.12625262899584866), ('59168', 0.12598553
66210318), ('38853', 0.1257362736511025), ('38376', 0.12571392435326506), ('178908', 0.1256509309682949), ('59098', 0.124400726
20485723), ('58910', 0.12426101729891523), ('58155', 0.12401417998269933), ('59206', 0.12377291419034682), ('59347', 0.12264430
587820517), ('59071', 0.12093979112836524), ('59125', 0.11998827031182324), ('59850', 0.11830927591539118), ('59252', 0.1182048
4005453752), ('58941', 0.11787401988865585), ('178314', 0.11731270991466204), ('58958', 0.11697921337861046), ('59511', 0.11680
72441910079 6), ('61253', 0.11667882807517638), ('59031', 0.11667576403539424), ('60252', 0.11647136534280325), ('59330', 0.1163
737004980421), ('179054', 0.11608118325147 57), ('59203', 0.11573179883549 64), ('178547', 0.11562854191246395), ('59913', 0.1156
0147491988758), ('58891', 0.11453473825568634)]
```

12. Star marked relevant documents

```
print("docs after 2nd iteration are: ",startlist2)
```

```
docs after 2nd iteration are:  ['59183*', '58578*', '58152*', '59871*', '59165', '59459*', '58897*', '59432*', '58577', '58849
*', '58850', '59189*', '59234', '58826', '58139*', '58907', '37920', '58053', '58838', '59034', '59009', '37919', '58052', '590
43*', '59069', '59434', '59064*', '59255', '58984', '61335', '59063', '59100', '59049', '59554', '58569', '59338', '58874', '58
568', '58880', '59333', '59323', '59238', '59499', '59332', '59488', '59632', '59256', '59209', '59322', '59178', '59527', '590
79', '58109', '59123', '59284', '59304', '59044', '59207', '58781', '59395', '179058', '59518', '59241', '38403', '59202', '598
48', '39638', '59286', '39078', '59160', '61435', '59380', '59548', '59905', '178571', '178293', '59228', '59318', '59470', '59
637', '59873', '178786', '59872', '176960', '179073', '61146', '59456', '38778', '59490', '59224', '61049', '59310', '59874',
'58953', '59168', '38853', '38376', '178908', '59098', '58910', '58155', '59206', '59347', '59071', '59125', '59850', '59252',
'58941', '178314', '58958', '59511', '61253', '59031', '60252', '59330', '179054', '59203', '178547', '59913', '58891']
```

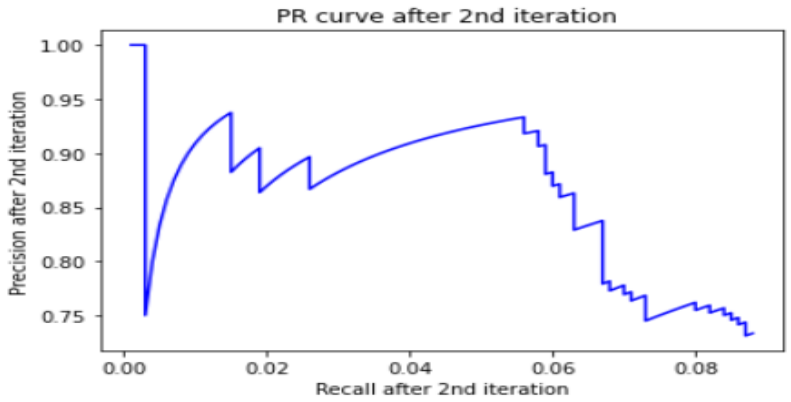
13. MAP, Precision and Recall after 2nd iteration

```
precision after 2nd iteration are:  [1.0, 1.0, 1.0, 0.75, 0.8, 0.8333333333333334, 0.8571428571428571, 0.875, 0.888888888888888
8, 0.9, 0.9090909090909091, 0.9166666666666666, 0.9230769230769231, 0.9285714285714286, 0.9333333333333333, 0.9375, 0.882352941
1764706, 0.8888888888888888, 0.8947368421052632, 0.9, 0.9047619047619048, 0.8636363636363636, 0.8695652173913043, 0.875, 0.88,
0.8846153846153846, 0.8888888888888888, 0.8928571428571429, 0.896551724137931, 0.8666666666666667, 0.8709677419354839, 0.875,
0.8787878787878788, 0.8823529411764706, 0.8857142857142857, 0.8888888888888888, 0.8918918918918919, 0.8947368421052632, 0.89743
58974358975, 0.9, 0.9024390243902439, 0.9047619047619048, 0.9069767441860465, 0.9090909090909091, 0.9111111111111111, 0.9130434
782608695, 0.9148936170212766, 0.9166666666666666, 0.9183673469387755, 0.92, 0.9215686274509803, 0.9230769230769231, 0.92452830
18867925, 0.9259259259259259, 0.9272727272727272, 0.9285714285714286, 0.9298245614035088, 0.9310344827586207, 0.932203389830508
4, 0.9333333333333333, 0.9180327868852459, 0.9193548387096774, 0.9206349206349206, 0.90625, 0.9076923076923077, 0.8939393939393
939, 0.8805970149253731, 0.8823529411764706, 0.8695652173913043, 0.8714285714285714, 0.8591549295774648, 0.8611111111111112, 0.
863013698630137, 0.8513513513513513, 0.84, 0.8289473684210527, 0.8311688311688312, 0.8333333333333334, 0.8354430379746836, 0.83
75, 0.8271604938271605, 0.8170731707317073, 0.8072289156626506, 0.7976190476190477, 0.788235294117647, 0.7790697674418605, 0.78
16091954022989, 0.7727272727272727, 0.7752808988764045, 0.7777777777777778, 0.7692307692307693, 0.7717391304347826, 0.763440860
2150538, 0.7659574468085106, 0.7684210526315789, 0.7604166666666666, 0.7525773195876289, 0.7448979591836735, 0.747474747474747
5, 0.75, 0.7524752475247525, 0.7549019607843137, 0.7572815533980582, 0.7596153846153846, 0.7619047619047619, 0.754716981132075
5, 0.7570093457943925, 0.7592592592592593, 0.7522935779816514, 0.7545454545454545, 0.7567567567567568, 0.75, 0.752212389380530
9, 0.7456140350877193, 0.7478260869565218, 0.7413793103448276, 0.7435897435897436, 0.7372881355932204, 0.7310924369747899, 0.73
33333333333333]


