

User Manual	2
1. Information on used packages	2
2. Data folder setup	2
3. How to execute Python files	2
4. Expected outputs for each Python file	4
5. Additional information	6
Design	7
• Description of your model for Task 1	7
• Assumption	7
• Algorithms (including inputs, outputs, and any other parameters)	8
• Description of your model for Task 2	8
• Assumption (if any)	9
• Algorithms (including inputs, outputs, and any other parameters)	9
• Description of your model for Task 3	10
• Assumption (if any)	10
• Algorithms (including inputs, outputs, and any other parameters)	11
Implementation	12
Model1 (BM25):	12
Model2 (JM_LM):	14
Model3 (My_PRM):	15
Results & Evaluation	15
Discussion	22
References	25
Appendix 1	25
Appendix 2	41
Appendix 3	58

Table of Contents

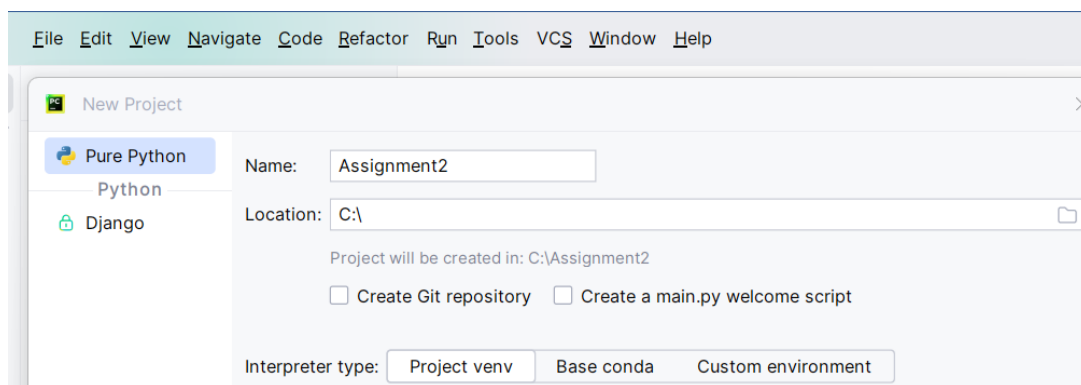
User Manual

1. Information on used packages

- **os** – used for reading from and writing to files.
- **math** – used for calculating the logarithm of values.
- **re** – used for text processing tasks.
- **string** – used for tasks like removing punctuation from text data.
- **numpy** – used for creating NumPy arrays.
- **pandas** – used for converting dictionaries to DataFrames.
- **nlTK** – used for stemming.
- **collections** – used for counting the frequency of each word in a list.
- **scipy.stats** – used for performing paired t-test

2. Data folder setup

2.1 Create a new PyCharm project.



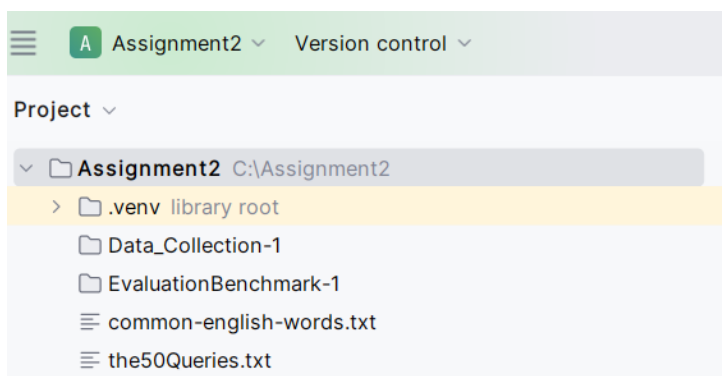
2.2 Unzip the “Data_collection.zip” file into the PyCharm project directory.

2.3 Unzip the “EvaluationBenchmark-1.zip” file into the PyCharm project directory.

2.4 Store the “the50Queries.txt” file in the PyCharm project directory.

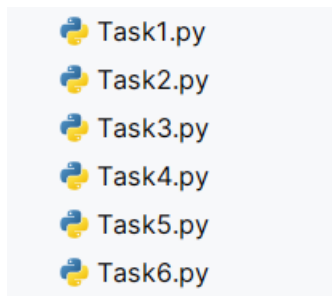
2.5 Store the “common-english-words.txt” file in the PyCharm project directory.

2.6 Ensure all files are in the same directory, as shown below:



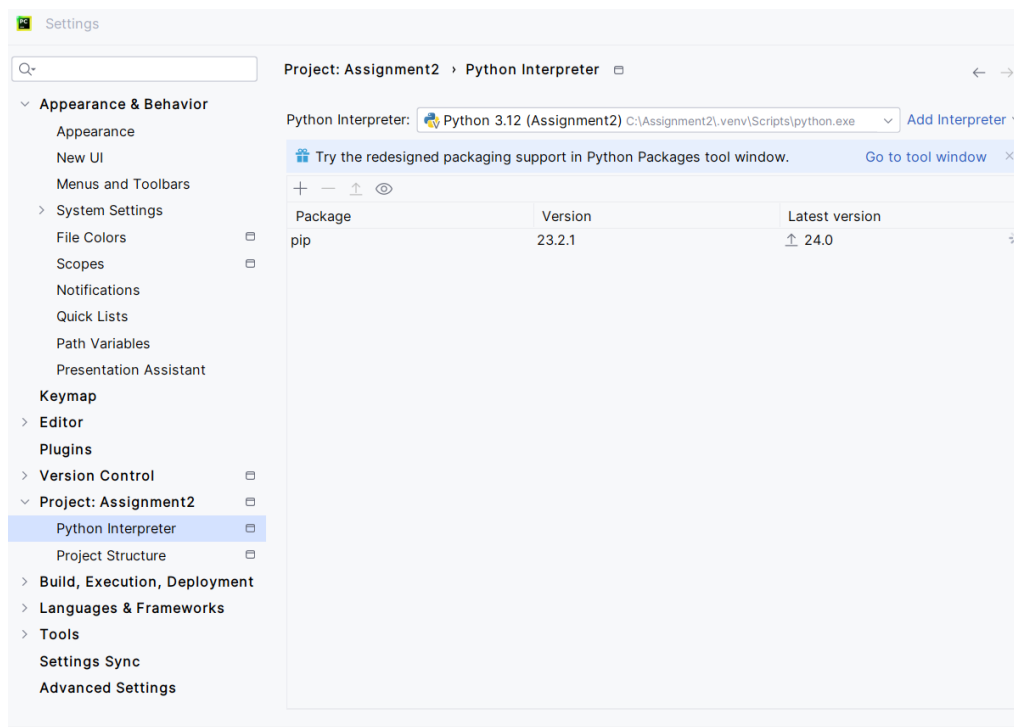
3. How to execute Python files

3.1 Load the Python files provided by us into the PyCharm project. The files are named from “task1.py” to “task6.py.”

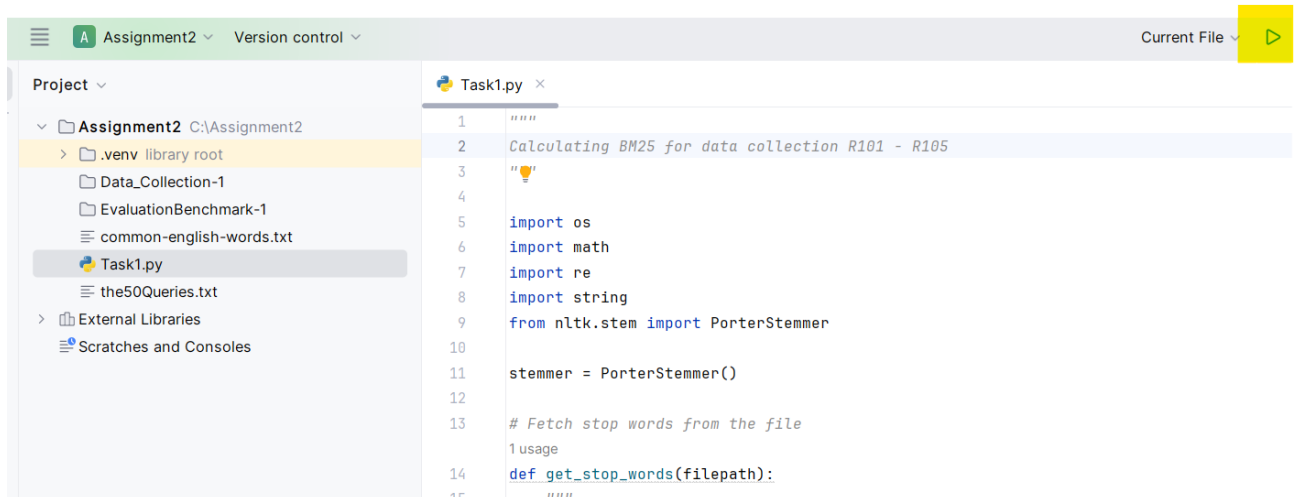


3.2 Install Required Packages –

- Go to File > Settings > Project: <Your Project> > Python Interpreter
- Press the + button to add new packages.
- Install the necessary packages (e.g., NumPy, pandas).



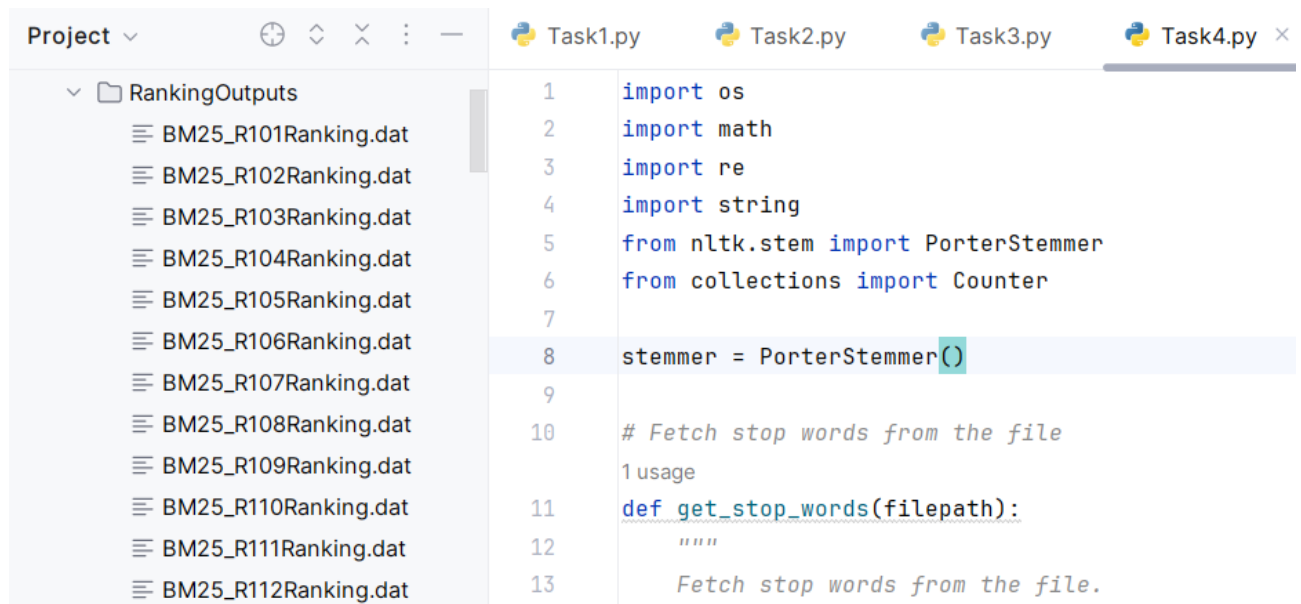
3.3 Load the script for Task 1 into the “Assignment2” file, Press the green play button (triangle) at the top right corner to run the script.



3.4 Repeat step 3.3 for the scripts for Task 2 to 6 and ensure each task script is loaded and executed separately.

4. Expected outputs for each Python file

4.1 [Task 4] After running task4.py, output files for all three models will be stored in the folder "RankingOutputs" in dat format.



Example output: BM25_R107Ranking.dat

Project	Task4.py	BM25_R107Ranking.dat	Task5
RankingOutputs	1	R107 Q0 51576 1 2.765639300454872 BM25	
BM25_R101Ranking.dat	2	R107 Q0 71157 2 2.386400256599359 BM25	
BM25_R102Ranking.dat	3	R107 Q0 77936 3 2.273108908141271 BM25	
BM25_R103Ranking.dat	4	R107 Q0 79950 4 2.144090202267175 BM25	
BM25_R104Ranking.dat	5	R107 Q0 59244 5 1.8983172340327779 BM25	
BM25_R105Ranking.dat	6	R107 Q0 67107 6 1.6074284655326316 BM25	
BM25_R106Ranking.dat	7	R107 Q0 67411 7 1.6074284655326316 BM25	
BM25_R107Ranking.dat	8	R107 Q0 86459 8 1.128090008008347 BM25	
BM25_R108Ranking.dat	9	R107 Q0 31404 9 0.5807469321665877 BM25	
BM25_R109Ranking.dat	10	R107 Q0 41791 10 0.5635155390134101 BM25	
BM25_R110Ranking.dat	11	R107 Q0 69377 11 0.5625603334593127 BM25	
BM25_R111Ranking.dat	12	R107 Q0 37330 12 0.5584756327943027 BM25	
BM25_R112Ranking.dat	13	R107 Q0 18297 13 0.5497333978620388 BM25	
BM25_R113Ranking.dat	14	R107 Q0 18529 14 0.5483670739951522 BM25	
BM25_R114Ranking.dat	15	R107 Q0 78164 15 0.5467090324257458 BM25	
	16	R107 Q0 67673 16 0.5418838837605672 BM25	

1.2 [Task 5] After running task5.py, the evaluation results, including average precision, precision@10, and DCG10, will be stored in “evaluation_results_[measure_name].csv” and can be exported to Excel for further analysis.

Project	Task5.py	Task6.py
pythonProject C:\Users\Fiona\IFNE	1	import os
> .venv library root	2	import math
> Data_Collection	3	import re
> EvaluationBenchmark	4	import string
> RankingOutputs	5	import numpy as np
common-english-words.txt	6	import pandas as pd
evaluation_results_ap.csv	7	from nltk.stem import PorterStemmer
evaluation_results_dcg10.csv	8	from collections import Counter
evaluation_results_p10.csv	9	
	10	stemmer = PorterStemmer()

1.3 [Task 6] After running `task6.py`, the paired t-test results, including the T-statistics, p-value, and comments, will be output in the PyCharm console.

[Task 6] Please *run task 5 before task 6* as the input of task 6 are “evaluation_results_[measure_name].csv” and output is the t-test result.

The screenshot displays a Jupyter Notebook environment. The top pane shows a file explorer on the left with a project named 'pythonProject' containing folders like '.venv', 'Data_Collection', 'EvaluationBenchmark', and 'RankingOutputs', and a file 'common-english-words.txt'. The main editor shows 'Task6.py' with the following code:

```
1 import pandas as pd
2 from scipy import stats
3
4 def generate_t_test(output1, output2):
```

The bottom pane, labeled 'Run', shows the execution output for 'Task6'. It contains three comparison results for Average Precision (AP) and two recommendations:

```
Comparing BM25 and My_PRM models for Average Precision (AP)
T-Statistic: -2.053547480047978
P-Value: 0.04537647250920848

Comparing JM_LM and My_PRM models for Average Precision (AP)
T-Statistic: -2.010204667297627
P-Value: 0.04993104651777437

Comparing BM25 and JM_LM models for Average Precision (AP)
T-Statistic: -0.14050994577913326
P-Value: 0.8888331211414475

Recommendation for Average Precision (AP): My_PRM model is significantly better than BM25.

Recommendation for Average Precision (AP): My_PRM model is significantly better than JM_LM.
```

5. Additional information

Due to spacing or formatting issues with the query numbers in the `the50Queries.txt` file, some of the queries could not be loaded from the original file. We have removed the spaces from the query numbers and uploaded the corrected file along with our Python scripts.

Design

Task 1: Design a BM25-based IR model (**BM25**) that ranks documents in each data collection using the corresponding topic (query) for all 50 data collections.

- Description of your model for Task 1

IR model purpose: To provide a mathematical framework for defining the search process.

functionality: Finds relevant documents for a given query

key Research Issue: Theories about relevance.

- Our BM25 Model Explanation:

Purpose: extends the scoring function for the binary independence model to include document and query term weights.

