Patrick Connolly

Assignment 3

For more details see results.txt

Query 1:

Using Query: KENNEDY ADMINISTRATION PRESSURE ON NGO DINH DIEM TO STOP

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Iteration | Precision | Recall | Average Precision | MAP |
| 0 | 4/10 | 4/7 | 4/9+3/7+1/4+2/6=0.208049887 | 0.208049887 |
| 1 | 7/10 | 7/7 | 1/1+2/2+3/3+4/4+5/5+6/6+7/7=1 | 0.604024944 |
| 2 | 7/10 | 7/7 | =1 | 0.736016629 |
| 3 | 7/10 | 7/7 | =1 | 0.802012472 |
| 4 | 7/10 | 7/7 | =1 | 0.841609977 |
| 5 | 7/10 | 7/7 | =1 | 0.868008314 |

MAP: 0.868008314

Query 2:

Using Query: CEREMONIAL SUICIDES COMMITTED BY SOME BUDDHIST MONKS IN SOUTH VIET NAM AND WHAT THEY ARE SEEKING TO GAIN BY SUCH ACTS .

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Iteration | Precision | Recall | Average Precision | MAP |
| 0 | 5/10 | 5/9 | 1+2/2+3/3+4/6+5/8+6/10 = 0.543518519 | 0.543518519 |
| 1 | 8/10 | 8/9 | =0.88888888888 | 0.716203704 |
| 2 | 8/10 | 8/9 | =0.88888888888 | 0.773765432 |
| 3 | 8/10 | 8/9 | =0.88888888888 | 0.802546296 |
| 4 | 8/10 | 8/9 | =0.88888888888 | 0.819814815 |
| 5 | 8/10 | 8/9 | =0.88888888888 | 0.831327161 |