recall after 2nd iteration are:  [0.001, 0.002, 0.003, 0.003, 0.004, 0.005, 0.006, 0.007, 0.008, 0.009, 0.01, 0.011, 0.012, 0.0
13, 0.014, 0.015, 0.015, 0.016, 0.017, 0.018, 0.019, 0.019, 0.02, 0.021, 0.022, 0.023, 0.024, 0.025, 0.026, 0.026, 0.027, 0.02
8, 0.029, 0.03, 0.031, 0.032, 0.033, 0.034, 0.035, 0.036, 0.037, 0.038, 0.039, 0.04, 0.041, 0.042, 0.043, 0.044, 0.045, 0.046,
0.047, 0.048, 0.049, 0.05, 0.051, 0.052, 0.053, 0.054, 0.055, 0.056, 0.056, 0.057, 0.058, 0.058, 0.059, 0.059, 0.059, 0.06, 0.0
6, 0.061, 0.061, 0.062, 0.063, 0.063, 0.063, 0.063, 0.064, 0.065, 0.066, 0.067, 0.067, 0.067, 0.067, 0.067, 0.067, 0.067, 0.06
8, 0.068, 0.069, 0.07, 0.07, 0.071, 0.071, 0.072, 0.073, 0.073, 0.073, 0.073, 0.074, 0.075, 0.076, 0.077, 0.078, 0.079, 0.08,
0.08, 0.081, 0.082, 0.082, 0.083, 0.084, 0.084, 0.085, 0.085, 0.086, 0.086, 0.087, 0.087, 0.087, 0.088]


MAP after 2nd iteration is:  0.8662607082356256
```

14. Precision-Recall curve after iteration 2

```
plt.plot(r2, p2,color="blue")
plt.xlabel("Recall after 2nd iteration")
plt.ylabel("Precision after 2nd iteration")
plt.title(" PR curve after 2nd iteration")
```

```
Text(0.5, 1.0, ' PR curve after 2nd iteration')
```



15. Updated query vector

```
print("New query vector after 3rd iteration is \n", new_query3)
```

```
New query vector after 3rd iteration is
 {'xref': 0.2066655319588635, 'cantaloupesrvcscmuedu': 0.2066655319588635, 'compgraphics37261': 0.0, 'altgraphics519': 0.0,
'compgraphicsanimation2614': 0.0, 'path': 0, 'cantaloupesrvcscmuedudasnewsharvardeduogicseuwmeduzaphodmpsohiostateedudarwinsu
ranetdtixdtnavymiloasyslipman': 0.0, 'lipmanoasysdtnavymil': 0.0, 'robert': 0.11644175071066559, 'lipman': 0.0, 'newsgroups':
0, 'compgraphicsaltgraphicscompgraphicsanimation': 0.0, 'subject': 0, 'call': 0.5878809485751273, 'presentation': 0.058207030
30953499, 'navy': 0, 'scivizvr': 0.0, 'seminar': 0.0, 'messageid': 0, '32850oasysdtnavymil': 0.0, 'date': 0, 'nineteen': 0.09
130959430058483, 'mar': 0.2486925993721647, 'ninety-three': 0.193081354738755, 'two hundred and one thousand and twenty-thre
e': 0.0, 'gmt': 0.07402546984839714, 'articleid': 0.03341290269001968, 'oasys32850': 0.0, 'expires': 0.3869018644567147, 'thi
rty': 0.029089527955846184, 'apr': 0.030256584762795684, 'forty thousand': 0.0, 'replyto': 0.07772821749028264, 'followupto':
0.3510290605369331, 'compgraphics': 0, 'distribution': 0.0683645917638605, 'usa': 0.1207505393954661, 'organization': 0.01150
287332393969, 'carderock': 0.0, 'division': 0, 'nswc': 0.0, 'bethesda': 0, 'md': 0.19125100698861092, 'line': 0.0013485695644
955614, 'sixty-five': 0, 'scientific': 0.35434961104406115, 'visualization': 0.23067075215850447, 'virtual': 0, 'reality': 0,
'tuesday': 0, 'june': 0, 'twenty-two': 0.22200612756346993, 'one thousand, nine hundred and ninety-three': 0.0871320412755081
4, 'naval': 0.0761358835006584, 'surface': 0, 'warfare': 0.0, 'center': 0.2820629625028808, 'formerly': 0, 'david': 0.0306396
68437123083, 'taylor': 0, 'research': 0.12008888067191678, 'maryland': 0, 'sponsor': 0, 'ness': 0.0, 'engineering': 0.0770472
1886487481, 'software': 0.02804202424753259, 'system': 0.49278897084136836, 'sponsoring': 0.0, 'oneday': 0.0, 'purpose': 0,
'present': 0.17622269454419506, 'exchange': 0.06203051008676764, 'information': 0.48587999652233216, 'navyrelated': 0.0, 'pro
gram': 0.14295659708155353, 'development': 0.13694510508603536, 'application': 0.11686470357580643, 'solicited': 0.0, 'aspec
t': 0.07169208131499015, 'current': 0.16073045790961632, 'work': 0.25079459915632607, 'worksinprogress': 0.0, 'proposed': 0.0
04585574493548704, 'considered': 0.11230002560120551, 'four': 0.35414873399428837, 'type': 0.4903770607980308, 'available':
```

16. Result after 3rd iteration according to the feedback provided by the user