The BM25 formula is as follows:

$$\sum_{i \in Q} \log \frac{(r_i + 0.5) / (R - r_i + 0.5)}{(n_i - r_i + 0.5) / (N - n_i - R + r_i + 0.5)} \cdot \frac{(k_1 + 1) f_i}{K + f_i} \cdot \frac{(k_2 + 1) q f_i}{k_2 + q f_i}$$

Variable	Description	Notes
ri	Relevance Information	Remains constant for a term across all queries but varies per term. Assumed as 0.
R	Total Relevant Documents	Constant for a collection during a single series of queries. Assumed as 0.
ni	Document Frequency	Changes per term, indicating how many documents contain the term.
N	Total Number of Documents	Fixed for a collection, to avoid log negative value, * 2 in the equation
fi	Frequency of Term in Document	Frequency weight (count) of term i in the document.
qfi	Frequency of Term in Query	Frequency weight (count) of term i in the query.
k1	Tuning Parameter	Constant, given 1.2
k2	Tuning Parameter	Constant, given 500.
b	Tuning Parameter	Constant, 0.75
dl	Document Length	Changes per XML document.
avdl	Average Document Length	Constant for a collection, calculated as the average length across all documents.
K	$\frac{(1 - b) + b * dl}{avdl} * K1$	

Assumption

- Document Collection: Pre-processed to remove stop words, perform stemming, and handle punctuation and numeric characters.
- Query Processing: Queries are parsed, cleaned, and stemmed to match the document collection processing.
- Relevance Information: We assume R and r as 0
- Total number of documents: adjusted to avoid log negative values by multiplying by 2.
- Standard Parameters: Uses typical values for BM25 parameters (k1 = 1.2, k2 = 500, b = 0.75).

Algorithms (including inputs, outputs, and any other parameters)

Inputs.

- Document Collection (coll): Collection of documents with unique IDs and lists of terms.
- Query (query_terms): Parsed and stemmed query as a dictionary of terms and their frequencies.
- Document Frequencies (df): Dictionary containing the document frequency of each term.
- Parameters. Typical values for BM25 parameters are used ($k_1 = 1.2$, $k_2 = 500$, $b = 0.75$).

Output: Document Scores (scores): Dictionary containing BM25 scores for each document.

Task 2: Design a Jelinek-Mercer based Language Model (JM_LM) that ranks documents in each data collection using the corresponding topic (query) for all 50 data collections.

- Description of your model for Task 2

The Jelinek-Mercer smoothing technique is employed to develop a language model that ranks documents according to their relevance to specific queries. This method effectively interpolates between the probability of observing terms within individual documents and their occurrence across the entire document collection. By doing so, it ensures that even documents that do not contain certain query terms are still considered by assigning them a non-zero probability of relevance.

Jelinek-Mercer Smoothing Formula model follows these steps:

1. Parsing the Queries-

- We parse the queries to include all contents (title, description, and narrative), not only the topic words. This provides a richer and more detailed query representation for better retrieval performance.

2. Document Collection Processing -

- For each document collection specified, read the document files. Tokenize the text in each document, remove stop words, and apply stemming to the terms. Create DataDoc objects for each document with the processed terms and update their term frequencies and document lengths.

3. Compute Term Frequencies -

- Calculate the frequency of each term within individual documents and update the DataDoc objects. Aggregate these frequencies to compute the total frequencies of each term across the entire document collection. Update the collection term frequencies and the total collection length accordingly.

4. Calculate JM Smoothing Scores -

- Use the Jelinek-Mercer smoothing formula to compute the conditional probability of each document given the query. Calculate the product of these probabilities for all terms in the query to derive each document's overall score.

JM Smoothing equation -

$$P(R | D) = \prod_{i=1}^n \left[(1 - \lambda) \frac{f(q_i, D)}{|D|} + \lambda \frac{c(q_i)}{|C|} \right]$$

Variable	Description	Notes
$p(R D)$	Probability of retrieving D given R.	Overall score for document D based on query R.
n	Number of query terms.	Total terms in the query.
q_i	i -th query term.	Each term in the query.
λ	Smoothing parameter.	Balances local and global term frequencies. Typical values around 0.4.
$f(q_i, D)$	Frequency of q_i in D.	Counts appearances of q_i in D.
D	Document being evaluated	Represents a single document
$c(q_i)$	Frequency of q_i in collection C.	Shows how common q_i is across all documents.
Data_Cx	Entire document collection	Encompasses all documents in the search

$$\text{Formula: } (1 - \lambda) \times \frac{f(t,D)}{|D|}$$

Document-Specific Probability:

- Assesses how often term t appears in document D , showing its relevance within that document. $|D|$, the total term count in D , converts the raw count into a probability.

$$\text{Formula: } \lambda \times \frac{f(t,C)}{|C|}$$

Collection-Wide Probability:

- Measures the frequency of term t across all documents, indicating its overall importance. $|C|$, the total term count in the collection, ensures the value is a probability.

Combining Probabilities:

- Calculate the product of these probabilities for all terms in the query to derive each document's overall score.

5. Rank Documents -

- Use the computed document scores for each query to rank the documents in descending order of relevance.

Assumption

- The model uses a constant lambda value of 0.4 for interpolation to balance between document-specific and collection-wide term probabilities.
- Document and collection terms are normalized through stemming and removal of stop words to enhance relevance calculations.
- Lengths of documents and the overall collection are precomputed to speed up the normalization of term frequencies.

- Algorithms (including inputs, outputs, and any other parameters)

Step	Inputs	Outputs	Parameters
Load Stop Words	File path to stop words list (common-english-words.txt)	List of stop words	None
Preprocess Queries	Query file (the50Queries.txt)	Parsed queries with stemmed terms, excluding stop words	None
Parse Document Collection	Document collection directory (Data_Collection)	DataDoc objects with term frequencies and document lengths	None
Compute Term Frequencies	Parsed document collection (DataColl object)	Aggregated term frequencies for the entire collection	None
Calculate JM Smoothing Scores	Parsed queries, document term frequencies, document lengths, collection term frequencies, total collection length	Document scores based on JM smoothing formula	lambda = 0.4
Rank Documents	Document scores for each query	Ranked list of documents for each query	None

Task 3. Based on the knowledge you gained from this unit, design a pseudo-relevance model (*My_PRM*) to rank documents in each data collection using the corresponding topic (query) for all 50 data collections.

- Description of your model for Task 3

The pseudo-relevance model (My_PRM) designed for this task is based on the concept of relevance feedback and pseudo-relevance feedback. The model aims to improve the ranking of documents by iteratively refining the query terms based on the relevance of top-ranked documents.

The My_PRM model follows these steps:

1. BM25 as the base language model -

- The BM25 algorithm is used to rank documents initially. BM25 calculates a score for each document based on the query terms and document term frequencies.

BM25 equation -

$$\sum_{i \in Q} \log \frac{(r_i + 0.5) / (R - r_i + 0.5)}{(n_i - r_i + 0.5) / (N - n_i - R + r_i + 0.5)} \cdot \frac{(k_1 + 1) f_i}{K + f_i} \cdot \frac{(k_2 + 1) q f_i}{k_2 + q f_i}$$

2. Parsing the Long Queries -

- We parse the queries to include all contents (title, description, and narrative), not only the topic words. This provides a richer and more detailed query representation for better retrieval performance.

3. Pseudo-Relevance Feedback (PRF) -

- The top-ranked documents from the initial ranking are considered as pseudo-relevant. The terms from these top documents are used to expand the query, adding new terms that are likely to be relevant.

4. Probabilistic Relevance Feedback -

- Further refinement of the query terms is performed by considering both relevant and non-relevant documents. The probabilities of terms appearing in relevant and non-relevant documents are used to adjust the weights of the query terms. [1]

Probabilities relevance feedback equation –

$$sim(d_j, q) = \frac{P(R|\vec{d}_j)}{P(\bar{R}|\vec{d}_j)}$$

5. Re-ranking Using Enhanced Query -

- The expanded and refined query is used to re-rank the documents. The BM25 algorithm is applied again using the enhanced query terms to get a more accurate ranking of documents.

6. Building Relevance Models -

- Relevance models are built from the top-ranked documents to calculate term probabilities and improve the ranking further. These models help to identify the most informative terms for the final re-ranking.

The probability of a term w in a relevance model is calculated as:

$$P(w|D) = (1 - \lambda) \frac{term_freq[w]}{doc_len} + \lambda \frac{collection_prob[w]}{total_doc_len}$$

- Assumption

1. The top-ranked documents in the initial ranking are considered pseudo-relevant for feedback.
2. The relevance feedback process assumes that terms appearing frequently in top-ranked documents are likely to be relevant.

3. A small epsilon value is used to avoid mathematical domain errors in probability calculations.

- Algorithms (including inputs, outputs, and any other parameters)

Step	Inputs	Outputs	Parameters
Initial BM25 Ranking	Document collection, query terms, document frequencies	Initial ranked documents based on BM25 scores	$k_1 = 1.2$, $k_2 = 500$, $b = 0.75$
Parsing the Queries	Query file	Parsed queries including title, description, narrative	None
Pseudo-Relevance Feedback	Initial query terms, top-ranked documents, document collection, doc frequencies	Expanded query terms with additional relevant terms	$\text{num_feedback_docs} = 50$, $\text{num_expansion_terms} = 200$
Probabilistic Relevance Feedback	Initial query terms, relevant documents, non-relevant documents, document frequencies, total number of documents	Updated query terms with adjusted weights	$\alpha = 10$, $\beta = 5$, $l_1_lambda = 0.99$
Building Relevance Models	Top-ranked documents, document collection, document frequencies, total document length, number of top terms	Relevance model with term probabilities	$\text{top_terms_count} = 245$, $\text{percentage} = 0.9$
Final Ranking Using Enhanced Query	Document collection, enhanced query terms, document frequencies	Final ranked documents based on the enhanced query	$\text{lambda_param} = 0.9$, k_1 (num of top rank) = 61, k_2 (num of rerank doc) = 4

Final output: Document Scores (scores) saved in My_PRM_R[Query_ID]Ranking for each document.

Implementation

Model1 (BM25):

- Python package or module (or any open-source software) you used:
 1. os: Used for interacting with the operating system, including reading directory contents and file paths.
 2. math: Provides mathematical functions, particularly used here for logarithmic calculations.
 3. re: Provides regular expression matching operations, used for parsing text and cleaning data.
 4. string: Contains common string operations, used for cleaning and manipulating strings.
 5. nltk: Natural Language Toolkit, used for text processing, including tokenization and stemming. Specifically, the PorterStemmer is used for stemming terms.
 6. nltk.download('punkt'): Downloads the necessary data for tokenization.

Data structures (used to represent a single document and a set of documents):

1. DataDoc: Represents a single document.
 - a. Attributes:
 - i. docID: Unique identifier for the document.
 - ii. terms: Dictionary where keys are terms and values are their frequencies within the document.
 - iii. doc_len: Length of the document in terms of the number of terms.
 - b. Methods:
 - i. set_doc_len(count): Sets the length of the document.
 - ii. getDocId(): Returns the document ID.
 - iii. getDocLen(): Returns the document length.
 - iv. get_term_list(): Returns the terms dictionary.
 - v. add_term(term): Adds a term to the terms dictionary.
2. DataColl: Represents a collection of documents.
 - a. Attributes:
 - i. collections: Dictionary where keys are document IDs and values are DataDoc objects.
 - ii. numOfDocs: Total number of documents in the collection.
 - iii. totalDocLength: Sum of the lengths of all documents in the collection.
 - iv. avgDocLen: Average length of documents in the collection.
 - b. Methods:
 - i. add_doc(doc): Adds a document to the collection.
 - ii. get_coll(): Returns the collections dictionary.
 - iii. getTotalDocLen(): Returns the total document length.
 - iv. getNumOfDocs(): Returns the number of documents.
 - v. setAvgDocLen(avg): Sets the average document length.
 - vi. getAvgDocLen(): Returns the average document length.

Steps:

1. Import necessary libraries and modules
 - os, math, re, string
 - stem from porter2
2. Function to fetch stop words from a file
 - Open file and read contents
 - Split content by comma to create stop words list
 - Return stop words list
3. Initialize variables and constants

- stop_words from "common-english-words.txt"
- queryfile = "the50Queries.txt"
- outputFolder = "RankingOutputs"
- data_collection_folder = "Data_Collection"
- evaluation_benchmark_folder = "EvaluationBenchmark"

4. Class DataDoc

- Initialize docID, terms dictionary, and doc_len
- Methods to set and get doc_len, docID, and term list
- Method to add term to terms dictionary

5. Class DataColl

- Initialize collections dictionary, numOfDocs, totalDocLength, and avgDocLen
- Methods to add doc, get collections, get total doc length, get number of docs, set and get average doc length

6. Function to parse a query

- Remove digits and punctuation from query
- Split query into terms
- Stem terms and filter based on length and stop words
- Return parsed terms dictionary

7. Function to load and parse queries from a file

- Open and read query file
- Use regex to extract query data
- Parse each query and store in a dictionary
- Return dictionary of parsed queries

8. Function to parse documents in a collection

- Initialize DataColl instance
- List files in input path
- For each XML file, read lines and process document
- Extract and stem terms, add to DataDoc instance
- Add DataDoc to DataColl
- Return DataColl instance

9. Function to calculate document frequency

- Initialize docFreq dictionary
- For each document, get term list and update docFreq
- Return docFreq dictionary

10. Function to calculate average document length

- Calculate average length and set in DataColl instance
- Return average length

11. Function to calculate BM25 scores

- Initialize scores dictionary and BM25 parameters
- For each document, calculate score for each query term
- Return scores dictionary

12. Function to save ranked results to a file

- Open output file for writing
- Write ranked documents to file in specified format

- Print confirmation message

13. Main block

- Ensure output directory exists
- Load and print queries
- List and process data collection folders
- For each folder, parse collection and calculate BM25 scores
- Save ranked results to output file
- Print completion message

Model2 (JM_LM):

Python package or module (or any open-source software) you used:

- nltk (Natural Language Toolkit): for text processing, including tokenization and stemming
- nltk.stem.PorterStemmer: For stemming words to their root form.
- os: For file and directory operations.
- numpy: For numerical operations.
- string: For string manipulation.
- re: For regular expression operations to parse text.
- collections.Counter: For counting and managing term frequencies.
- Data Structures

1. DataDoc Class:

- Represents a single document.
- Attributes:
 - docID: Document ID.
 - terms: Dictionary of terms and their frequencies.
 - doc_len: Length of the document in terms of the number of words.
- Methods:
 - set_doc_len(count): Sets the document length.
 - getDocId(): Returns the document ID.
 - getDocLen(): Returns the document length.
 - get_term_list(): Returns the dictionary of terms and their frequencies.
 - add_term(term): Adds a term to the document.

2. DataColl Class:

- Represents a collection of documents.
- Attributes:
 - collections: Dictionary of documents.
 - numOfDocs: Number of documents in the collection.
 - totalDocLength: Total length of all documents in the collection.
 - avgDocLen: Average length of the documents in the collection.
- Methods:
 - add_doc(doc): Adds a document to the collection.
 - get_coll(): Returns the dictionary of documents.
 - getTotalDocLen(): Returns the total document length.
 - getNumOfDocs(): Returns the number of documents.
 - setAvgDocLen(avg): Sets the average document length.
 - getAvgDocLen(): Returns the average document length.

Functions

1. Fetching Stop Words:

- Function: get_stop_words(filepath)
- Reads stop words from a file and returns a list of stop words.

2. **Parsing Document Collection:**
 - Function: `parse_collection(inputpath)`
 - Parses documents from a specified folder and adds them to a `DataColl` object.
3. **Parsing Queries:**
 - Function: `parse_query_long(query)`
 - Parses a long query and returns a dictionary of query terms and their frequencies.
4. **Loading Queries:**
 - Function: `load_queries_long(query_file)`
 - Loads long queries from a file and returns a dictionary of queries.
5. **Calculating Document Frequencies:**
 - Function: `my_df(coll)`
 - Calculates the frequency of each term in the collection.
6. **Calculating Average Document Length:**
 - Function: `avg_length(coll)`
 - Calculates and sets the average document length in the collection.

7. **JM Smoothing Score Calculation:**
 - Function: `jm_smoothing(query, doc_term_freq, doc_length, collection_term_freq, collection_length, lambda_param)`
 - Initializes a dictionary for scores.
 - Sets lambda parameter to 0.4.
 - For each document in the collection, calculates the Jelinek-Mercer score for each query term.
 - Aggregates scores for terms to get a total document score.
 - Uses the Jelinek-Mercer smoothing formula to compute the conditional probability of each document given the query.
8. **Performing JM LM Ranking:**
 - Function: `rank_documents(query_id, query, documents, collection_term_freq, collection_length)`
 - Uses the computed document scores for each query to rank the documents in descending order of relevance.
 - Initializes a list for scores.
 - For each document, calculates the score using `jm_smoothing`.
 - Sorts the documents based on scores
9. **Saving Ranked Results:**
 - Function: `save_ranked_results(system_name, query_id, ranked_docs)`
 - Saves the ranked list of documents for each query to an output file.