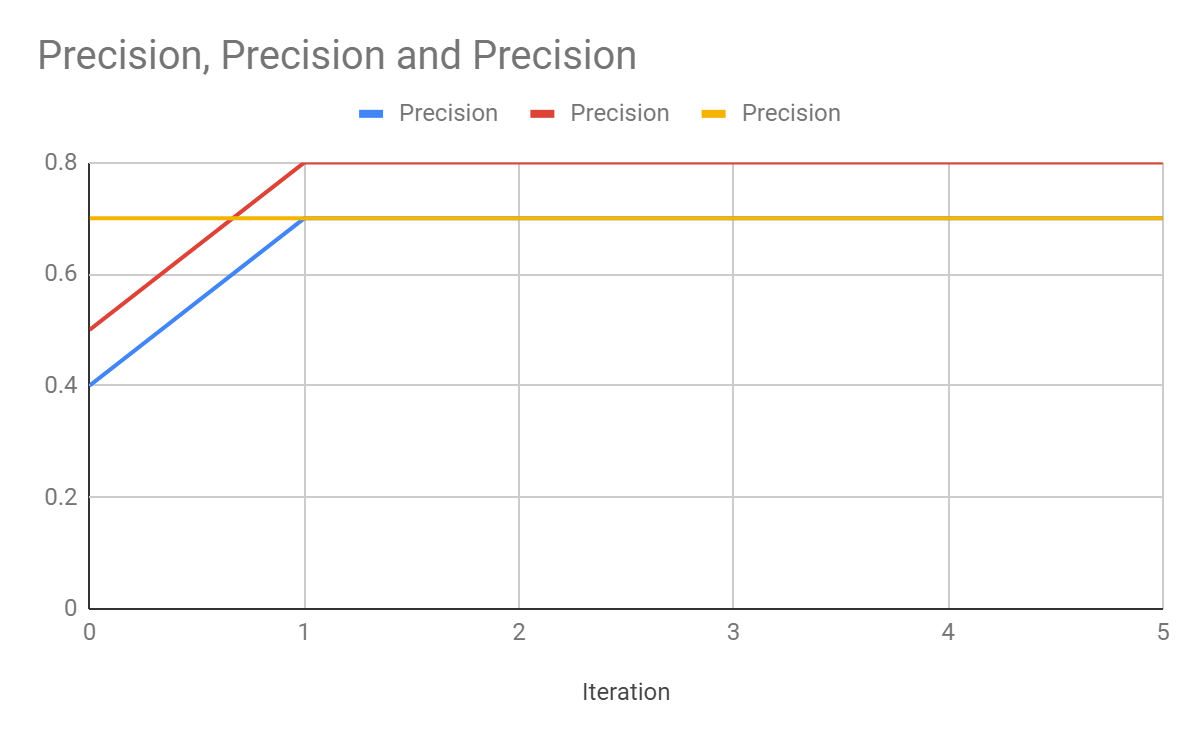
MAP: 0.831327161

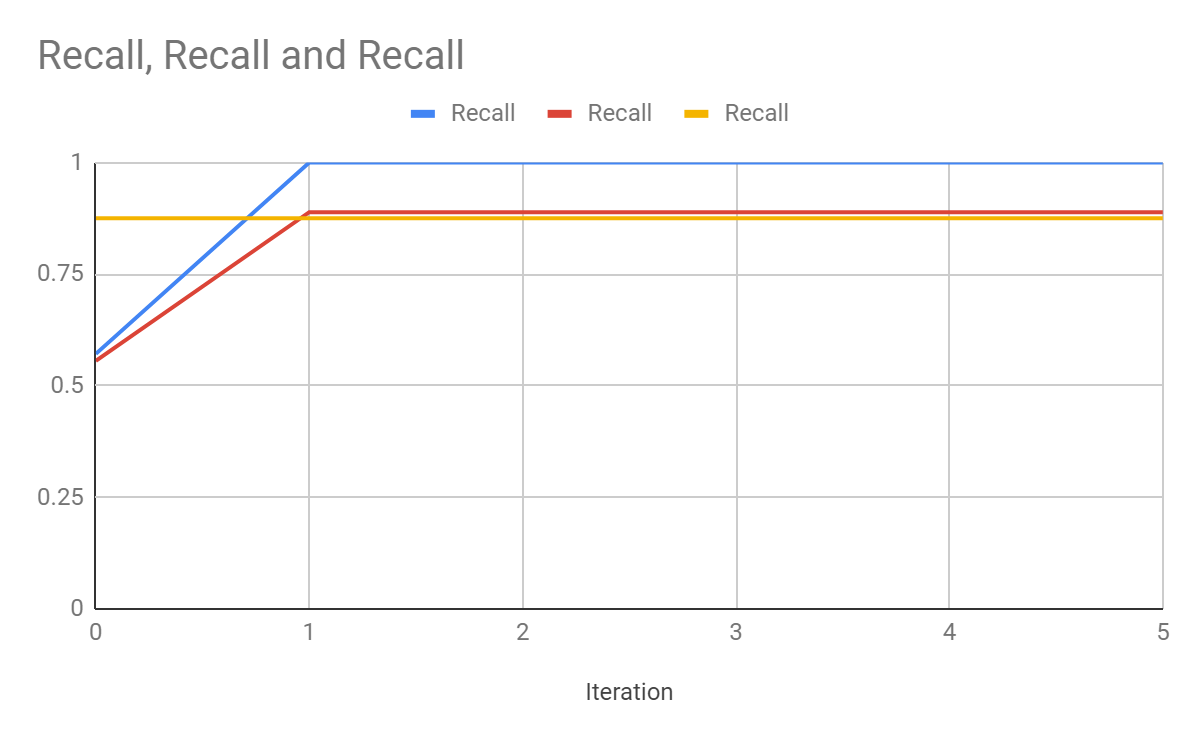
Query 3:

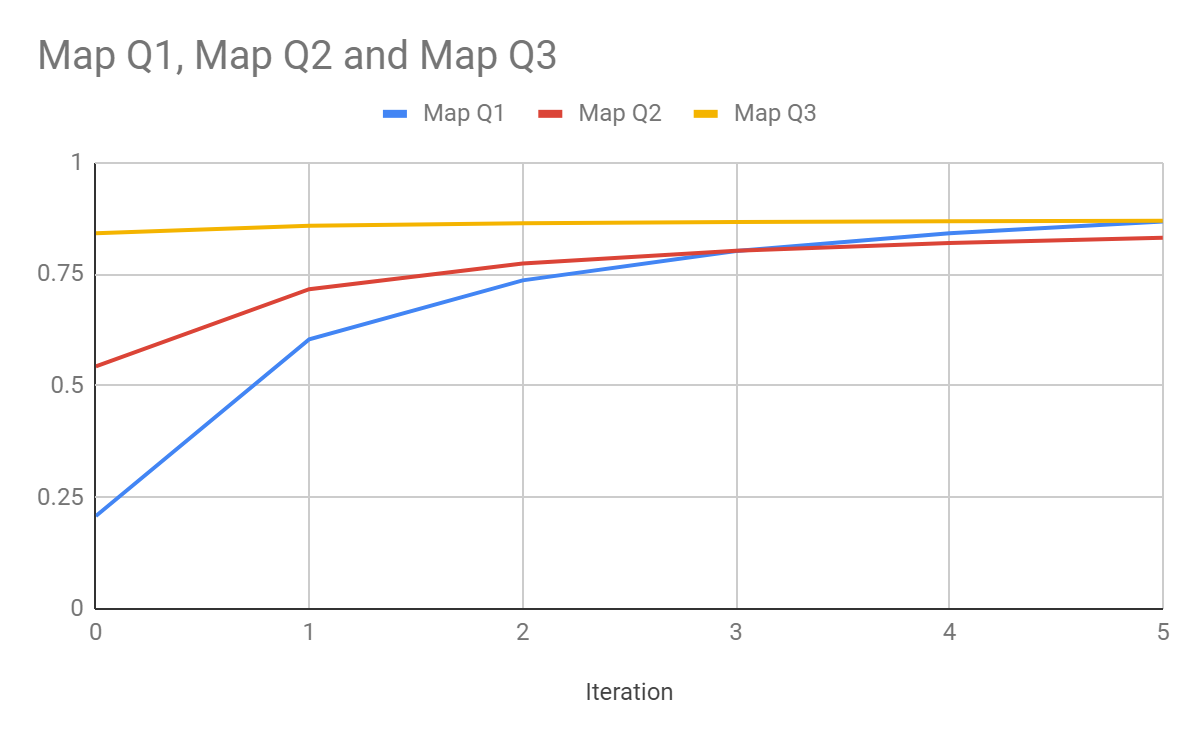
Using Query: OPPOSITION OF INDONESIA TO THE NEWLY-CREATED MALAYSIA .

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Iteration | Precision | Recall | Average Precision | MAP |
| 0 | 7/10 | 7/8 | 1+1+1+1+1+6/7+7/8 =0.841517857 | 0.841517857 |
| 1 | 7/10 | 7/8 | =0.875 | 0.858258928 |
| 2 | 7/10 | 7/8 | =0.875 | 0.863839286 |
| 3 | 7/10 | 7/8 | =0.875 | 0.866629464 |
| 4 | 7/10 | 7/8 | =0.875 | 0.868303571 |
| 5 | 7/10 | 7/8 | =0.875 | 0.869419643 |

MAP: 0.869419643





Pseudo relevance tends to have lower precision and recall

Pseudo Relevance feedback tends to have less recall and precision, due to the fact that it only chooses the top 3 documents as relevant even if they are not. A more detailed look into this can be seen in results.txt