```
Enter the  12.0 documents you want to mark as relevant

58577,58838,58984,59049,59256,59241,59202,59873,59310,59874,59098,59031
docs after 3rd iteration are:  [('59874', 0.43792731001518076), ('58578', 0.42600457231324373), ('59183', 0.42300062363883045),
('59873', 0.37500597719055856), ('59871', 0.35884957274884494), ('58897', 0.3261536979063995), ('58984', 0.32025344140969003),
('58577', 0.31842130000655927), ('59256', 0.3084064732997529), ('58152', 0.3046479024821894), ('58907', 0.2868579133869832),
('59049', 0.275727436493768), ('58850', 0.27268437654485167), ('59069', 0.27107368740575216), ('59165', 0.27033520652606596),
('58849', 0.25660030084100416), ('58838', 0.25543257034906725), ('59034', 0.2551376057227737), ('59459', 0.2399595209161226),
('59202', 0.2366611582849465), ('58826', 0.23362607142971384), ('59189', 0.23160310860478728), ('59234', 0.23109129833543174),
('59432', 0.22568690544655626), ('59241', 0.2235506991300974), ('59009', 0.211144761253313), ('59031', 0.2095472605628554 1),
('59168', 0.19951498954511881), ('58139', 0.19411579535970133), ('58569', 0.19391319676487812), ('58568', 0.1928091381110123),
('59063', 0.19188551752881272), ('59872', 0.18777849340058092), ('59255', 0.18618755495663286), ('58874', 0.1841480677084916),
('38853', 0.1785296703660058), ('38376', 0.17844741485783733), ('59848', 0.17776914493285584), ('59043', 0.1774930372177256),
('59913', 0.17641372950069134), ('38778', 0.17594844973716414), ('59905', 0.17501809242186436), ('59380', 0.17173289619866955),
('37920', 0.17129644803689284), ('58053', 0.17129644803689284), ('61435', 0.17076407590669168), ('58781', 0.16868860148126588),
('59098', 0.16783793795058088), ('59338', 0.167041953006311), ('38403', 0.1668853524550255), ('39638', 0.16677048455485924),
('59434', 0.1666448608126465), ('39078', 0.16645526211561842), ('61146', 0.1641674910771591), ('61335', 0.1638190985475084),
('59284', 0.16379536815491413), ('59870', 0.16315078395380248), ('59252', 0.16223195336222354), ('59160', 0.16222690002414625),
('59554', 0.1613653778635953), ('59123', 0.16117456705656616), ('38377', 0.15968426490452264), ('38851', 0.1595943419756752),
('59310', 0.15871906409593511), ('59028', 0.15854964336461727), ('59238', 0.1578503693132087), ('37919', 0.1576669555486288),
('58052', 0.1576669555486288), ('59527', 0.15741130672330883), ('59209', 0.15691375404641655), ('59286', 0.15616290191212204),
('59499', 0.15392752216792502), ('59064', 0.1528115464663774), ('179058', 0.15254715713921552), ('59488', 0.1520364622859636),
('61253', 0.15158162587137955), ('59323', 0.15157342153281306), ('58833', 0.15036244753177316), ('58917', 0.15026903650381368),
('59333', 0.14997605183080828), ('59179', 0.14943474892375644), ('59044', 0.1489181892792022), ('38375', 0.14827819878310242),
('38852', 0.1479480596753659), ('176960', 0.14716216690814954), ('59100', 0.146808864414391), ('58109', 0.14621475390266842),
('61207', 0.14588401091460398), ('59332', 0.1456391248002023), ('178571', 0.14501546109081906), ('58759', 0.14485182594141766),
('59849', 0.14444008885570375), ('59850', 0.14325084661608906), ('59908', 0.14313931493296808), ('60103', 0.14213319018313966),
('178293', 0.14101237422279314), ('59271', 0.1408773151029035), ('59322', 0.14086765964630632), ('59303', 0.14064093365848704),
('59632', 0.14061436418431064), ('179073', 0.14051884330274805), ('58880', 0.14049408289381732), ('61049', 0.1397556992701222
4), ('59347', 0.13782989289982275), ('59228', 0.13707944594113203), ('59073', 0.1367418613420932), ('178786', 0.136410636970604
46), ('59548', 0.13600650650270463), ('59518', 0.13458309460162984), ('59909', 0.13448810352229074), ('59904', 0.13408310592305
586), ('38733', 0.13390598914980412), ('59237', 0.13373480587449688), ('59113', 0.13371954988021784), ('58834', 0.1336217435030
1215), ('178908', 0.1332114979748799), ('59126', 0.13260681806892663), ('59125', 0.13259901298913088), ('59071', 0.132258840333
50394), ('61244', 0.1317968018402353)]
```

17. Star marked relevant documents