Model3 (My_PRM):

Data Structures

1. **DataDoc Class:**
 - Represents a single document.
 - Attributes: `docID`, `terms` (dictionary of terms and frequencies), `doc_len`.
 - Methods: `set_doc_len(count)`, `getDocId()`, `getDocLen()`, `get_term_list()`, `add_term(term)`.
2. **DataColl Class:**
 - Represents a collection of documents.
 - Attributes: `collections`, `numOfDocs`, `totalDocLength`, `avgDocLen`.
 - Methods: `add_doc(doc)`, `get_coll()`, `getTotalDocLen()`, `getNumOfDocs()`, `setAvgDocLen(avg)`, `getAvgDocLen()`.

Functions

1. **Fetching Stop Words:**
 - Function: `get_stop_words(filepath)`

- Reads stop words from a file.
- 2. **Parsing Document Collection:**
 - Function: `parse_collection(inputpath)`
 - Parses documents and adds them to a `DataColl` object.
- 3. **Parsing Queries:**
 - Function: `parse_query_long(query)`
 - Parses a long query.
- 4. **Loading Queries:**
 - Function: `load_queries_long(query_file)`
 - Loads long queries from a file.
- 5. **Calculating Document Frequencies:**
 - Function: `my_df(coll)`
 - Calculates term frequencies in the collection.
- 6. **Calculating Average Document Length:**
 - Function: `avg_length(coll)`
 - Calculates the average document length.
- 7. **BM25 Weight Calculation:**
 - Function: `calculate_bm25_weights(coll, query_terms, df)`
 - Calculates BM25 weights for terms in documents.
- 8. **Probabilistic Relevance Feedback:**
 - Function: `probabilistic_relevance_feedback(query_terms, relevant_docs, non_relevant_docs, df, N, alpha, beta, l1_lambda)`
 - Updates query terms using probabilistic relevance feedback.
- 9. **Pseudo-Relevance Feedback:**
 - Function: `pseudo_relevance_feedback(query_terms, top_docs, coll, df, num_feedback_docs, num_expansion_terms)`
 - Expands query terms using pseudo-relevant documents.
- 10. **Building Relevance Models:**
 - Function: `build_relevance_model(top_k_docs, collection, collection_term_freq, total_doc_length, top_terms_count, percentage)`
 - Builds a relevance model from top documents.
- 11. **Performing PRM Ranking:**
 - Function: `my_prm(coll, query_id, query_terms, df, top_terms_count)`
 - Performs PRM ranking.
- 12. **Saving Ranked Results:**
 - Function: `save_ranked_results(system_name, query_id, ranked_docs)`
 - Saves the ranked results to a file.

Results & Evaluation

Task 5. Use three effectiveness measures to evaluate the three models.

Table 1: The performance of 3 models on average precision (MAP)

Topic	BM25	JM_LM	My_PRM
R101	0.665	0.665	0.536
R102	0.785	0.782	0.724
R103	0.421	0.310	0.437
R104	0.914	0.856	0.951
R105	0.765	0.647	0.914
R106	0.442	0.507	0.617
R107	0.303	0.326	0.462
R108	0.139	0.151	0.583
R109	0.641	0.668	0.753
R110	0.553	0.562	0.900
R111	0.131	0.131	0.216
R112	0.577	0.807	0.623
R113	0.635	0.657	0.776
R114	0.858	0.927	0.778
R115	0.325	0.288	0.152
R116	0.300	0.291	0.413
R117	0.269	0.483	0.556
R118	0.322	0.290	0.403
R119	0.196	0.152	0.601
R120	0.523	0.552	0.271
R121	0.606	0.742	0.750
R122	0.553	0.558	0.912
R123	0.444	0.493	0.189
R124	0.391	0.590	0.449
R125	0.686	0.531	0.703
R126	0.594	0.615	0.637
R127	0.575	0.463	0.316
R128	0.328	0.329	0.364
R129	0.530	0.523	0.725
R130	0.698	0.643	0.806
R131	0.425	0.252	0.208
R132	0.292	0.295	0.331
R133	1.000	1.000	0.452
R134	0.184	0.195	0.257
R135	0.430	0.545	0.416
R136	0.437	0.482	0.681
R137	1.000	1.000	0.767
R138	0.210	0.204	0.470
R139	0.867	0.738	0.867
R140	0.795	0.693	0.744
R141	0.850	0.725	0.733

R142	0.177	0.177	0.198
R143	0.130	0.107	0.081
R144	0.381	0.487	0.933
R145	0.046	0.048	0.040
R146	0.398	0.359	0.869
R147	0.624	0.597	0.678
R148	0.349	0.349	0.732
R149	0.200	0.209	0.205
R150	0.419	0.466	0.296
Average	0.488	0.489	0.549

Graph 1: The comparison of 3 models on MAP

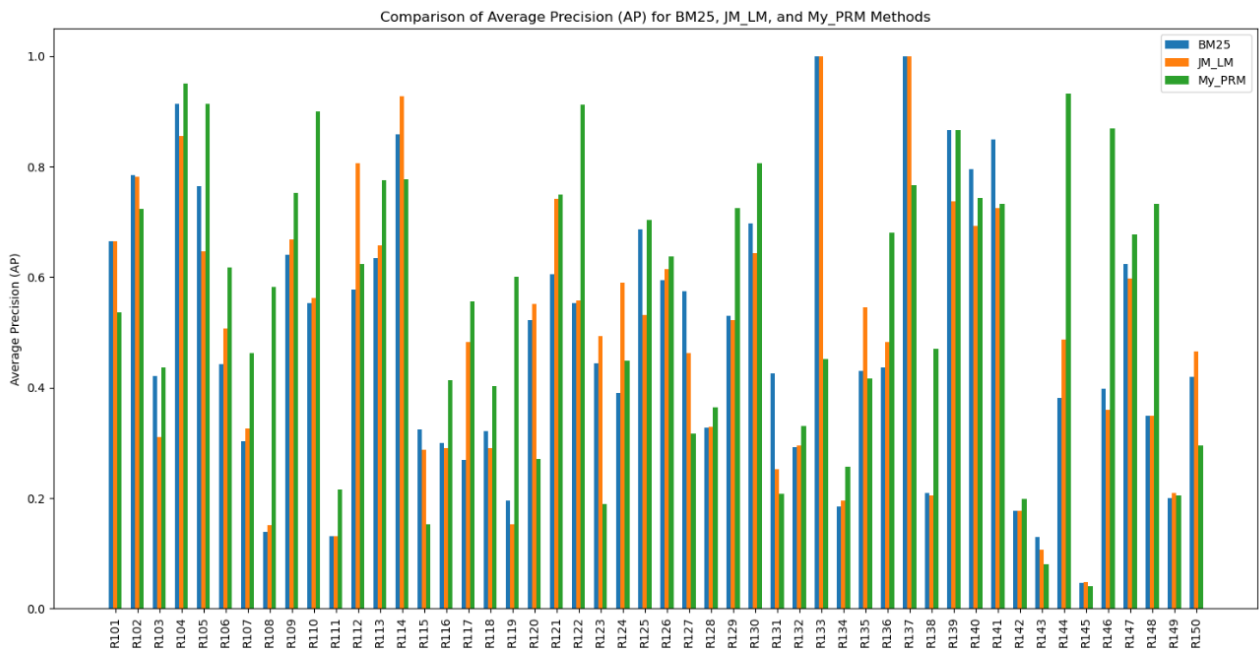


Table 2: The performance of 3 models on precision@10

Topic	BM25	JM_LM	My_PRM
R101	0.500	0.500	0.500
R102	0.600	0.600	0.600
R103	0.500	0.300	0.500
R104	1.000	0.800	1.000
R105	0.700	0.700	1.000
R106	0.200	0.300	0.300
R107	0.200	0.200	0.200
R108	0.100	0.200	0.200
R109	0.600	0.600	0.700
R110	0.400	0.400	0.500
R111	0.000	0.000	0.300
R112	0.600	0.600	0.600
R113	0.600	0.800	0.800

R114	0.500	0.500	0.500
R115	0.200	0.200	0.100
R116	0.200	0.100	0.400
R117	0.200	0.200	0.300
R118	0.300	0.200	0.300
R119	0.100	0.000	0.300
R120	0.600	0.500	0.300
R121	0.600	0.600	0.700
R122	0.400	0.500	0.900
R123	0.200	0.200	0.200
R124	0.400	0.400	0.400
R125	0.700	0.700	0.700
R126	0.700	0.600	0.600
R127	0.500	0.500	0.400
R128	0.100	0.100	0.100
R129	0.500	0.600	0.800
R130	0.300	0.300	0.300
R131	0.100	0.100	0.100
R132	0.200	0.200	0.300
R133	0.500	0.500	0.500
R134	0.200	0.100	0.300
R135	0.400	0.400	0.200
R136	0.400	0.400	0.500
R137	0.300	0.300	0.300
R138	0.100	0.200	0.400
R139	0.300	0.200	0.300
R140	0.900	0.900	0.700
R141	0.900	0.700	0.800
R142	0.100	0.100	0.200
R143	0.000	0.000	0.000
R144	0.400	0.500	0.600
R145	0.000	0.000	0.000
R146	0.100	0.100	0.800
R147	0.500	0.500	0.400
R148	0.100	0.100	0.700
R149	0.200	0.200	0.200
R150	0.400	0.300	0.300
Average	0.372	0.360	0.442

Graph 2: The comparison of 3 models on precision@10

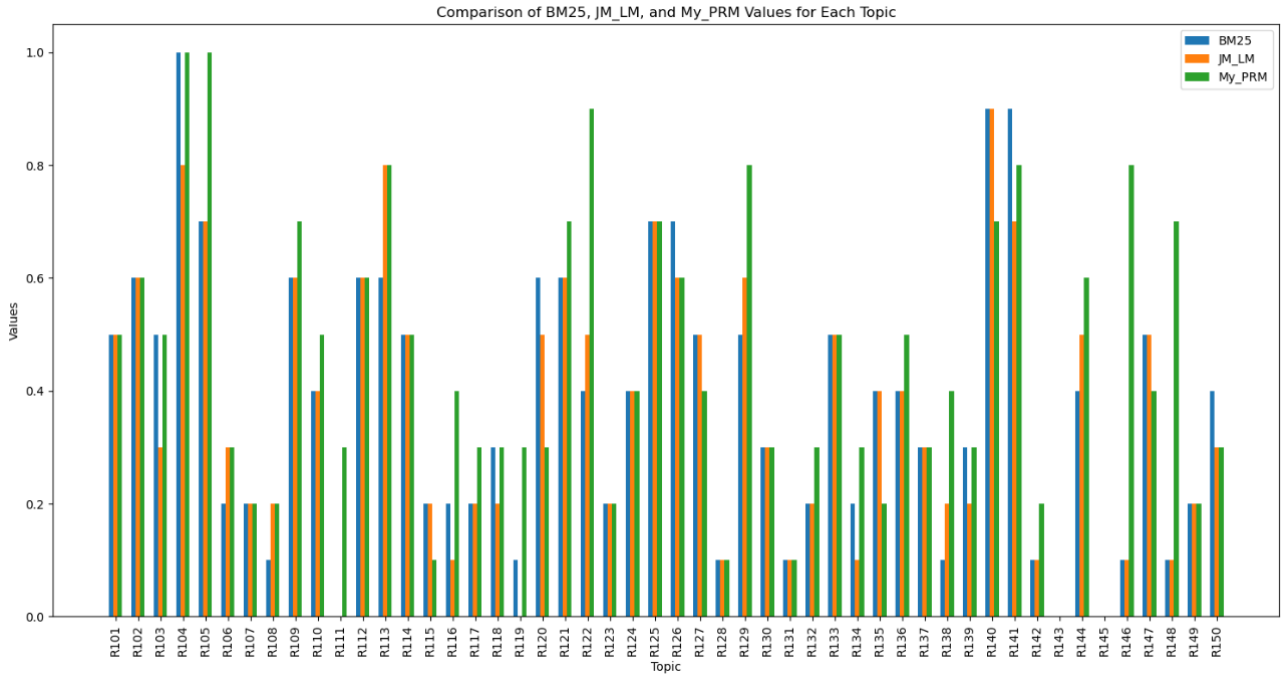
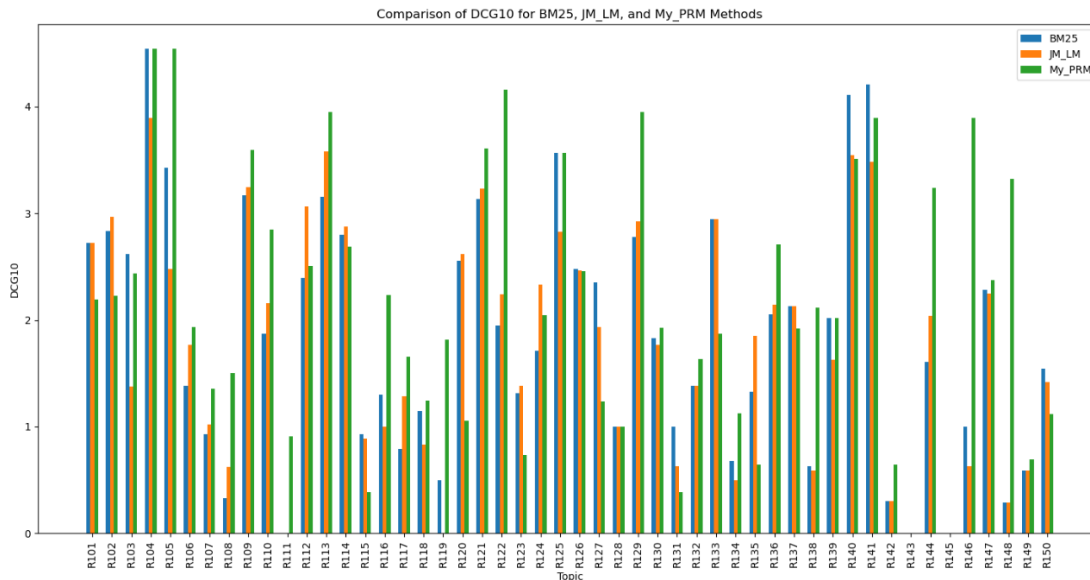


Table 3: The performance of 3 models on DCG₁₀

Topic	BM25	JM_LM	My_PRM
R101	2.721	2.721	2.196
R102	2.833	2.964	2.225
R103	2.620	1.374	2.436
R104	4.544	3.898	4.544
R105	3.425	2.482	4.544
R106	1.387	1.764	1.932
R107	0.931	1.018	1.356
R108	0.333	0.622	1.500
R109	3.167	3.249	3.594
R110	1.874	2.157	2.851
R111	0.000	0.000	0.906
R112	2.398	3.065	2.508
R113	3.152	3.579	3.953
R114	2.803	2.874	2.689
R115	0.931	0.887	0.387
R116	1.301	1.000	2.235
R117	0.789	1.289	1.657
R118	1.145	0.833	1.246
R119	0.500	0.000	1.815
R120	2.552	2.619	1.053
R121	3.134	3.233	3.606
R122	1.950	2.239	4.157
R123	1.315	1.387	0.732
R124	1.710	2.333	2.047
R125	3.565	2.824	3.567

R126	2.482	2.464	2.456
R127	2.354	1.937	1.239
R128	1.000	1.000	1.000
R129	2.776	2.922	3.953
R130	1.833	1.764	1.931
R131	1.000	0.631	0.387
R132	1.387	1.387	1.634
R133	2.948	2.948	1.869
R134	0.676	0.500	1.122
R135	1.325	1.851	0.645
R136	2.053	2.145	2.707
R137	2.131	2.131	1.920
R138	0.631	0.590	2.119
R139	2.018	1.631	2.018
R140	4.113	3.544	3.511
R141	4.210	3.485	3.895
R142	0.301	0.301	0.645
R143	0.000	0.000	0.000
R144	1.604	2.040	3.238
R145	0.000	0.000	0.000
R146	1.000	0.631	3.898
R147	2.282	2.251	2.377
R148	0.289	0.289	3.323
R149	0.590	0.590	0.690
R150	1.544	1.418	1.116
Average	1.833	1.777	2.149

Graph 3: The comparison of 3 models on DCG@10



1. Mean Average Precision (MAP):

- MAP is a widely used evaluation metric in information retrieval (IR). It summarizes rankings from multiple queries by calculating the average of the precision values at each relevant document position.