```
print("docs after 3rd iteration are: ",startlist3)
```

```
docs after 3rd iteration are:  ['59874*', '58578', '59183', '59873*', '59871', '58897', '58984*', '58577*', '59256*', '58152',
'58907', '59049*', '58850', '59069', '59165', '58849', '58838*', '59034', '59459', '59202*', '58826', '59189', '59234', '5943
2', '59241*', '59009', '59031*', '59168', '58139', '58569', '58568', '59063', '59872', '59255', '58874', '38853', '38376', '598
48', '59043', '59913', '38778', '59905', '59380', '37920', '58053', '61435', '58781', '59098*', '59338', '38403', '39638', '594
34', '39078', '61146', '61335', '59284', '59870', '59252', '59160', '59554', '59123', '38377', '38851', '59310*', '59028', '592
38', '37919', '58052', '59527', '59209', '59286', '59499', '59064', '179058', '59488', '61253', '59323', '58833', '58917', '593
33', '59179', '59044', '38375', '38852', '176960', '59100', '58109', '61207', '59332', '178571', '58759', '59849', '59850', '59
908', '60103', '178293', '59271', '59322', '59303', '59632', '179073', '58880', '61049', '59347', '59228', '59073', '178786',
'59548', '59518', '59909', '59904', '38733', '59237', '59113', '58834', '178908', '59126', '59125', '59071', '61244']
```

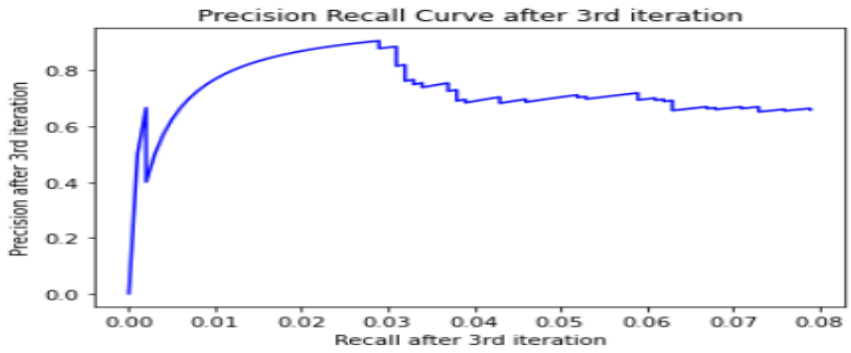
18. MAP, Precision and Recall after 3rd iteration

```
precision after 3rd iteration are:  [0.0, 0.5, 0.6666666666666666, 0.5, 0.4, 0.5, 0.5714285714285714, 0.625, 0.6666666666666666, 0.7, 0.7272727272727273, 0.75, 0.7692307692307693, 0.7857142857142857, 0.8, 0.8125, 0.8235294117647058, 0.8333333333333334, 0.8421052631578947, 0.85, 0.8571428571428571, 0.8636363636363636, 0.8695652173913043, 0.875, 0.88, 0.8846153846153846, 0.8888888888888888, 0.8928571428571429, 0.896551724137931, 0.9, 0.9032258064516129, 0.90625, 0.8787878787878788, 0.8823529411764706, 0.8857142857142857, 0.8611111111111112, 0.8378378378378378, 0.8157894736842105, 0.8205128205128205, 0.8, 0.7804878048780488, 0.7619047619047619, 0.7674418604651163, 0.75, 0.7555555555555555, 0.7391304347826086, 0.7446808510638298, 0.75, 0.7551020408163265, 0.74, 0.7254901960784313, 0.7307692307692307, 0.7169811320754716, 0.7037037037037037, 0.6909090909090909, 0.6964285714285714, 0.6842105263157895, 0.6896551724137931, 0.6949152542372882, 0.7, 0.7049180327868853, 0.6935483870967742, 0.6825396825396826, 0.6875, 0.6923076923076923, 0.696969696969697, 0.6865671641791045, 0.6911764705882353, 0.6956521739130435, 0.7, 0.70422535211267 6, 0.7083333333333334, 0.7123287671232876, 0.7027027027027027, 0.7066666666666667, 0.6973684210526315, 0.7012987012987013, 0.7051282051282052, 0.7088607594936709, 0.7125, 0.7160493827160493, 0.7195121951219512, 0.7108433734939759, 0.7023809523809523, 0.6941176470588235, 0.6976744186046512, 0.7011494252873564, 0.6931818181818182, 0.6966292134831461, 0.6888888888888889, 0.6923076923076923, 0.6847826086956522, 0.6774193548387096, 0.6702127659574468, 0.6631578947368421, 0.65625, 0.6597938144329897, 0.6632653061224489, 0.6666666666666666, 0.67, 0.6633663366336634, 0.6666666666666666, 0.6601941747572816, 0.6634615384615384, 0.6666666666666666, 0.6698113207547169, 0.6635514018691588, 0.6666666666666666, 0.6697247706422018, 0.6636363636363637, 0.65765765765765 77, 0.6517857142857143, 0.6548672566371682, 0.6578947368421053, 0.6608695652173913, 0.6551724137931034, 0.6581196581196581, 0.6610169491525424, 0.6638655462184874, 0.6583333333333333]