2. Precision@10:

- Precision@10 focuses on the top 10 documents in the ranking. It measures how many of the top 10 retrieved documents are relevant.
3. **Discounted Cumulative Gain at Rank 10 (DCG@10):**
- DCG@10 uses graded relevance as a measure of usefulness. It considers both the relevance of documents and their position in the ranking.

From the above three tables and bar graphs, we note that My_PRM achieved the highest average score than BM25 and JM_LM.

Discussion

Task 6. Recommend a model based on significance test and your analysis.

To compare the performance of different models, we conducted t-tests on the evaluation results, specifically focusing on Average Precision (AP), DCG@10, and Precision@10 (P@10). The t-tests were performed to determine if there were significant differences between the models.

A t-test assesses whether the means of two groups are statistically different from each other. We used the t-statistic and p-value to determine the significance. A lower p-value (typically ≤ 0.05) indicates stronger evidence against the null hypothesis, meaning there is a significant difference between the models.

For each pair, the hypotheses are defined specifically as below:

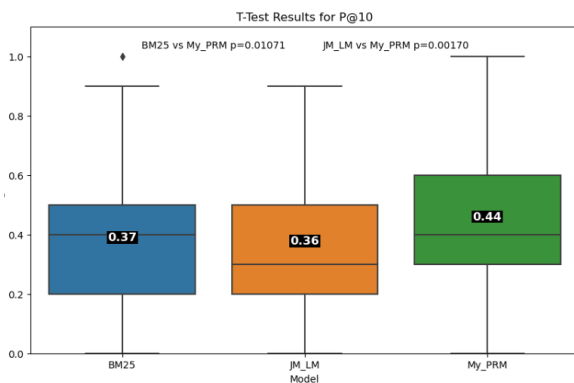
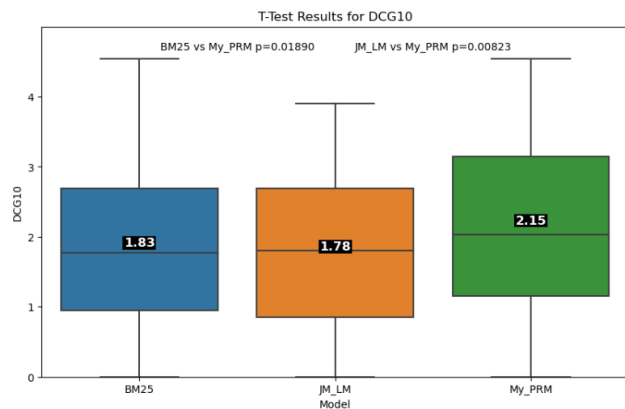
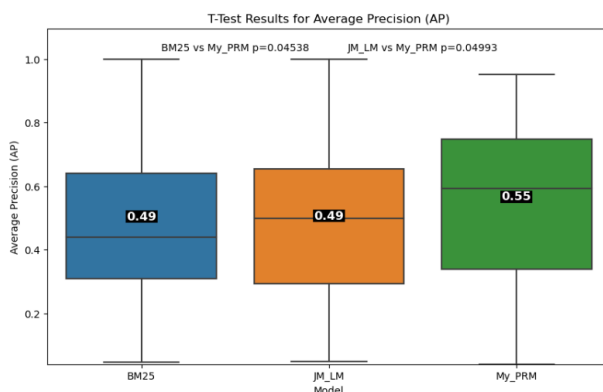
1. **BM25 vs. My_PRM:**
 - **H0:** The mean performance of the BM25 model is equal to the mean performance of the My_PRM model.
 - **Ha:** The mean performance of the My_PRM model is better than BM25 model.
2. **JM_LM vs. My_PRM:**
 - **H0:** The mean performance of the JM_LM model is equal to the mean performance of the My_PRM model.
 - **Ha:** The mean performance of the My_PRM model is better than JM_LM model.
3. **BM25 vs. JM_LM:**
 - **H0:** The mean performance of the BM25 model is equal to the mean performance of the JM_LM model.
 - **Ha:** The mean performance of the JM_LM model is better than BM25 model.

Below is the table summarizing the evaluation metric, t-statistics, p-value, t-test results of all paired models.

Comparison	Evaluation Metric	T-Statistic	P-Value	T-test result	Conclusion
BM25 vs. My_PRM	Average Precision	-2.05354748	0.04537647	Since p-value is lower than 0.05, reject H0	My_PRM is better
JM_LM vs. My_PRM	Average Precision	-2.01020467	0.04993105	Since p-value is lower than 0.05, reject H0	My_PRM is better
BM25 vs. JM_LM	Average Precision	-0.14050995	0.88883312	Since p-value is larger than 0.05, cannot reject H0	No significant difference
BM25 vs. My_PRM	DCG10	-2.42794368	0.01890322	Since p-value is lower than 0.05, reject H0	My_PRM is better

JM_LM vs. My_PRM	DCG10	-2.75450873	0.00822638	Since p-value is lower than 0.05, reject H0	My_PRM is better
BM25 vs. JM_LM	DCG10	1.00584786	0.31943021	Since p-value is larger than 0.05, cannot reject H0	No significant difference
BM25 vs. My_PRM	P@ 10	-2.65349877	0.01070939	Since p-value is lower than 0.05, reject H0	My_PRM is better
JM_LM vs. My_PRM	P@ 10	-3.32086076	0.00170073	Since p-value is lower than 0.05, reject H0	My_PRM is better
BM25 vs. JM_LM	P@ 10	1.097687584	0.27770587	Since p-value is larger than 0.05, cannot reject H0	No significant difference

Below boxplots shows the comparison of Precision at 10 (P@10), DCG10 and average precision scores for three different models: BM25, JM_LM, and My_PRM.



The results of the paired t-test indicate that My_PRM outperforms BM25 and JM_LM and shows its excellent performance.

• Analysis

Better Performance: BM25 and JM_LM models were regularly beaten by the My_PRM model on all important parameters, such as average precision, DCG10, and precision@10. This suggests that My_PRM offers document rankings that are more accurate and pertinent. As observed from the boxplots above, the mean of the three measures in My_PRM is higher than those in the other two models.

Statistical Significance: My_PRM's superiority is confirmed by the paired t-tests performed for each metric. The tests' p-values were less than the conventional significance level of 0.05, proving that the gains made by My_PRM are statistically significant and not the result of

chance.

Accuracy: My PRM is a powerful and accurate model for document ranking, as seen by its consistent performance across a variety of evaluation metrics. For real-world applications where several factors of ranking quality matter, this quality is essential.

Better model structure: Below is the feature comparison table highlighting the advantages of the My_PRM model over the BM25 and JM_LM –

Feature	BM25	JM_LM	My_PRM
Initial Ranking	Based on BM25 formula	Based on Jelinek-Mercer smoothing	Based on BM25 formula
Handling Long Queries	Parse query topic word only	Parse query topic word only	Parses full query content for richer representation and better retrieval
Query Processing	Parsed, cleaned, and stemmed	Parsed, cleaned, and stemmed	Parsed, cleaned, and stemmed; includes full query content (title, description, narrative)
Term Frequency Weighting	Uses document and query term weights	Interpolates term frequency in document and collection	Uses document and query term weights initially
Relevance Feedback	N/A	N/A	Pseudo-Relevance Feedback with top-ranked documents considered pseudo-relevant
Query Expansion	N/A	N/A	Expanded with terms from top-ranked documents, refining the query iteratively
Relevance Model Building	N/A	N/A	Builds relevance models from top-ranked documents to calculate term probabilities
Re-Ranking	Based on initial BM25 scores	Based on JM smoothing scores	Re-ranks using enhanced query terms for more accurate document ranking
Adaptability	Less adaptive to varying query lengths	Moderate adaptability with λ smoothing parameter	Highly adaptive due to iterative query refinement and expansion

From the above, we note that My_PRM enhances query representation, makes use of relevance feedback algorithms, and adopts to longer query length than BM25 and JM_LM.

- **Recommendations**

Based on the analysis, we recommend using the My_PRM model for document ranking tasks within this collection. The justification for this recommendation is supported by several factors.


- ✚ **Empirical Results:** My_PRM outperforms the other models in terms of ranking quality, with the highest scores in average precision, DCG10, and precision@10.

- ✚ **Statistical Evidence:** The strong evidence from the statistical tests indicates that the observed improvements are not due to chance.

- ✚ **Model Structure:** From a structural perspective, My_PRM not only considers the topic words of queries but also includes narratives and descriptions. Additionally, it uses relevance feedback mechanisms to identify more relevant words.

In conclusion, **My PRM** is the best model for document ranking in this task due to its superior empirical performance, statistically significant results, and better model structure.

References

- [1]  'Probabilistic relevance model'. [Online]. Available: https://en.wikipedia.org/wiki/Probabilistic_relevance_model#:~:text=It%20is%20a%20theoretical%20model,the%20query%20and%20document%20representations.

Appendix 1 (listing the top-15 documents for all queries for Model1)

Rank	QueryID	DocID	Score
1	R101	46547	2.651464274
2	R101	46974	2.651464274
3	R101	62325	1.990353845
4	R101	6146	1.173185981
5	R101	61329	1.16687995
6	R101	22170	1.023842557
7	R101	61780	0.932183641
8	R101	22513	0.794108373
9	R101	82330	0.597484129
10	R101	39496	0.484338415
11	R101	18586	0
12	R101	26642	0
13	R101	26847	0
14	R101	27577	0
15	R101	30647	0
1	R102	58476	3.386302117
2	R102	26061	3.247747158
3	R102	73038	3.216170177
4	R102	57914	3.158613138
5	R102	78836	3.12666839
6	R102	76635	2.881988681
7	R102	12769	2.779535551
8	R102	12767	2.705406809
9	R102	25096	2.473729742
10	R102	82227	2.455810094
11	R102	24515	2.380987214
12	R102	65414	2.24994154
13	R102	26611	2.129383845
14	R102	33172	2.008245519
15	R102	29908	1.982222486

Rank	QueryID	DocID	Score
1	R103	81463	2.905129679
2	R103	14314	2.852625848
3	R103	27426	2.307092001
4	R103	27106	2.158280475
5	R103	54533	2.007140512
6	R103	59459	1.977089439
7	R103	83370	1.863718176
8	R103	26385	1.798296565
9	R103	20159	1.772496027
10	R103	14069	1.680971944
11	R103	26386	1.579201301
12	R103	79396	1.558277568
13	R103	26432	1.506371928
14	R103	80988	1.226645477
15	R103	9272	1.212798943
1	R104	11930	1.374403492
2	R104	11923	1.357110167
3	R104	11481	1.3023767
4	R104	6636	1.266686915
5	R104	7502	1.261415152
6	R104	49815	1.259775417
7	R104	3837	1.25061146
8	R104	12812	1.24722216
9	R104	11485	1.228496979
10	R104	3827	1.203962188
11	R104	12479	1.191235444
12	R104	15077	1.184543708
13	R104	25205	1.18225264
14	R104	65394	1.181248673
15	R104	4306	1.18092976
1	R105	86042	1.35411281
2	R105	86961	1.35411281
3	R105	15744	1.346324815
4	R105	33961	1.332964876
5	R105	64966	1.320461519
6	R105	51493	1.29460174
7	R105	49633	1.259673547
8	R105	48148	1.248336741
9	R105	80483	1.233897787
10	R105	80484	1.233897787
11	R105	80884	1.233897787
12	R105	65210	1.20961987
13	R105	3008	1.1854294
14	R105	5225	1.163059363
15	R105	5226	1.163059363
1	R106	16841	3.697987271
2	R106	73731	3.061971633

Rank	QueryID	DocID	Score
3	R106	49023	2.896310348
4	R106	51441	2.183860409
5	R106	16723	2.105639473
6	R106	55891	2.103149821
7	R106	7914	2.058474333
8	R106	35782	1.808822637
9	R106	68829	1.794772158
10	R106	6386	1.240065472
11	R106	76556	1.230942833
12	R106	6409	1.222570799
13	R106	8085	1.165201147
14	R106	41521	1.072555396
15	R106	2800	0.996795864
1	R107	51576	2.7656393
2	R107	71157	2.386400257
3	R107	77936	2.273108908
4	R107	79950	2.144090202
5	R107	59244	1.898317234
6	R107	67107	1.607428466
7	R107	67411	1.607428466
8	R107	86459	1.128090008
9	R107	31404	0.580746932
10	R107	41791	0.563515539
11	R107	69377	0.562560333
12	R107	37330	0.558475633
13	R107	18297	0.549733398
14	R107	18529	0.548367074
15	R107	78164	0.546709032
1	R108	82290	2.483188881
2	R108	5176	2.408862898
3	R108	5177	2.408862898
4	R108	2829	2.39449922
5	R108	10759	2.244000119
6	R108	50314	2.062641805
7	R108	10810	1.958086348
8	R108	80518	1.919786526
9	R108	22169	1.885783004
10	R108	83932	1.845192038
11	R108	19961	1.671955795
12	R108	22662	1.645174313
13	R108	26058	1.626669878
14	R108	54765	1.606848653
15	R108	8780	1.515907465
1	R109	16953	1.768944146
2	R109	26073	1.645079531
3	R109	61540	1.634666753
4	R109	4933	1.487301014

Rank	QueryID	DocID	Score
5	R109	64476	1.400285897
6	R109	16575	1.383425954
7	R109	67717	1.362442345
8	R109	78626	1.286355435
9	R109	34684	1.242986422
10	R109	23398	1.216861035
11	R109	73598	1.152816598
12	R109	15776	1.144048448
13	R109	51139	1.05596251
14	R109	24340	1.046927985
15	R109	29314	1.033922206
1	R110	38920	3.229986301
2	R110	41791	2.766713475
3	R110	42439	2.760910465
4	R110	72320	2.727402057
5	R110	85147	2.702983631
6	R110	38356	2.671161749
7	R110	78595	2.487134112
8	R110	16827	2.435664157
9	R110	71167	2.262263839
10	R110	25590	2.235513804
11	R110	28185	2.204992552
12	R110	18552	2.149502673
13	R110	23408	2.149502673
14	R110	6163	2.001254581
15	R110	82926	1.668397948
1	R111	18100	1.562450296
2	R111	18163	1.562450296
3	R111	18164	1.562450296
4	R111	19440	1.34599745
5	R111	28923	1.331432472
6	R111	18706	1.32246953
7	R111	75763	1.320247619
8	R111	13057	1.206711843
9	R111	22382	1.160381711
10	R111	52020	1.147064375
11	R111	18859	1.130896683
12	R111	47049	0.931936937
13	R111	66338	0.893777313
14	R111	45422	0.850258355
15	R111	46485	0.84386677
1	R112	58748	2.626203105
2	R112	78115	2.562252296
3	R112	53046	2.54098399
4	R112	70466	2.477339664
5	R112	59464	2.477237359
6	R112	81350	2.405163946

Rank	QueryID	DocID	Score
7	R112	78269	2.379263034
8	R112	85553	2.264113874
9	R112	83201	2.262591881
10	R112	5166	1.817766758
11	R112	12858	1.813431667
12	R112	48882	0.86054008
13	R112	54208	0.836871498
14	R112	18246	0.831000204
15	R112	22021	0.816612095
1	R113	17030	2.628178984
2	R113	17217	2.628178984
3	R113	13685	2.470849083
4	R113	25500	2.363039104
5	R113	2323	1.765126597
6	R113	4053	1.765126597
7	R113	7722	1.751147829
8	R113	2451	1.649810472
9	R113	79404	1.635858565
10	R113	13690	1.609857561
11	R113	40234	1.576244545
12	R113	37816	1.445056253
13	R113	86541	1.393912854
14	R113	76103	1.376900661
15	R113	36588	1.365711863
1	R114	18879	2.521145463
2	R114	38452	2.15576867
3	R114	60810	2.111750496
4	R114	72354	1.665320481
5	R114	26700	1.658176336
6	R114	25303	1.64783122
7	R114	28014	1.623867271
8	R114	21110	1.515506858
9	R114	38330	1.393540381
10	R114	78708	1.318776251
11	R114	24474	1.048589634
12	R114	58634	1.020564435
13	R114	37685	0.885897118
14	R114	55340	0.641670596
15	R114	57956	0.641670596
1	R115	50020	3.33678666
2	R115	58137	2.543417731
3	R115	76605	2.49885127
4	R115	61647	2.156390054
5	R115	48443	2.064260704
6	R115	52066	2.045812251
7	R115	5671	1.918745105
8	R115	38068	1.902177649