recall after 3rd iteration are:  [0.0, 0.001, 0.002, 0.002, 0.002, 0.003, 0.004, 0.005, 0.006, 0.007, 0.008, 0.009, 0.01, 0.01 1, 0.012, 0.013, 0.014, 0.015, 0.016, 0.017, 0.018, 0.019, 0.02, 0.021, 0.022, 0.023, 0.024, 0.025, 0.026, 0.027, 0.028, 0.029, 0.029, 0.03, 0.031, 0.031, 0.031, 0.031, 0.032, 0.032, 0.032, 0.032, 0.033, 0.033, 0.034, 0.034, 0.035, 0.036, 0.037, 0.037, 0.037, 0.038, 0.038, 0.038, 0.038, 0.039, 0.039, 0.04, 0.041, 0.042, 0.043, 0.043, 0.043, 0.044, 0.045, 0.046, 0.046, 0.047, 0.048, 0.049, 0.05, 0.051, 0.052, 0.052, 0.053, 0.053, 0.054, 0.055, 0.056, 0.057, 0.058, 0.059, 0.059, 0.059, 0.06, 0.061, 0.061, 0.062, 0.062, 0.063, 0.063, 0.063, 0.063, 0.063, 0.063, 0.064, 0.065, 0.066, 0.067, 0.067, 0.068, 0.068, 0.069, 0.07, 0.071, 0.071, 0.072, 0.073, 0.073, 0.073, 0.073, 0.074, 0.075, 0.076, 0.076, 0.077, 0.078, 0.079, 0.079]


MAP after 3rd iteration is:  0.7352513287724209
```

19. Precision-Recall curve after iteration 3

```
plt.plot(r3, p3,color="blue")
plt.xlabel("Recall after 3rd iteration")
plt.ylabel("Precision after 3rd iteration")
plt.title("Precision Recall Curve after 3rd iteration")
```

```
Text(0.5, 1.0, 'Precision Recall Curve after 3rd iteration')
```



20. Updated query vector

```
print("New query vector after 4th iteration is \n", new_query4)
```

```
New query vector after 4th iteration is
 {'xref': 0.23722162356030996, 'cantaloupesrvcscmuedu': 0.23722162356030996, 'compgraphics37261': 0.0, 'altgraphics519': 0.0, 'compgraphicsanimation2614': 0.0, 'path': 0, 'cantaloupesrvcscmuedudasnewsharvardeduogicseuwmeduzaphodmpsohiostateedudarwinsuranetdtixdtnavymiloasyslipman': 0.0, 'lipmanoasysdtnavymil': 0.0, 'robert': 0.06682308386150337, 'lipman': 0.0, 'newsgroups': 0, 'compgraphicsaltgraphicscompgraphicsanimation': 0.0, 'subject': 0, 'call': 0.7369963594030335, 'presentation': 0.035162657 93912807, 'navy': 0, 'scivizvr': 0.0, 'seminar': 0.0, 'messageid': 0, '32850oasysdtnavymil': 0.0, 'date': 0, 'nineteen': 0.09 527142891286816, 'mar': 0.1479300108598552, 'ninety-three': 0.2779769268118479, 'two hundred and one thousand and twenty-three': 0.0, 'gmt': 0.09221946369333615, 'articleid': 0.020574504963168084, 'oasys32850': 0.0, 'expires': 0.3979219876819391, 'thirty': 0, 'apr': 0.03918332855748826, 'forty thousand': 0.0, 'replyto': 0.10366662692915551, 'followupto': 0.432219418500445 7, 'compgraphics': 0, 'distribution': 0.03971758649356299, 'usa': 0.10427857578961246, 'organization': 0.014881823285231398, 'carderock': 0.0, 'division': 0, 'nswc': 0.0, 'bethesda': 0, 'md': 0.08783557751256268, 'line': 0.0016923432639366983, 'sixty-five': 0, 'scientific': 0.40294977160535445, 'visualization': 0.6029738867576298, 'virtual': 0, 'reality': 0, 'tuesday': 0, 'june': 0, 'twenty-two': 0.3067190226419948, 'one thousand, nine hundred and ninety-three': 0.08580371631476347, 'naval': 0.05988830994623773, 'surface': 0, 'warfare': 0.0, 'center': 0.10845931957341826, 'formerly': 0, 'david': 0, 'taylor': 0, 'research': 0, 'maryland': 0, 'sponsor': 0, 'ness': 0.0, 'engineering': 0.012499538144417094, 'software': 0, 'system': 0.4129048514 9832445, 'sponsoring': 0.0, 'oneday': 0.0, 'purpose': 0, 'present': 0.059635339654768194, 'exchange': 0.024048731928289158, 'information': 0.31588320562348265, 'navyrelated': 0.0, 'program': 0, 'development': 0.004797760781731553, 'application': 0, 'solicited': 0.0, 'aspect': 0.022871143472749912, 'current': 0.045187256395293 86, 'work': 0.1554048288246669, 'worksinprogress': 0.0, 'proposed': 0, 'considered': 0.09206391082271852, 'four': 0.22148349416242435, 'type': 0.46698088568226664, 'availab
```

21. Result after 4th iteration according to the feedback provided by the user

```
Enter the  12.0 documents you want to mark as relevant

59905,58053,58781,38403,59284,59064,59323,59333,59044,59303,59347,59347
docs after 4th iteration are:  [('58578', 0.3972651637886252 5), ('58577', 0.3671117085865749), ('59874', 0.363176999775656),
('59183', 0.351837219593905), ('59873', 0.3041894595779897), ('59871', 0.3003435295750512), ('58838', 0.29991724793265545), ('5
8897', 0.2825443901149836), ('58984', 0.28004574639998897), ('59256', 0.27250040563289657), ('58152', 0.26829852017324995), ('5
9009', 0.2610027636341577), ('37920', 0.25344717599450417), ('58053', 0.25344717599450417), ('58907', 0.24879305714804614), ('3
7919', 0.2387690103863361), ('58052', 0.2387690103863361), ('58850', 0.23304273976698422), ('59063', 0.23146044135556867), ('58
849', 0.22672801851219118), ('59049', 0.22589936458848497), ('59434', 0.22571003570567555), ('59069', 0.22196594749481993), ('6
1335', 0.22105593508041366), ('59165', 0.2205981404025266), ('59459', 0.21845766638197653), ('59432', 0.2108214022229425), ('59
034', 0.2052910832633891), ('59202', 0.20118913248225667), ('58826', 0.20050998821947633), ('59189', 0.19879326283654256), ('59
310', 0.19767471307146162), ('59234', 0.19734925455437363), ('59031', 0.19024680783316253), ('59100', 0.18915291853035537), ('5
9043', 0.18727667746066384), ('59241', 0.18508191946683805), ('58880', 0.184845422358848), ('59098', 0.18283420459097635), ('58
139', 0.17512341062372125), ('58569', 0.17423438685084938), ('59064', 0.1741121447511684), ('58568', 0.17286683344848366), ('59
168', 0.16941439101629255), ('59238', 0.16040324864361719), ('58874', 0.15799541017272525), ('59255', 0.15793811270260302), ('5
8781', 0.15033886024477428), ('59872', 0.1497503791425163), ('59028', 0.14757945264090336), ('61146', 0.14710356102568733), ('5
9380', 0.14681053630219076), ('59079', 0.14470594653231778), ('58958', 0.14142074391723009), ('59178', 0.1406537891599706), ('5
9026', 0.1367933624627457), ('38853', 0.13587967545721513), ('38376', 0.1358560264351622), ('38778', 0.13526798694624664), ('59
160', 0.13481270142408813), ('58833', 0.13477063657891306), ('59848', 0.13461145445079012), ('59913', 0.1322989877572268), ('59
252', 0.1321325699775614), ('59338', 0.13195862905387615), ('58917', 0.13091380642127542), ('59554', 0.12776142853538056), ('58
109', 0.12753973653532943), ('59905', 0.12720669122489592), ('59870', 0.1265327494343464), ('59209', 0.12637525657702972), ('59
527', 0.12634216874505108), ('58759', 0.1260373943797616), ('59284', 0.12499028618456266), ('39638', 0.12476300737255602), ('38
403', 0.124685521144783), ('39078', 0.12410496221322186), ('61435', 0.12379291284314475), ('59123', 0.12302425916447576), ('383
77', 0.12283643781059196), ('59044', 0.12279340864219117), ('38851', 0.12270235497789238), ('59303', 0.12263261896311876), ('59
323', 0.12237991855753323), ('59179', 0.1219660111358916), ('59113', 0.11959797489786728), ('59499', 0.11932875757837387), ('59
488', 0.11844102235847785), ('59073', 0.1172519951603392), ('59286', 0.1168150153365477), ('179058', 0.11656424020261026), ('38
733', 0.11601888622933283), ('59271', 0.11551069520085634), ('58834', 0.11533860515104417), ('178571', 0.11495254588795951),
('176960', 0.11435717797353152), ('59363', 0.11430203828368124), ('59131', 0.1123815750129402), ('59133', 0.11234120179762597),
('59228', 0.11168141196823254), ('59347', 0.11134864612138402), ('58785', 0.11086700050719935), ('179073', 0.1108051124575039
5), ('59333', 0.11041818561771928), ('61049', 0.11035625483490241), ('178786', 0.11023628464121557), ('59071', 0.11023199083488
14), ('59024', 0.11004317291974391), ('59237', 0.1094747727905021), ('38375', 0.10943784451121318), ('38852', 0.109154269124939
45), ('178293', 0.10877115104977877), ('59332', 0.10831397647472624), ('59359', 0.1075583889712474), ('59632', 0.10737869243332
986), ('59908', 0.10737825512287116), ('60103', 0.1072578632628505), ('178908', 0.10724912490111287), ('59637', 0.1070602903852
3086), ('61253', 0.10623852445121833)]
```

22. Star marked relevant documents