Rank	QueryID	DocID	Score
9	R115	43604	1.899000787
10	R115	79396	1.881035977
11	R115	46714	1.857329907
12	R115	53345	1.37122173
13	R115	15606	0.549590875
14	R115	44746	0.5494638
15	R115	77790	0.544628721
1	R116	22679	1.564239627
2	R116	20188	1.254362809
3	R116	77250	1.090480824
4	R116	79559	1.062484882
5	R116	85143	1.021690747
6	R116	73431	0.947001065
7	R116	76422	0.919536751
8	R116	79585	0.912776442
9	R116	74010	0.90166392
10	R116	71555	0.87675883
11	R116	79369	0.876431906
12	R116	86514	0.867283144
13	R116	82616	0.867146146
14	R116	78863	0.854885896
15	R116	74017	0.849615243
1	R117	54844	0.25321075
2	R117	9544	0.247941647
3	R117	71102	0.239557255
4	R117	67204	0.236626344
5	R117	37425	0.226954599
6	R117	81694	0.209788612
7	R117	79525	0.204432284
8	R117	65692	0.180588576
9	R117	48566	0.174602795
10	R117	60811	0.152985024
11	R117	62586	0.088242145
12	R117	21113	0
13	R117	74800	0
1	R118	80908	1.976576877
2	R118	61677	1.769247318
3	R118	20763	1.759235094
4	R118	64892	1.665194763
5	R118	28903	1.54376843
6	R118	43173	1.488702311
7	R118	2906	1.45273181
8	R118	84529	1.171363498
9	R118	47129	1.064295125
10	R118	50269	1.027112004
11	R118	28826	0.990581547
12	R118	35104	0.917985739

Rank	QueryID	DocID	Score
13	R118	33112	0.837031016
14	R118	9544	0.818160119
15	R118	43241	0.797815525
1	R119	13151	0
2	R119	13902	0
3	R119	26074	0
4	R119	28125	0
5	R119	34033	0
6	R119	35105	0
7	R119	36308	0
8	R119	38614	0
9	R119	41275	0
10	R119	51080	0
11	R119	51082	0
12	R119	51488	0
13	R119	52752	0
14	R119	62314	0
15	R119	62324	0
1	R120	39184	3.049824137
2	R120	48981	2.488634003
3	R120	62781	2.408626725
4	R120	43393	1.61003349
5	R120	64800	1.432281201
6	R120	4527	1.373040591
7	R120	75135	1.251796632
8	R120	12555	1.152432437
9	R120	24062	1.132216337
10	R120	78091	1.098768498
11	R120	21067	1.013763447
12	R120	78092	0.955535707
13	R120	25918	0.930473746
14	R120	20830	0.780338887
15	R120	69629	0.714907838
1	R121	17004	3.323044535
2	R121	13294	3.228569518
3	R121	15074	3.228569518
4	R121	7320	2.994416716
5	R121	13309	2.917883018
6	R121	49980	2.877006711
7	R121	3885	2.822242465
8	R121	79604	2.786527765
9	R121	14772	2.770211578
10	R121	7289	2.688405387
11	R121	75621	2.64415644
12	R121	79669	2.64415644
13	R121	6604	2.607401274
14	R121	43938	2.576230593

Rank	QueryID	DocID	Score
15	R121	44187	2.560998134
1	R122	38244	4.175799835
2	R122	74775	3.682578344
3	R122	77971	3.166082065
4	R122	84331	3.097277318
5	R122	13984	2.981940016
6	R122	79418	2.708162598
7	R122	72052	2.673678771
8	R122	73186	2.673678771
9	R122	76564	2.673678771
10	R122	77921	2.441223215
11	R122	78866	2.409982835
12	R122	80435	2.19715322
13	R122	54175	2.122677443
14	R122	44818	1.992538942
15	R122	54747	1.896893729
1	R123	37440	12.64265295
2	R123	26681	10.65844913
3	R123	29079	9.22957832
4	R123	78018	8.656518171
5	R123	13485	8.121667616
6	R123	30089	8.080668089
7	R123	55119	7.176396379
8	R123	10515	6.453643209
9	R123	20315	6.239663542
10	R123	47025	6.212425277
11	R123	85953	4.766398205
12	R123	70379	4.661642359
13	R123	48085	3.65312047
14	R123	42609	3.576055454
15	R123	42697	3.576055454
1	R124	21511	1.110137409
2	R124	19285	1.045484164
3	R124	25328	0.99421539
4	R124	74036	0.994062165
5	R124	65604	0.95309197
6	R124	4528	0.934777318
7	R124	9981	0.885513144
8	R124	7225	0.846195516
9	R124	53830	0.82145596
10	R124	61262	0.798129142
11	R124	30653	0.78868252
12	R124	46194	0.786076528
13	R124	14377	0.786021983
14	R124	13147	0.652006852
15	R124	78418	0.579118246
1	R125	77907	2.050065436

Rank	QueryID	DocID	Score
2	R125	82519	1.993801907
3	R125	77381	1.979386846
4	R125	82085	1.792770016
5	R125	80922	1.789830657
6	R125	25833	1.751533553
7	R125	76899	1.729695688
8	R125	82106	1.710842169
9	R125	27974	1.695289617
10	R125	9011	1.33109325
11	R125	7970	1.301665254
12	R125	73371	0.925385124
13	R125	27960	0.870289051
14	R125	39122	0.859755207
15	R125	28549	0.852232379
1	R126	15228	0.114619034
2	R126	45159	0.111615522
3	R126	22250	0.108325542
4	R126	78479	0.100569667
5	R126	38191	0.10035524
6	R126	14933	0.098816755
7	R126	76216	0.098590908
8	R126	59083	0.098090797
9	R126	67522	0.095451155
10	R126	60588	0.09436934
11	R126	80102	0.093547594
12	R126	75770	0.092110161
13	R126	56499	0.091132702
14	R126	9384	0.091101612
15	R126	4984	0.090834521
1	R127	67104	2.333486271
2	R127	10706	2.077826298
3	R127	40234	2.012489508
4	R127	75738	1.772199595
5	R127	59095	1.759333647
6	R127	60141	1.759333647
7	R127	40292	1.733572967
8	R127	70010	1.297742196
9	R127	26026	1.280289195
10	R127	26144	1.280289195
11	R127	56741	1.182195721
12	R127	58487	0.926571463
13	R127	62215	0.903612592
14	R127	68825	0.844554615
15	R127	62401	0.83957316
1	R128	3740	2.591790988
2	R128	66961	2.385942323
3	R128	64070	2.373185034

Rank	QueryID	DocID	Score
4	R128	64033	2.351826191
5	R128	70233	2.342396829
6	R128	61170	2.317291906
7	R128	24214	2.283049579
8	R128	66923	2.228015097
9	R128	75201	2.224400432
10	R128	27215	2.184970172
11	R128	60528	1.649372625
12	R128	4933	1.099882321
13	R128	26085	1.099054383
14	R128	26088	1.098345025
15	R128	58476	1.037751331
1	R129	46494	2.295425603
2	R129	72826	2.144206022
3	R129	72899	2.121658772
4	R129	76669	1.866534269
5	R129	16900	1.535876221
6	R129	75465	1.366882121
7	R129	20126	1.366770011
8	R129	44081	1.350107836
9	R129	41297	1.342234451
10	R129	35148	1.338905829
11	R129	21791	1.333512295
12	R129	75593	1.326833394
13	R129	55197	1.281072563
14	R129	83428	1.221065926
15	R129	83946	1.207486623
1	R130	15776	2.645302119
2	R130	70556	2.527291313
3	R130	18235	2.506491539
4	R130	69863	2.36536535
5	R130	60528	2.05198504
6	R130	75483	2.026333686
7	R130	46458	1.692311292
8	R130	60483	1.578951089
9	R130	4063	0.840495118
10	R130	47530	0.718375289
11	R130	68037	0.680748932
12	R130	73550	0.680280393
13	R130	11715	0.671476024
14	R130	22585	0.668538232
15	R130	22584	0.485117149
1	R131	26568	1.479391818
2	R131	59502	1.43949269
3	R131	58957	1.429960456
4	R131	60118	1.41435057
5	R131	16170	1.381368068

Rank	QueryID	DocID	Score
6	R131	62536	1.348957367
7	R131	31504	1.190531203
8	R131	77349	1.185391118
9	R131	80756	1.176774112
10	R131	31515	1.162345378
11	R131	41975	1.081425312
12	R131	73034	1.078037079
13	R131	40119	0.878472915
14	R131	41857	0.878472915
15	R131	63056	0.455017459
1	R132	7486	5.118993094
2	R132	85553	2.900947633
3	R132	48981	2.847285094
4	R132	9843	2.639263584
5	R132	67115	2.143734966
6	R132	79565	1.821296537
7	R132	68370	1.679151337
8	R132	66981	1.617801141
9	R132	39184	1.560833678
10	R132	68780	1.489576828
11	R132	53133	1.48447926
12	R132	64294	1.464809524
13	R132	65174	1.428340381
14	R132	78577	1.392302367
15	R132	68784	1.372560936
1	R133	74695	3.513071333
2	R133	35853	3.208168757
3	R133	54844	2.661023707
4	R133	74681	2.51909614
5	R133	84398	2.089916126
6	R133	71102	1.84534273
7	R133	9544	1.776202467
8	R133	32662	1.606768061
9	R133	6939	1.496247665
10	R133	82290	1.357407797
11	R133	35755	1.328246801
12	R133	54925	1.222895687
13	R133	54814	1.188199601
14	R133	62586	1.156064678
15	R133	37425	1.038057631
1	R134	59659	2.73995694
2	R134	58438	1.422090941
3	R134	84627	1.381391743
4	R134	75562	1.259237765
5	R134	73397	1.217771056
6	R134	26253	1.179766014
7	R134	21172	1.165139233

Rank	QueryID	DocID	Score
8	R134	46142	1.128931677
9	R134	69112	1.106750256
10	R134	74790	1.092443687
11	R134	60928	1.077812955
12	R134	14567	1.06330593
13	R134	74778	0.990218718
14	R134	53308	0.985593193
15	R134	43806	0.985225326
1	R135	57876	0.857414612
2	R135	71894	0.653758942
3	R135	42848	0.651103485
4	R135	2738	0.504589943
5	R135	31784	0
6	R135	34733	0
7	R135	40131	0
8	R135	41556	0
9	R135	42847	0
10	R135	44221	0
11	R135	46122	0
12	R135	49229	0
13	R135	52721	0
14	R135	59518	0
15	R135	65311	0
1	R136	35104	2.495617364
2	R136	79605	2.450067593
3	R136	41487	2.228695659
4	R136	43962	2.012676129
5	R136	19488	1.834762162
6	R136	8768	1.748782307
7	R136	46422	1.566879783
8	R136	42907	1.473670401
9	R136	48692	1.404649913
10	R136	15404	1.32420615
11	R136	83238	1.2264247
12	R136	76750	1.145595312
13	R136	3641	1.144969669
14	R136	76779	1.14450895
15	R136	3610	1.08549696
1	R137	66879	3.180333495
2	R137	85322	2.348860697
3	R137	61252	2.338889199
4	R137	39136	1.918936666
5	R137	81418	1.792995721
6	R137	80107	1.772818126
7	R137	78455	1.681480899
8	R137	80039	1.624970303
9	R137	46874	1.606104832

Rank	QueryID	DocID	Score
10	R137	27297	1.577639215
11	R137	82828	1.577392393
12	R137	4650	1.546827627
13	R137	78483	1.469195525
14	R137	74200	1.462540767
15	R137	21644	1.39747779
1	R138	77971	1.617063408
2	R138	2552	1.366978697
3	R138	20899	1.365587108
4	R138	37732	1.350615631
5	R138	49872	1.300705108
6	R138	50734	1.249580696
7	R138	40341	1.189265519
8	R138	65378	1.159873315
9	R138	56318	1.149054871
10	R138	22407	1.131554114
11	R138	47714	1.06926811
12	R138	62620	1.042132215
13	R138	21088	1.010089636
14	R138	74775	1.00905545
15	R138	56045	0.947001538
1	R139	71102	2.332051272
2	R139	67204	2.160306624
3	R139	50076	1.464346036
4	R139	41231	1.43149091
5	R139	62586	1.317039304
6	R139	58507	1.308397223
7	R139	54844	1.185385294
8	R139	9544	1.182691037
9	R139	79525	1.044845556
10	R139	81694	0.973833402
11	R139	37425	0.926774451
12	R139	65692	0.922816805
13	R139	48566	0.888517774
14	R139	35853	0.855064278
15	R139	79033	0.642264202
1	R140	25063	1.89197778
2	R140	58343	1.872777477
3	R140	45396	1.840233253
4	R140	6857	1.807676071
5	R140	22321	1.79231058
6	R140	44400	1.752694583
7	R140	43139	1.565789781
8	R140	11825	1.499977226
9	R140	18151	1.282663589
10	R140	18393	1.282663589
11	R140	35886	0.361060504

Rank	QueryID	DocID	Score
12	R140	13885	0.360728517
13	R140	33779	0.358034719
14	R140	35214	0.358034719
15	R140	36005	0.351591579
1	R141	65284	2.104286349
2	R141	60155	1.74131528
3	R141	62292	1.729773946
4	R141	60221	1.680630751
5	R141	58503	1.570731331
6	R141	62212	1.416435674
7	R141	75561	1.401710212
8	R141	55848	1.38627943
9	R141	52215	1.295105132
10	R141	62213	1.161609904
11	R141	35053	0.996739196
12	R141	58514	0.872946118
13	R141	39864	0.816448638
14	R141	59351	0.807358209
15	R141	75471	0.799000272
1	R142	83869	1.103335866
2	R142	26011	1.087163755
3	R142	54126	1.044304543
4	R142	68798	1.026953594
5	R142	53188	0.999282079
6	R142	54675	0.902057449
7	R142	34410	0.81042304
8	R142	7397	0.704216448
9	R142	28354	0.583933323
10	R142	54716	0.544690823
11	R142	60740	0.522734744
12	R142	45201	0.506404515
13	R142	6158	0.470172222
14	R142	9382	0.436026886
15	R142	86201	0.42368085
1	R143	72134	2.028145273
2	R143	64896	1.899612172
3	R143	74605	1.74528662
4	R143	23698	1.7260235
5	R143	3666	1.724179269
6	R143	75564	1.638636815
7	R143	11852	1.609842646
8	R143	11744	1.512850783
9	R143	63379	1.493174385
10	R143	21265	1.355968364
11	R143	72032	1.299755362
12	R143	76550	1.299755362
13	R143	44232	0.87213933

Rank	QueryID	DocID	Score
14	R143	30864	0.840235391
15	R143	37657	0.794748607
1	R144	43899	1.990814443
2	R144	27525	1.917383791
3	R144	13135	1.760346175
4	R144	7359	1.731737804
5	R144	47896	1.671704642
6	R144	69511	1.612251697
7	R144	74200	1.507424999
8	R144	36318	1.46364926
9	R144	35258	1.445539326
10	R144	12579	1.443424382
11	R144	12658	1.440678787
12	R144	26257	1.440662093
13	R144	19626	1.440623724
14	R144	3364	1.419913403
15	R144	27535	1.418246651
1	R145	17654	1.732564954
2	R145	14501	1.677940491
3	R145	4300	0.490999132
4	R145	54291	0.463491979
5	R145	11106	0.462860613
6	R145	28025	0.452475604
7	R145	6237	0.451411111
8	R145	32472	0.449863298
9	R145	67785	0.440350979
10	R145	41461	0.435489085
11	R145	35545	0.432489046
12	R145	51917	0.432400515
13	R145	22691	0.431378699
14	R145	21077	0.426821266
15	R145	27970	0.421778691
1	R146	12035	1.979319782
2	R146	23687	1.786119134
3	R146	72072	1.677451142
4	R146	49904	1.657095999
5	R146	58000	1.417073018
6	R146	38821	1.297347999
7	R146	24287	1.153856617
8	R146	70836	1.083189393
9	R146	12196	1.050122011
10	R146	54065	0.990547406
11	R146	12396	0.937543545
12	R146	11944	0
13	R146	12476	0
14	R146	14541	0
15	R146	15478	0