```
docs after 4th iteration are:  ['58578', '58577*', '59874', '59183', '59873', '59871', '58838*', '58897', '58984', '59256', '58
152', '59009*', '37920', '58053*', '58907', '37919', '58052*', '58850', '59063*', '58849', '59049', '59434*', '59069', '61335',
'59165', '59459', '59432', '59034', '59202', '58826', '59189', '59310*', '59234', '59031', '59100*', '59043', '59241', '58880
*', '59098', '58139', '58569', '59064', '58568', '59168', '59238', '58874', '59255', '58781', '59872', '59028', '61146', '5938
0', '59079', '58958*', '59178', '59026', '38853', '38376', '38778', '59160', '58833', '59848', '59913', '59252', '59338', '5891
7', '59554', '58109', '59905', '59870', '59209', '59527', '58759', '59284', '39638', '38403', '39078', '61435', '59123', '3837
7', '59044', '38851', '59303', '59323', '59179', '59113', '59499', '59488', '59073', '59286', '179058', '38733', '59271', '5883
4', '178571', '176960', '59363', '59131', '59133', '59228', '59347', '58785', '179073', '59333', '61049', '178786', '59071', '5
9024', '59237', '38375', '38852', '178293', '59332', '59359', '59632', '59908', '60103', '178908', '59637', '61253']
```

23. Final result with all the marked relevant documents by user (docs are marked with ‘*’)

```
print("Relevant documents after 4 iterations are: \n ",final_star)
```

```
Relevant documents after 4 iterations are:
  ['58578*', '58577*', '59874*', '59183*', '59873*', '59871*', '58838*', '58897*', '58984*', '59256*', '58152*', '59009*', '379
20', '58053*', '58907', '37919', '58052*', '58850', '59063*', '58849*', '59049*', '59434*', '59069', '61335', '59165', '59459
*', '59432*', '59034', '59202*', '58826', '59189*', '59310*', '59234', '59031*', '59100*', '59043*', '59241*', '58880*', '59098
*', '58139*', '58569', '59064*', '58568', '59168', '59238', '58874', '59255', '58781', '59872', '59028', '61146', '59380', '590
79', '58958*', '59178', '59026', '38853', '38376', '38778', '59160', '58833', '59848', '59913', '59252', '59338', '58917', '595
54', '58109', '59905', '59870', '59209', '59527', '58759', '59284', '39638', '38403', '39078', '61435', '59123', '38377', '5904
4', '38851', '59303', '59323', '59179', '59113', '59499', '59488', '59073', '59286', '179058', '38733', '59271', '58834', '1785
71', '176960', '59363', '59131', '59133', '59228', '59347', '58785', '179073', '59333', '61049', '178786', '59071', '59024', '5
9237', '38375', '38852', '178293', '59332', '59359', '59632', '59908', '60103', '178908', '59637', '61253']
```

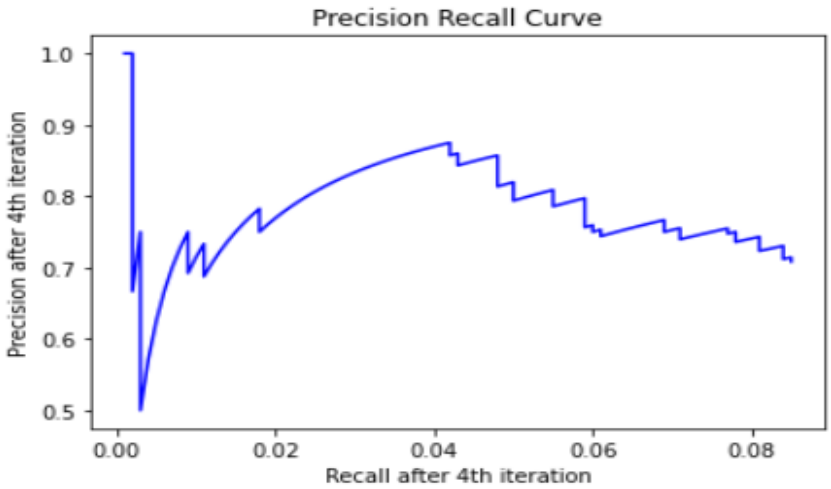
24. MAP, Precision and Recall after 4th iteration