Rank	QueryID	DocID	Score
1	R147	56691	1.812699399
2	R147	27441	1.713431319
3	R147	64153	1.661168729
4	R147	34675	1.6367091
5	R147	73146	1.588804055
6	R147	59813	1.441932023
7	R147	27911	1.423014118
8	R147	15581	1.384859869
9	R147	15582	1.384859869
10	R147	41481	1.363482675
11	R147	81325	1.336152601
12	R147	76089	1.308418256
13	R147	76092	1.308418256
14	R147	75754	1.307356389
15	R147	23293	1.288798366
1	R148	84829	0.639297941
2	R148	58009	0.62014696
3	R148	8622	0.619289412
4	R148	9133	0.619289412
5	R148	24044	0.614801635
6	R148	31619	0.595444104
7	R148	51453	0.593130067
8	R148	51260	0.570085274
9	R148	52489	0.569012004
10	R148	60293	0.567409655
11	R148	75205	0.543827206
12	R148	85344	0.502338299
13	R148	52661	0.457204993
14	R148	85663	0.45227588
15	R148	55609	0.407049215
1	R149	41538	2.339040479
2	R149	14196	2.216204952
3	R149	39661	2.204788322
4	R149	39926	2.06269117
5	R149	32400	1.937492416
6	R149	25529	1.640603466
7	R149	62392	1.404618924
8	R149	3936	1.383938569
9	R149	17390	1.269764442
10	R149	17360	1.239753445
11	R149	72985	0.909204744
12	R149	25220	0.864123364
13	R149	30390	0.615662342
14	R149	52385	0.580922113
15	R149	40902	0.561350061
1	R150	19409	2.229113962
2	R150	64482	2.147733513

Rank	QueryID	DocID	Score
3	R150	53361	2.04456991
4	R150	19411	2.010314138
5	R150	19420	1.969608478
6	R150	19885	1.802241607
7	R150	3148	1.717723042
8	R150	50614	1.572531585
9	R150	54538	1.381875437
10	R150	53269	1.249603218
11	R150	39907	0.95757898
12	R150	26648	0.881635072
13	R150	85997	0.77405376
14	R150	23361	0.758220013
15	R150	6318	0.605529534

Appendix 2 (listing the top-15 documents for all queries for Model2)

Rank	QueryID	DocID	Score
1	R101	46547	0.000108455
2	R101	46974	0.000108455
3	R101	62325	1.82614E-05
4	R101	61329	2.52275E-06
5	R101	6146	1.91459E-06
6	R101	22170	9.53326E-07
7	R101	61780	9.29325E-07
8	R101	22513	8.58564E-07
9	R101	82330	4.52415E-07
10	R101	39496	3.34373E-07
11	R101	18586	1.01008E-07
12	R101	26642	1.01008E-07
13	R101	26847	1.01008E-07
14	R101	27577	1.01008E-07
15	R101	30647	1.01008E-07
1	R102	78836	8.52054E-08
2	R102	58476	6.18248E-08
3	R102	57914	3.16462E-08
4	R102	26061	2.67083E-08
5	R102	76635	2.15866E-08
6	R102	73038	1.87103E-08
7	R102	12769	1.36214E-08
8	R102	12767	1.07416E-08
9	R102	25096	6.61491E-09
10	R102	24515	5.28588E-09
11	R102	82227	4.37239E-09
12	R102	26611	2.51089E-09
13	R102	65414	1.99265E-09
14	R102	33172	1.84017E-09
15	R102	29908	1.72393E-09

Rank	QueryID	DocID	Score
1	R103	14314	9.7013E-08
2	R103	81463	8.47014E-08
3	R103	14069	1.43147E-08
4	R103	26385	1.33807E-08
5	R103	27106	1.167E-08
6	R103	27426	9.96663E-09
7	R103	26386	7.49263E-09
8	R103	79396	6.03245E-09
9	R103	20159	5.35923E-09
10	R103	26432	4.48869E-09
11	R103	54533	4.29157E-09
12	R103	59459	3.90615E-09
13	R103	83370	3.21685E-09
14	R103	80988	2.50256E-09
15	R103	63966	2.19021E-09
1	R104	11930	1.62339E-07
2	R104	3837	1.56537E-07
3	R104	11923	1.31614E-07
4	R104	49815	9.31508E-08
5	R104	25205	8.75657E-08
6	R104	11960	7.92527E-08
7	R104	11481	7.10877E-08
8	R104	65394	6.65781E-08
9	R104	12812	5.8857E-08
10	R104	9591	5.59802E-08
11	R104	7502	5.10274E-08
12	R104	6636	4.87342E-08
13	R104	76690	4.58405E-08
14	R104	11485	4.53967E-08
15	R104	75949	4.15528E-08
1	R105	64966	4.11891E-06
2	R105	33961	1.73949E-06
3	R105	15744	1.50138E-06
4	R105	65210	1.25588E-06
5	R105	51493	1.05109E-06
6	R105	86042	8.00399E-07
7	R105	86961	8.00399E-07
8	R105	3008	7.52248E-07
9	R105	49633	3.62036E-07
10	R105	48148	2.62927E-07
11	R105	80483	9.69633E-08
12	R105	80484	9.69633E-08
13	R105	80884	9.69633E-08
14	R105	2493	9.09413E-08
15	R105	2494	9.09413E-08
1	R106	16841	2.23572E-09
2	R106	73731	4.29516E-11

Rank	QueryID	DocID	Score
3	R106	49023	3.89769E-11
4	R106	16723	3.13941E-11
5	R106	55891	2.29156E-11
6	R106	35782	1.85095E-11
7	R106	6409	1.18736E-11
8	R106	68829	1.16822E-11
9	R106	76556	1.06976E-11
10	R106	8085	1.00695E-11
11	R106	6386	8.70409E-12
12	R106	51441	6.37282E-12
13	R106	76481	4.52632E-12
14	R106	41521	4.43871E-12
15	R106	31281	4.37696E-12
1	R107	71157	1.38067E-09
2	R107	79950	1.13857E-10
3	R107	59244	1.00877E-10
4	R107	51576	9.25742E-11
5	R107	77936	3.2457E-11
6	R107	67673	1.85159E-11
7	R107	31404	1.74936E-11
8	R107	37330	1.48712E-11
9	R107	69377	1.40045E-11
10	R107	78164	1.18545E-11
11	R107	66686	1.10655E-11
12	R107	28185	1.09614E-11
13	R107	17736	1.00214E-11
14	R107	71159	1.00214E-11
15	R107	36212	9.41785E-12
1	R108	82290	1.35607E-10
2	R108	10759	1.00528E-10
3	R108	2829	8.88779E-11
4	R108	5176	4.22142E-11
5	R108	5177	4.22142E-11
6	R108	50314	1.48985E-11
7	R108	10810	1.19494E-11
8	R108	80518	8.96682E-12
9	R108	54765	5.30658E-12
10	R108	19961	4.9195E-12
11	R108	22662	3.6187E-12
12	R108	26058	3.29473E-12
13	R108	33970	2.92449E-12
14	R108	22169	2.57305E-12
15	R108	8780	2.35975E-12
1	R109	26073	1.68338E-07
2	R109	16953	1.26452E-07
3	R109	64476	9.86969E-08
4	R109	61540	8.02409E-08

Rank	QueryID	DocID	Score
5	R109	67717	7.35661E-08
6	R109	16575	3.98547E-08
7	R109	4933	2.17192E-08
8	R109	23398	1.99764E-08
9	R109	24340	1.93035E-08
10	R109	34684	1.57171E-08
11	R109	78626	1.5456E-08
12	R109	65289	1.40858E-08
13	R109	29314	1.30645E-08
14	R109	25832	1.12813E-08
15	R109	73598	7.88659E-09
1	R110	38356	1.33082E-11
2	R110	38920	1.31513E-11
3	R110	42439	1.05424E-11
4	R110	71167	1.0042E-11
5	R110	78595	4.79864E-12
6	R110	85147	4.25834E-12
7	R110	72320	3.64814E-12
8	R110	16827	2.90652E-12
9	R110	41791	2.61178E-12
10	R110	30597	7.61754E-13
11	R110	27404	6.65877E-13
12	R110	25590	5.93171E-13
13	R110	28185	3.88986E-13
14	R110	86950	2.39816E-13
15	R110	13264	2.29903E-13
1	R111	28923	1.14023E-06
2	R111	18706	1.06595E-06
3	R111	75763	1.0489E-06
4	R111	18100	6.13245E-07
5	R111	18163	6.13245E-07
6	R111	18164	6.13245E-07
7	R111	13057	5.59285E-07
8	R111	22382	4.62316E-07
9	R111	19440	4.28534E-07
10	R111	18859	3.43916E-07
11	R111	52020	3.19601E-07
12	R111	47049	2.27897E-07
13	R111	66338	2.06825E-07
14	R111	45422	1.85956E-07
15	R111	46485	1.83135E-07
1	R112	78269	0.000119907
2	R112	78115	0.000117828
3	R112	58748	0.000109968
4	R112	59464	0.000105539
5	R112	85553	8.78216E-05
6	R112	53046	7.72202E-05

Rank	QueryID	DocID	Score
7	R112	81350	3.35305E-05
8	R112	70466	3.11658E-05
9	R112	83201	9.16297E-06
10	R112	12858	2.18632E-06
11	R112	5166	1.99857E-06
12	R112	48882	1.28629E-06
13	R112	68780	9.79874E-07
14	R112	22021	9.59784E-07
15	R112	39944	8.17733E-07
1	R113	13685	2.46254E-07
2	R113	37816	2.15248E-07
3	R113	25500	1.25069E-07
4	R113	17030	1.09066E-07
5	R113	17217	1.09066E-07
6	R113	13690	6.23753E-08
7	R113	7722	5.21504E-08
8	R113	40234	3.02763E-08
9	R113	37815	2.95481E-08
10	R113	79404	2.28426E-08
11	R113	2323	2.08448E-08
12	R113	4053	2.08448E-08
13	R113	43262	1.89325E-08
14	R113	76103	1.52627E-08
15	R113	2451	1.37849E-08
1	R114	18879	5.7218E-07
2	R114	38452	1.06879E-07
3	R114	60810	5.34066E-08
4	R114	72354	4.41038E-08
5	R114	21110	1.99241E-08
6	R114	25303	1.05313E-08
7	R114	26700	9.712E-09
8	R114	28014	8.66541E-09
9	R114	38330	6.12195E-09
10	R114	78708	5.8958E-09
11	R114	58634	2.17377E-09
12	R114	24474	2.08404E-09
13	R114	37685	1.42344E-09
14	R114	55340	9.24827E-10
15	R114	57956	9.24827E-10
1	R115	58137	3.292E-08
2	R115	50020	2.03439E-08
3	R115	61647	1.48104E-08
4	R115	48443	1.45304E-08
5	R115	76605	1.24086E-08
6	R115	52066	1.12143E-08
7	R115	38068	5.50238E-09
8	R115	43604	4.8016E-09

Rank	QueryID	DocID	Score
9	R115	5671	4.60037E-09
10	R115	46714	4.4625E-09
11	R115	79396	2.76855E-09
12	R115	77790	6.69188E-10
13	R115	44746	6.60401E-10
14	R115	15606	4.05072E-10
15	R115	53345	4.00256E-10
1	R116	22679	1.98418E-05
2	R116	71542	1.4982E-05
3	R116	71543	1.4982E-05
4	R116	71555	1.45094E-05
5	R116	20188	1.10951E-05
6	R116	74017	1.08798E-05
7	R116	82616	1.0673E-05
8	R116	79369	9.27068E-06
9	R116	78863	8.97758E-06
10	R116	86514	8.6766E-06
11	R116	66894	7.41078E-06
12	R116	52940	6.84267E-06
13	R116	82598	6.76326E-06
14	R116	82279	6.73203E-06
15	R116	52741	6.46409E-06
1	R117	71102	0.00024074
2	R117	54844	0.00012408
3	R117	9544	0.000106319
4	R117	67204	9.91609E-05
5	R117	79525	6.6344E-05
6	R117	65692	5.3156E-05
7	R117	81694	4.93887E-05
8	R117	48566	3.96456E-05
9	R117	37425	2.93562E-05
10	R117	60811	1.76248E-05
11	R117	21113	4.38996E-06
12	R117	62586	4.20933E-06
13	R117	74800	2.41636E-06
1	R118	80908	4.8545E-09
2	R118	28903	2.19122E-09
3	R118	20763	1.79375E-09
4	R118	64892	1.45989E-09
5	R118	61677	1.25558E-09
6	R118	2906	9.59175E-10
7	R118	43173	8.58565E-10
8	R118	35104	7.0724E-10
9	R118	47129	5.68271E-10
10	R118	84529	5.26292E-10
11	R118	43241	4.5338E-10
12	R118	50269	4.42332E-10

Rank	QueryID	DocID	Score
13	R118	28826	3.60104E-10
14	R118	9544	3.13509E-10
15	R118	33112	1.90455E-10
1	R119	71547	0.064162185
2	R119	82567	0.063352571
3	R119	38614	0.06275894
4	R119	28125	0.061794445
5	R119	52752	0.061794445
6	R119	13151	0.059598905
7	R119	51080	0.030332749
8	R119	51082	0.030332749
9	R119	62324	0.029652381
10	R119	68177	0.0226185
11	R119	65942	0.022178332
12	R119	51488	0.020540593
13	R119	67199	0.018017877
14	R119	26074	0.016051233
15	R119	13902	0.012955548
1	R120	48981	1.53115E-06
2	R120	39184	7.50282E-07
3	R120	62781	6.45555E-07
4	R120	43393	9.24805E-08
5	R120	4527	6.81079E-08
6	R120	78091	4.24981E-08
7	R120	78092	1.85502E-08
8	R120	12555	1.79026E-08
9	R120	75135	1.66575E-08
10	R120	21067	1.52645E-08
11	R120	69629	1.26912E-08
12	R120	71406	1.26912E-08
13	R120	69633	1.2112E-08
14	R120	16241	8.17138E-09
15	R120	45819	7.79566E-09
1	R121	17004	8.45576E-09
2	R121	13294	7.47154E-09
3	R121	15074	7.47154E-09
4	R121	13309	2.02756E-09
5	R121	7320	1.84618E-09
6	R121	15415	8.94698E-10
7	R121	49980	8.43792E-10
8	R121	14772	5.32187E-10
9	R121	43938	5.04378E-10
10	R121	3885	3.87521E-10
11	R121	15410	3.75391E-10
12	R121	79604	3.34168E-10
13	R121	14556	3.26639E-10
14	R121	9384	1.83065E-10

Rank	QueryID	DocID	Score
15	R121	9820	1.817E-10
1	R122	38244	1.39764E-07
2	R122	84331	3.04152E-08
3	R122	77971	2.31609E-08
4	R122	74775	2.2879E-08
5	R122	79418	2.07388E-08
6	R122	13984	1.30743E-08
7	R122	72052	6.98793E-09
8	R122	73186	6.98793E-09
9	R122	76564	6.98793E-09
10	R122	78866	2.75845E-09
11	R122	77921	1.8385E-09
12	R122	80435	9.83672E-10
13	R122	54175	6.45085E-10
14	R122	44818	5.28693E-10
15	R122	54747	3.9862E-10
1	R123	37440	4.65907E-43
2	R123	29079	2.92338E-45
3	R123	26681	1.34882E-45
4	R123	30089	7.36688E-46
5	R123	10515	7.57216E-47
6	R123	42609	6.76298E-47
7	R123	42697	6.76298E-47
8	R123	13485	6.40431E-47
9	R123	48085	4.38879E-47
10	R123	29677	2.52421E-47
11	R123	9382	2.23028E-47
12	R123	85067	1.98624E-47
13	R123	85953	1.65027E-47
14	R123	15623	1.62925E-47
15	R123	48360	1.54362E-47
1	R124	9981	2.96724E-05
2	R124	74036	1.99557E-05
3	R124	21511	1.66987E-05
4	R124	53830	1.53421E-05
5	R124	7225	1.46579E-05
6	R124	25328	1.12896E-05
7	R124	4528	9.4462E-06
8	R124	19285	9.26525E-06
9	R124	65604	6.80386E-06
10	R124	65633	5.14045E-06
11	R124	46194	4.78553E-06
12	R124	61262	4.55367E-06
13	R124	2345	4.44466E-06
14	R124	2346	4.37972E-06
15	R124	13147	3.83566E-06
1	R125	76899	0.000100335

Rank	QueryID	DocID	Score
2	R125	77907	7.53613E-05
3	R125	82519	7.24978E-05
4	R125	25833	6.92748E-05
5	R125	77381	5.42136E-05
6	R125	27974	5.20628E-05
7	R125	80922	3.14713E-05
8	R125	82085	3.05701E-05
9	R125	82106	2.7198E-05
10	R125	71913	1.99706E-05
11	R125	28549	1.05312E-05
12	R125	9011	9.03327E-06
13	R125	7970	8.12932E-06
14	R125	73371	7.61198E-06
15	R125	27960	7.55874E-06
1	R126	45159	0.000354373
2	R126	38191	0.000347036
3	R126	14933	0.000273648
4	R126	15228	0.000271295
5	R126	67522	0.000252163
6	R126	60588	0.000213088
7	R126	9384	0.000208273
8	R126	78479	0.000179993
9	R126	59083	0.000158025
10	R126	56499	0.000157599
11	R126	83225	0.000149079
12	R126	71223	0.000144515
13	R126	22250	0.000143121
14	R126	80102	0.000126086
15	R126	76216	0.000118732
1	R127	40234	2.50924E-08
2	R127	67104	1.68526E-08
3	R127	10706	1.23436E-08
4	R127	59095	8.78223E-09
5	R127	60141	8.78223E-09
6	R127	40292	7.73368E-09
7	R127	75738	5.60571E-09
8	R127	26026	1.93383E-09
9	R127	26144	1.93383E-09
10	R127	58487	1.50848E-09
11	R127	56741	1.49306E-09
12	R127	70010	1.4792E-09
13	R127	62401	1.41201E-09
14	R127	62215	1.27994E-09
15	R127	68825	5.66643E-10
1	R128	3740	3.83772E-07
2	R128	64033	1.42527E-07
3	R128	27215	1.38889E-07

Rank	QueryID	DocID	Score
4	R128	64070	1.00216E-07
5	R128	24214	9.6701E-08
6	R128	66961	7.51705E-08
7	R128	61170	6.07933E-08
8	R128	70233	4.92888E-08
9	R128	66923	3.40429E-08
10	R128	75201	3.32545E-08
11	R128	26085	2.98719E-08
12	R128	26088	2.95248E-08
13	R128	60528	2.57857E-08
14	R128	4933	1.87626E-08
15	R128	32366	1.77805E-08
1	R129	46494	1.40489E-07
2	R129	76669	8.12493E-08
3	R129	72826	7.45015E-08
4	R129	72899	6.77157E-08
5	R129	35148	2.92425E-08
6	R129	44081	2.34882E-08
7	R129	41297	1.96268E-08
8	R129	75465	1.68582E-08
9	R129	83428	1.66677E-08
10	R129	21791	1.42561E-08
11	R129	75593	1.32418E-08
12	R129	67400	1.03601E-08
13	R129	83268	8.987E-09
14	R129	82198	6.51442E-09
15	R129	75590	6.48001E-09
1	R130	15776	1.82658E-07
2	R130	70556	5.6008E-08
3	R130	69863	2.99223E-08
4	R130	18235	2.96402E-08
5	R130	60528	2.91906E-08
6	R130	75483	9.01884E-09
7	R130	46458	5.03276E-09
8	R130	60483	3.87724E-09
9	R130	4063	1.29531E-09
10	R130	73550	1.25805E-09
11	R130	22585	8.99453E-10
12	R130	68037	3.72625E-10
13	R130	11715	3.44062E-10
14	R130	47530	1.75978E-10
15	R130	22584	1.02109E-10
1	R131	16170	8.5161E-05
2	R131	26568	8.14441E-05
3	R131	60118	5.26107E-05
4	R131	59502	5.01189E-05
5	R131	58957	4.20447E-05

Rank	QueryID	DocID	Score
6	R131	62536	3.96567E-05
7	R131	77349	2.88644E-05
8	R131	80756	2.55683E-05
9	R131	31504	1.99992E-05
10	R131	31515	1.39136E-05
11	R131	41975	1.182E-05
12	R131	13484	7.96647E-06
13	R131	73034	7.09031E-06
14	R131	62835	2.87393E-06
15	R131	62836	2.79535E-06
1	R132	7486	1.56943E-07
2	R132	85553	3.78027E-09
3	R132	48981	2.58012E-09
4	R132	9843	8.62757E-10
5	R132	67115	2.76017E-10
6	R132	68370	8.71627E-11
7	R132	68780	7.98564E-11
8	R132	79565	7.33158E-11
9	R132	53133	4.71936E-11
10	R132	39184	4.66081E-11
11	R132	66981	4.52203E-11
12	R132	78577	4.37447E-11
13	R132	64294	3.76518E-11
14	R132	65174	3.32154E-11
15	R132	4017	3.21532E-11
1	R133	74695	1.29761E-09
2	R133	35853	1.14582E-09
3	R133	54844	6.23415E-10
4	R133	84398	2.08145E-10
5	R133	74681	1.03643E-10
6	R133	6939	2.80104E-11
7	R133	9544	2.20926E-11
8	R133	71102	1.64259E-11
9	R133	82290	1.61845E-11
10	R133	35755	1.57076E-11
11	R133	32662	1.14341E-11
12	R133	65692	9.17966E-12
13	R133	54814	7.25212E-12
14	R133	54925	6.0216E-12
15	R133	37425	4.43422E-12
1	R134	59659	5.95882E-11
2	R134	58438	7.65705E-12
3	R134	73397	6.26798E-12
4	R134	75562	3.27268E-12
5	R134	84627	2.78368E-12
6	R134	21172	2.71095E-12
7	R134	69112	2.23644E-12

Rank	QueryID	DocID	Score
8	R134	60928	1.77905E-12
9	R134	46142	1.72036E-12
10	R134	53308	1.33415E-12
11	R134	5870	1.32904E-12
12	R134	3169	1.2191E-12
13	R134	74790	8.43561E-13
14	R134	14567	7.83793E-13
15	R134	43806	7.60188E-13
1	R135	78305	1.20018E-08
2	R135	65465	1.10039E-08
3	R135	66118	1.10039E-08
4	R135	31784	8.5212E-09
5	R135	68756	8.10666E-09
6	R135	80891	7.81152E-09
7	R135	78998	6.48971E-09
8	R135	79698	6.48971E-09
9	R135	52721	6.43472E-09
10	R135	65311	6.31059E-09
11	R135	65501	6.28358E-09
12	R135	67275	5.8136E-09
13	R135	67433	5.8136E-09
14	R135	34733	5.19186E-09
15	R135	86691	5.02387E-09
1	R136	35104	4.55976E-08
2	R136	79605	1.60546E-08
3	R136	43962	7.34645E-09
4	R136	41487	7.33954E-09
5	R136	19488	3.74061E-09
6	R136	46422	3.4645E-09
7	R136	48692	2.70419E-09
8	R136	8768	2.41394E-09
9	R136	42907	2.05901E-09
10	R136	58438	1.88005E-09
11	R136	34428	1.46899E-09
12	R136	83238	1.11197E-09
13	R136	15404	1.09912E-09
14	R136	76750	8.30234E-10
15	R136	76779	8.27626E-10
1	R137	66879	3.76227E-07
2	R137	85322	4.84431E-08
3	R137	61252	2.22858E-08
4	R137	81418	1.22814E-08
5	R137	80107	1.21215E-08
6	R137	46874	9.46144E-09
7	R137	80039	7.51064E-09
8	R137	39136	6.76808E-09
9	R137	74200	5.99868E-09

Rank	QueryID	DocID	Score
10	R137	4650	5.97058E-09
11	R137	78455	5.89673E-09
12	R137	82828	5.12051E-09
13	R137	78483	4.76179E-09
14	R137	53354	4.45804E-09
15	R137	27297	4.01997E-09
1	R138	20899	1.3244E-10
2	R138	49872	1.30982E-10
3	R138	77971	8.37849E-11
4	R138	37732	4.98736E-11
5	R138	56318	2.63679E-11
6	R138	50734	1.66269E-11
7	R138	40341	6.42739E-12
8	R138	65378	6.01947E-12
9	R138	2552	5.88752E-12
10	R138	75463	5.39635E-12
11	R138	56045	4.66962E-12
12	R138	56248	4.66962E-12
13	R138	47714	4.34157E-12
14	R138	62620	1.88909E-12
15	R138	48174	1.66475E-12
1	R139	71102	1.31702E-06
2	R139	67204	3.50724E-07
3	R139	54844	4.90632E-08
4	R139	9544	4.2018E-08
5	R139	79525	2.60285E-08
6	R139	50076	2.16924E-08
7	R139	65692	2.0693E-08
8	R139	81694	1.90681E-08
9	R139	48566	1.5253E-08
10	R139	41231	1.36501E-08
11	R139	58507	1.31795E-08
12	R139	35853	1.17033E-08
13	R139	37425	1.08441E-08
14	R139	62586	9.85863E-09
15	R139	79033	3.33834E-09
1	R140	6857	9.45286E-05
2	R140	58343	3.65313E-05
3	R140	44400	2.91816E-05
4	R140	25063	2.39094E-05
5	R140	18151	1.85722E-05
6	R140	18393	1.85722E-05
7	R140	45396	1.83627E-05
8	R140	22321	1.21414E-05
9	R140	43139	9.25288E-06
10	R140	11825	6.21626E-06
11	R140	13885	6.14847E-06

Rank	QueryID	DocID	Score
12	R140	35886	5.10091E-06
13	R140	16137	4.50412E-06
14	R140	33779	4.0927E-06
15	R140	35214	4.0927E-06
1	R141	65284	1.62655E-07
2	R141	60155	1.27048E-07
3	R141	62292	9.84932E-08
4	R141	60221	9.43336E-08
5	R141	51691	8.79418E-08
6	R141	59351	6.79155E-08
7	R141	39864	4.94111E-08
8	R141	9954	4.86336E-08
9	R141	55848	4.64237E-08
10	R141	39304	4.04094E-08
11	R141	39972	3.69365E-08
12	R141	75561	3.6746E-08
13	R141	75471	3.28511E-08
14	R141	39883	2.88136E-08
15	R141	75641	2.79573E-08
1	R142	34410	8.85958E-06
2	R142	83869	8.4155E-06
3	R142	26011	8.00366E-06
4	R142	54126	7.0515E-06
5	R142	68798	6.7144E-06
6	R142	53188	6.22449E-06
7	R142	7397	5.74982E-06
8	R142	54675	4.85866E-06
9	R142	28354	4.05968E-06
10	R142	54716	3.30958E-06
11	R142	60740	2.98587E-06
12	R142	45201	2.77757E-06
13	R142	6158	2.13036E-06
14	R142	9382	2.09956E-06
15	R142	86201	2.00746E-06
1	R143	72134	1.27901E-07
2	R143	23698	9.74067E-08
3	R143	74605	8.98254E-08
4	R143	3666	5.76002E-08
5	R143	64896	5.6716E-08
6	R143	75564	4.96344E-08
7	R143	44232	3.03282E-08
8	R143	11744	2.20846E-08
9	R143	63379	2.07468E-08
10	R143	30864	1.7804E-08
11	R143	11852	1.4223E-08
12	R143	37657	1.12452E-08
13	R143	21265	1.05102E-08

Rank	QueryID	DocID	Score
14	R143	7041	9.42041E-09
15	R143	72032	9.30244E-09
1	R144	43899	1.08317E-07
2	R144	36318	6.52181E-08
3	R144	13135	6.44364E-08
4	R144	69511	5.33192E-08
5	R144	19626	5.29495E-08
6	R144	27525	5.08809E-08
7	R144	35258	3.91531E-08
8	R144	26257	3.78862E-08
9	R144	7359	3.60059E-08
10	R144	27535	2.83828E-08
11	R144	47896	2.28405E-08
12	R144	12579	2.08626E-08
13	R144	12658	2.05927E-08
14	R144	3364	1.9008E-08
15	R144	74200	1.79036E-08
1	R145	32232	2.03581E-08
2	R145	17654	1.83027E-08
3	R145	4300	1.7669E-08
4	R145	14501	1.58682E-08
5	R145	11106	1.04654E-08
6	R145	32472	1.02144E-08
7	R145	6237	8.5308E-09
8	R145	51917	6.74818E-09
9	R145	28025	6.48357E-09
10	R145	67785	5.35934E-09
11	R145	54291	4.84956E-09
12	R145	27970	4.49194E-09
13	R145	35545	4.0582E-09
14	R145	41447	3.96936E-09
15	R145	26072	3.52563E-09
1	R146	23687	5.23648E-06
2	R146	12035	4.197E-06
3	R146	72072	2.63989E-06
4	R146	58000	1.88255E-06
5	R146	49904	1.38602E-06
6	R146	38821	9.97722E-07
7	R146	70836	6.739E-07
8	R146	24287	6.0206E-07
9	R146	54065	4.86938E-07
10	R146	12196	4.51079E-07
11	R146	12396	4.14241E-07
12	R146	11944	4.33003E-08
13	R146	12476	4.33003E-08
14	R146	14541	4.33003E-08
15	R146	15478	4.33003E-08

Rank	QueryID	DocID	Score
1	R147	56691	0.000494668
2	R147	27441	6.49915E-05
3	R147	64153	5.96259E-05
4	R147	34675	4.70826E-05
5	R147	81325	3.19196E-05
6	R147	73146	3.18726E-05
7	R147	27911	2.78727E-05
8	R147	59813	2.72609E-05
9	R147	23293	1.73438E-05
10	R147	41481	1.51852E-05
11	R147	82440	1.49883E-05
12	R147	27845	1.09838E-05
13	R147	15581	9.50584E-06
14	R147	15582	9.50584E-06
15	R147	27921	8.7118E-06
1	R148	84829	0.023365634
2	R148	8622	0.018915834
3	R148	9133	0.018915834
4	R148	58009	0.018537854
5	R148	24044	0.017730125
6	R148	31619	0.015930759
7	R148	51260	0.01464175
8	R148	51453	0.014539221
9	R148	52489	0.014438222
10	R148	60293	0.014144072
11	R148	75205	0.013771343
12	R148	52661	0.009005386
13	R148	85344	0.007550841
14	R148	85663	0.005923419
15	R148	83678	0.005550841
1	R149	41538	4.69713E-08
2	R149	39661	3.58726E-08
3	R149	14196	3.36838E-08
4	R149	25529	2.77986E-08
5	R149	32400	2.4064E-08
6	R149	39926	1.8911E-08
7	R149	62392	8.74993E-09
8	R149	3936	4.6783E-09
9	R149	17390	3.74908E-09
10	R149	17360	3.38804E-09
11	R149	30390	1.6574E-09
12	R149	25220	1.38331E-09
13	R149	72985	1.29886E-09
14	R149	40902	1.03933E-09
15	R149	16941	1.01736E-09
1	R150	19409	3.43982E-05
2	R150	53361	3.40979E-05

Rank	QueryID	DocID	Score
3	R150	19411	2.89463E-05
4	R150	19420	2.70281E-05
5	R150	64482	2.47384E-05
6	R150	19885	1.24703E-05
7	R150	3148	9.89138E-06
8	R150	26648	5.51052E-06
9	R150	39907	4.00438E-06
10	R150	50614	3.96895E-06
11	R150	54538	3.7199E-06
12	R150	53269	2.13789E-06
13	R150	23361	1.28521E-06
14	R150	85997	1.10145E-06
15	R150	6318	7.57907E-07

Appendix 3 (listing the top-15 documents for all queries for Model3)

Rank	Query ID	Document ID	Score
1	R101	61780	430.0039276
2	R101	46547	367.9841156
3	R101	46974	367.9841156
4	R101	62325	317.3144338
5	R101	61329	157.2099012
6	R101	6146	131.7186683
7	R101	39496	130.9467229
8	R101	22170	127.8490127
9	R101	82330	125.7794003
10	R101	22513	119.5702293
11	R101	18586	0
12	R101	26642	0
13	R101	26847	0
14	R101	27577	0
15	R101	30647	0
1	R102	12769	207.0649371
2	R102	12767	199.805883
3	R102	58476	190.7148872
4	R102	78836	159.0622373
5	R102	26061	156.0355009
6	R102	11922	155.6935342
7	R102	73038	127.298878
8	R102	76635	124.159189
9	R102	70306	106.3312391
10	R102	24485	100.7674467
11	R102	82227	97.83255062
12	R102	4358	97.68036079
13	R102	4933	90.67539
14	R102	29911	85.14420209
15	R102	10182	83.97150076
1	R103	81463	416.5279546
2	R103	26386	321.0010535
3	R103	26385	267.241234
4	R103	82912	248.3090282
5	R103	26432	229.8749764
6	R103	27426	194.0320146
7	R103	59459	191.7583605
8	R103	83370	185.4663203
9	R103	14069	170.6260074
10	R103	20159	158.308242
11	R103	79396	140.8837976
12	R103	14314	135.6859681
13	R103	78569	129.705095
14	R103	54533	128.4390323
15	R103	27106	128.0135245
1	R104	14713	154.956595