```
precision after 4th iteration are:  [1.0, 1.0, 0.6666666666666666, 0.75, 0.6, 0.5, 0.5714285714285714, 0.625, 0.6666666666666666, 0.7, 0.7272727272727273, 0.75, 0.6923076923076923, 0.7142857142857143, 0.7333333333333333, 0.6875, 0.7058823529411765, 0.7222222222222222, 0.7368421052631579, 0.75, 0.7619047619047619, 0.7727272727272727, 0.782608695652174, 0.75, 0.76, 0.7692307692307693, 0.7777777777777778, 0.7857142857142857, 0.7931034482758621, 0.8, 0.8064516129032258, 0.8125, 0.8181818181818182, 0.8235294117647058, 0.8285714285714286, 0.8333333333333334, 0.8378378378378378, 0.8421052631578947, 0.8461538461538461, 0.85, 0.8536585365853658, 0.8571428571428571, 0.8604651162790697, 0.8636363636363636, 0.8666666666666667, 0.8695652173913043, 0.8723404255319149, 0.875, 0.8571428571428571, 0.86, 0.8431372549019608, 0.8461538461538461, 0.8490566037735849, 0.8518518518518519, 0.8545454545454545, 0.8571428571428571, 0.8421052631578947, 0.8275862068965517, 0.8135593220338984, 0.8166666666666667, 0.819672131147541, 0.8064516129032258, 0.7936507936507936, 0.796875, 0.8, 0.803030303030303, 0.8059701492537313, 0.8088235294117647, 0.7971014492753623, 0.7857142857142857, 0.7887323943661971, 0.7916666666666666, 0.7945205479452054, 0.7972972972972973, 0.7866666666666666, 0.7763157894736842, 0.7662337662337663, 0.7564102564102564, 0.75493670886076, 0.75, 0.7530864197530864, 0.7439024390243902, 0.7469879518072289, 0.75, 0.7529411764705882, 0.7558139534883721, 0.7586206896551724, 0.7613636363636364, 0.7640449438202247, 0.7666666666666667, 0.7582417582417582, 0.75, 0.7526881720430108, 0.7553191489361702, 0.7473684210526316, 0.7395833333333334, 0.7422680412371134, 0.7448979591836735, 0.7474747474747475, 0.75, 0.7524752475247525, 0.7549019607843137, 0.7475728155339806, 0.75, 0.7428571428571429, 0.7358490566037735, 0.7383177570093458, 0.7407407407407407, 0.7431192660550459, 0.7363636363636363, 0.7297297297297297, 0.7232142857142857, 0.7256637168141593, 0.7280701754385965, 0.7304347826086957, 0.7241379310344828, 0.717948717948718, 0.711864406779661, 0.7142857142857143, 0.7083333333333334]


recall after 4th iteration are:  [0.001, 0.002, 0.002, 0.003, 0.003, 0.003, 0.004, 0.005, 0.006, 0.007, 0.008, 0.009, 0.009, 0.01, 0.011, 0.011, 0.012, 0.013, 0.014, 0.015, 0.016, 0.017, 0.018, 0.018, 0.019, 0.02, 0.021, 0.022, 0.023, 0.024, 0.025, 0.026, 0.027, 0.028, 0.029, 0.03, 0.031, 0.032, 0.033, 0.034, 0.035, 0.036, 0.037, 0.038, 0.039, 0.04, 0.041, 0.042, 0.042, 0.043, 0.043, 0.044, 0.045, 0.046, 0.047, 0.048, 0.048, 0.048, 0.048, 0.049, 0.05, 0.05, 0.05, 0.051, 0.052, 0.053, 0.054, 0.055, 0.055, 0.055, 0.056, 0.057, 0.058, 0.059, 0.059, 0.059, 0.059, 0.059, 0.06, 0.06, 0.061, 0.061, 0.062, 0.063, 0.064, 0.065, 0.066, 0.067, 0.068, 0.069, 0.069, 0.069, 0.07, 0.071, 0.071, 0.071, 0.072, 0.073, 0.074, 0.075, 0.076, 0.077, 0.077, 0.078, 0.078, 0.078, 0.079, 0.08, 0.081, 0.081, 0.081, 0.081, 0.082, 0.083, 0.084, 0.084, 0.084, 0.084, 0.085, 0.085]


MAP after 4th iteration is:  0.7853508503309906
```

25. Precision-Recall curve after iteration 4

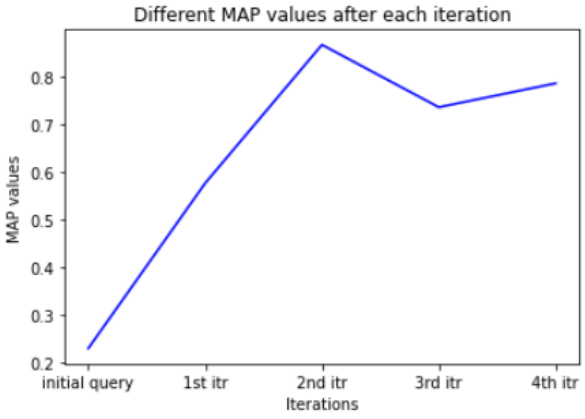
```
plt.plot(r4, p4,color="blue")
plt.xlabel("Recall after 4th iteration")
plt.ylabel("Precision after 4th iteration")
plt.title("Precision Recall Curve")
```

```
Text(0.5, 1.0, 'Precision Recall Curve')
```



26. MAP at the end of 4 iterations

```
[0.22934867106598114, 0.5756985745719785, 0.8662607082356256, 0.7352513287724209, 0.7853508503309906]

Text(0.5, 1.0, 'Different MAP values after each iteration')
```



27. T-SNE plot for all the query vectors

```
plt.figure(figsize=(7,7))
for i in range(len(x)):
    plt.scatter(x[i],y[i])
    plt.annotate(labels[i],xy=(x[i], y[i]),xytext=(5, 5),textcoords='offset points',ha='right',va='bottom')
plt.show()
```