Rank	Query ID	Document ID	Score
2	R104	15082	153.2396143
3	R104	4881	145.7321018
4	R104	11930	140.3782037
5	R104	11485	140.2693911
6	R104	11923	137.3247011
7	R104	25096	129.2262888
8	R104	6636	128.4742302
9	R104	7502	127.6566651
10	R104	3833	127.541767
11	R104	13313	126.9368709
12	R104	29904	125.2831318
13	R104	4306	125.2097002
14	R104	13314	124.4422491
15	R104	35396	124.236803
1	R105	5225	299.9938575
2	R105	5226	299.9938575
3	R105	5004	290.2472956
4	R105	2493	286.0219529
5	R105	2494	286.0219529
6	R105	3008	153.1970359
7	R105	86042	130.1364074
8	R105	86961	130.1364074
9	R105	48148	127.1125066
10	R105	49633	125.6746943
11	R105	64966	92.7575608
12	R105	15744	85.5623092
13	R105	5223	81.00280983
14	R105	80483	79.64356741
15	R105	80484	79.64356741
1	R106	16841	851.5053202
2	R106	16723	444.798149
3	R106	47424	438.1928661
4	R106	56686	436.837038
5	R106	76481	432.9096817
6	R106	53259	426.0842392
7	R106	50516	424.7549543
8	R106	86068	395.48427
9	R106	41355	349.8090253
10	R106	35782	302.9525028
11	R106	55891	296.5114851
12	R106	67169	286.3122146
13	R106	32528	274.7448108
14	R106	31281	268.5996936
15	R106	32963	266.4122682
1	R107	79950	329.3356735
2	R107	86459	163.477524
3	R107	71157	121.8111087

Rank	Query ID	Document ID	Score
4	R107	59244	116.829879
5	R107	47338	110.9466637
6	R107	77936	96.22776786
7	R107	59521	86.66661571
8	R107	48786	84.19050756
9	R107	63648	84.01829151
10	R107	45168	74.4469344
11	R107	23437	73.06814924
12	R107	78164	69.88603751
13	R107	3169	69.39064776
14	R107	20760	62.0729145
15	R107	25590	60.70391738
1	R108	19961	444.3308644
2	R108	82290	336.7488708
3	R108	10810	323.473219
4	R108	5176	311.5727482
5	R108	5177	311.5727482
6	R108	33970	267.4358955
7	R108	2829	265.4853072
8	R108	49600	236.9651425
9	R108	80518	236.3652956
10	R108	22169	227.5958533
11	R108	54925	220.1452716
12	R108	54814	219.6331549
13	R108	26058	202.9589881
14	R108	82392	200.8324058
15	R108	8780	199.9611844
1	R109	64476	476.5691098
2	R109	16953	459.5569024
3	R109	67717	436.0153723
4	R109	26073	420.5587878
5	R109	65289	324.2083426
6	R109	61540	284.3495237
7	R109	4933	283.9321985
8	R109	15776	268.1357225
9	R109	51139	249.6670107
10	R109	82229	239.2220707
11	R109	24340	223.1178214
12	R109	23398	213.651725
13	R109	16575	208.4296582
14	R109	58676	192.9653243
15	R109	34684	190.7098648
1	R110	41791	308.2458843
2	R110	42439	295.8101788
3	R110	38356	289.3004524
4	R110	85147	261.3955034
5	R110	28185	192.953565

Rank	Query ID	Document ID	Score
6	R110	6163	188.8331779
7	R110	25590	185.5592082
8	R110	18552	181.7524006
9	R110	23408	181.7524006
10	R110	82926	160.7583551
11	R110	60334	143.3174328
12	R110	71167	34.83482382
13	R110	27404	33.92265527
14	R110	69330	33.03496957
15	R110	30597	31.63748883
1	R111	18100	395.4443622
2	R111	18163	395.4443622
3	R111	18164	395.4443622
4	R111	19440	298.1459746
5	R111	18859	281.7118885
6	R111	49617	87.78592443
7	R111	7776	80.85485468
8	R111	45422	69.40330614
9	R111	46485	69.10496698
10	R111	46968	69.10496698
11	R111	52020	54.08422186
12	R111	23643	49.81136481
13	R111	49650	49.40036469
14	R111	66338	49.04144072
15	R111	22660	48.94738512
1	R112	38350	167.3708568
2	R112	78115	136.1172202
3	R112	78269	111.0172212
4	R112	83201	88.74312291
5	R112	38383	85.22094301
6	R112	59464	77.27337389
7	R112	39944	70.975562
8	R112	81350	68.33296968
9	R112	85553	68.0256428
10	R112	70466	61.07793779
11	R112	53046	60.77422425
12	R112	58748	50.69701104
13	R112	12858	45.82200375
14	R112	39961	43.46427998
15	R112	5166	39.47674423
1	R113	17030	671.3183669
2	R113	17217	671.3183669
3	R113	13685	644.1565275
4	R113	25500	452.6500011
5	R113	24251	356.4307604
6	R113	40234	330.9179997
7	R113	79404	330.8359744

Rank	Query ID	Document ID	Score
8	R113	13690	326.8686299
9	R113	86541	280.7356455
10	R113	36588	277.6033074
11	R113	2451	270.8906816
12	R113	7722	261.6572203
13	R113	2323	259.2279122
14	R113	4053	259.2279122
15	R113	76103	250.3448239
1	R114	25303	561.8405165
2	R114	38452	561.2018584
3	R114	26700	547.6845155
4	R114	28014	539.8163174
5	R114	60810	485.4052439
6	R114	18879	423.2842914
7	R114	38330	384.4432167
8	R114	21110	372.3584921
9	R114	78708	316.7949145
10	R114	58634	308.5764416
11	R114	72354	281.3158667
12	R114	55340	271.810596
13	R114	57956	271.810596
14	R114	18405	194.1978151
15	R114	16921	175.4231435
1	R115	86068	533.9985859
2	R115	79601	530.9947013
3	R115	67715	529.3332751
4	R115	70456	528.5131428
5	R115	76605	462.0192133
6	R115	44554	437.0730949
7	R115	62192	433.3284209
8	R115	65206	430.2944535
9	R115	35782	427.4639656
10	R115	43604	424.5459937
11	R115	9611	415.5076177
12	R115	57152	408.4275599
13	R115	62221	386.5566665
14	R115	58137	385.8239266
15	R115	79396	383.032056
1	R116	22679	247.2600948
2	R116	74010	168.5850343
3	R116	82679	153.8578841
4	R116	66894	153.6803658
5	R116	52940	114.0444553
6	R116	86514	109.5620862
7	R116	79369	107.6853828
8	R116	68718	107.5835132
9	R116	78863	105.445299

Rank	Query ID	Document ID	Score
10	R116	68716	104.6455173
11	R116	82598	104.2565666
12	R116	71542	98.50230816
13	R116	82279	95.37082275
14	R116	71555	95.18706678
15	R116	81084	93.6947156
1	R117	9544	17.32681684
2	R117	54844	15.67439886
3	R117	67204	14.98254065
4	R117	71102	14.90000884
5	R117	37425	12.605373
6	R117	79525	11.89583439
7	R117	81694	11.78909193
8	R117	65692	10.59704129
9	R117	48566	10.24177773
10	R117	60811	9.694818325
11	R117	62586	7.068722524
12	R117	21113	0
13	R117	74800	0
1	R118	80908	87.11612696
2	R118	20763	72.84174744
3	R118	28903	67.84589946
4	R118	2906	64.97441642
5	R118	43173	57.71783354
6	R118	35104	42.07006469
7	R118	43241	39.44263184
8	R118	9544	38.19248892
9	R118	72892	29.7784624
10	R118	37614	29.59376917
11	R118	84529	29.34586013
12	R118	50269	29.26149324
13	R118	63369	28.8049282
14	R118	28826	28.47981208
15	R118	26058	27.90117324
1	R119	13151	0
2	R119	13902	0
3	R119	26074	0
4	R119	28125	0
5	R119	34033	0
6	R119	35105	0
7	R119	36308	0
8	R119	38614	0
9	R119	41275	0
10	R119	51080	0
11	R119	51082	0
12	R119	51488	0
13	R119	52752	0

Rank	Query ID	Document ID	Score
14	R119	62314	0
15	R119	62324	0
1	R120	39184	79.64735301
2	R120	62781	69.41028943
3	R120	48981	68.78828285
4	R120	43393	42.17191514
5	R120	64800	27.74956259
6	R120	4527	27.01358467
7	R120	24062	25.13260968
8	R120	78091	24.35714791
9	R120	12555	23.81506382
10	R120	75135	22.82443402
11	R120	78092	21.5789596
12	R120	21067	20.21862644
13	R120	20830	19.78512777
14	R120	25918	19.24331906
15	R120	69629	18.26000166
1	R121	13294	151.1386837
2	R121	15074	151.1386837
3	R121	17004	133.2701921
4	R121	13309	131.0318007
5	R121	7320	123.7239241
6	R121	14772	120.3200004
7	R121	49980	112.3271458
8	R121	3885	111.8271942
9	R121	7289	110.5330183
10	R121	75621	105.0157125
11	R121	79669	105.0157125
12	R121	79604	103.4887996
13	R121	6604	103.2614494
14	R121	44187	96.25907055
15	R121	43938	95.43170932
1	R122	38244	97.96681626
2	R122	72052	63.14597214
3	R122	73186	63.14597214
4	R122	76564	63.14597214
5	R122	74775	49.32195392
6	R122	77971	41.40768612
7	R122	13984	41.40163858
8	R122	79418	39.89507751
9	R122	84331	38.89051674
10	R122	77921	37.9101438
11	R122	78866	36.69757125
12	R122	80435	35.71401356
13	R122	54175	33.4320253
14	R122	54747	32.24006002
15	R122	44818	30.94613884

Rank	Query ID	Document ID	Score
1	R123	37440	460.2299645
2	R123	26681	411.9375469
3	R123	29079	361.3302558
4	R123	78018	336.4009634
5	R123	30089	325.0616203
6	R123	13485	263.6047008
7	R123	47025	257.3662872
8	R123	20315	248.8622932
9	R123	55119	230.7941163
10	R123	10515	204.3245234
11	R123	70379	187.138837
12	R123	48085	165.4224393
13	R123	48360	154.1556897
14	R123	47027	133.7762219
15	R123	46933	130.6096944
1	R124	21511	60.96312592
2	R124	19285	53.5800827
3	R124	4528	51.8947786
4	R124	74036	50.17109503
5	R124	65604	45.08722934
6	R124	25328	45.01984495
7	R124	7225	42.83514914
8	R124	9981	42.67039288
9	R124	61262	42.57129242
10	R124	14377	41.99054305
11	R124	53830	40.27154621
12	R124	46194	39.77357171
13	R124	30653	37.12635091
14	R124	13147	36.58081521
15	R124	69236	35.17280508
1	R125	77907	104.4958435
2	R125	77381	97.19872928
3	R125	82519	92.52730828
4	R125	27974	89.21556618
5	R125	82085	85.85860545
6	R125	82106	84.77844313
7	R125	80922	84.46136944
8	R125	25833	84.03790257
9	R125	76899	63.89322314
10	R125	9011	63.19852189
11	R125	7970	62.88493814
12	R125	39122	51.9049363
13	R125	73371	49.46589229
14	R125	27960	45.60694978
15	R125	28549	41.40433975
1	R126	15228	8.478698625
2	R126	22250	8.080731999

Rank	Query ID	Document ID	Score
3	R126	14933	7.786751015
4	R126	38191	7.663162555
5	R126	67522	6.875891851
6	R126	59083	6.827799952
7	R126	45159	6.643827697
8	R126	76216	6.440392439
9	R126	78479	6.418165719
10	R126	63280	6.223426802
11	R126	56499	6.162664925
12	R126	80102	6.079148351
13	R126	60588	6.047493722
14	R126	9384	6.027548066
15	R126	4984	5.977923125
1	R127	67104	124.1579176
2	R127	10706	104.6315087
3	R127	59095	101.6038492
4	R127	60141	101.6038492
5	R127	40234	79.31700409
6	R127	75738	78.58315236
7	R127	40292	76.67106988
8	R127	26026	70.4516285
9	R127	26144	70.4516285
10	R127	58487	56.85179481
11	R127	62215	56.20917052
12	R127	62401	53.69469559
13	R127	56741	41.17271833
14	R127	59087	37.06324385
15	R127	70010	36.98138994
1	R128	3740	129.4141039
2	R128	64070	118.8333448
3	R128	61170	114.695594
4	R128	64033	100.5556859
5	R128	70233	96.46779137
6	R128	66961	95.78605271
7	R128	27215	92.70202333
8	R128	24214	89.61456416
9	R128	66923	83.62784206
10	R128	75201	81.46148755
11	R128	60528	58.23139708
12	R128	26085	36.04014102
13	R128	26088	35.98646304
14	R128	4933	35.67279564
15	R128	58476	34.83880889
1	R129	75593	82.38608453
2	R129	75465	82.02380461
3	R129	44081	77.31305734
4	R129	72826	77.03382313

Rank	Query ID	Document ID	Score
5	R129	72899	74.9196622
6	R129	46494	71.36055528
7	R129	21791	65.18562276
8	R129	35148	63.59308234
9	R129	41297	63.5515691
10	R129	75590	63.22007005
11	R129	78782	60.7595851
12	R129	83428	59.55666061
13	R129	83958	59.27925422
14	R129	82198	58.99524574
15	R129	83312	58.75250915
1	R130	15776	105.1854442
2	R130	70556	94.58644333
3	R130	69863	85.55229634
4	R130	18235	83.04179175
5	R130	75483	65.22166306
6	R130	46458	59.81038145
7	R130	60528	43.72044705
8	R130	60483	33.57777109
9	R130	47530	26.68359523
10	R130	73550	23.45159918
11	R130	4063	22.96920581
12	R130	11715	21.70562299
13	R130	68037	21.46142109
14	R130	22585	20.84391304
15	R130	22584	19.88099479
1	R131	26568	33.4136302
2	R131	58957	30.94230411
3	R131	60118	30.30117036
4	R131	59502	28.93082366
5	R131	62536	27.361806
6	R131	77349	16.87016456
7	R131	16170	16.72839647
8	R131	31504	15.32952014
9	R131	31515	14.75829481
10	R131	41975	14.71176252
11	R131	73034	13.11712691
12	R131	80756	11.60559522
13	R131	40119	8.829248844
14	R131	41857	8.780690246
15	R131	13484	7.793213361
1	R132	7486	164.8554939
2	R132	85553	92.60325146
3	R132	9843	91.85248941
4	R132	48981	89.09806778
5	R132	79565	52.91452391
6	R132	67115	46.6384239

Rank	Query ID	Document ID	Score
7	R132	53133	42.72033723
8	R132	64294	42.62344577
9	R132	65174	41.64274416
10	R132	53131	41.22955342
11	R132	68784	41.0669274
12	R132	68777	40.61186485
13	R132	4017	40.47892812
14	R132	66981	39.70212429
15	R132	71710	34.29947846
1	R133	74695	164.4741685
2	R133	35853	142.2679363
3	R133	74681	124.3622346
4	R133	54844	110.4063946
5	R133	84398	90.46364989
6	R133	9544	82.24072664
7	R133	71102	73.36843252
8	R133	54925	60.4256274
9	R133	54814	59.55289489
10	R133	67204	54.94015974
11	R133	74652	53.38029156
12	R133	65692	51.68439877
13	R133	37425	50.88088799
14	R133	6939	50.63626352
15	R133	32662	48.43333256
1	R134	58438	64.03516447
2	R134	73397	56.726021
3	R134	75562	53.83627997
4	R134	84627	48.41336201
5	R134	21172	43.09083733
6	R134	69112	42.23716698
7	R134	74790	41.49440091
8	R134	53308	41.35177743
9	R134	3169	41.02858848
10	R134	43806	40.87067591
11	R134	59659	40.54797663
12	R134	60928	40.53621796
13	R134	46142	38.97597559
14	R134	5870	37.56848194
15	R134	26253	34.96394075
1	R135	2738	0
2	R135	31784	0
3	R135	34733	0
4	R135	40131	0
5	R135	41556	0
6	R135	42847	0
7	R135	42848	0
8	R135	44221	0

Rank	Query ID	Document ID	Score
9	R135	46122	0
10	R135	49229	0
11	R135	52721	0
12	R135	57876	0
13	R135	59518	0
14	R135	65311	0
15	R135	65465	0
1	R136	35104	106.9176264
2	R136	43962	74.82230725
3	R136	46422	63.0296912
4	R136	41487	62.06325271
5	R136	79605	53.47528235
6	R136	48692	48.56523234
7	R136	8768	46.93369935
8	R136	15404	42.53390822
9	R136	19488	42.50898983
10	R136	60305	39.31656844
11	R136	76779	39.0908661
12	R136	76750	39.04744041
13	R136	60227	37.0176331
14	R136	58438	34.5569292
15	R136	55797	32.5277531
1	R137	66879	107.9223448
2	R137	80107	69.15675878
3	R137	61252	67.39404386
4	R137	85322	65.55199824
5	R137	81418	53.13191203
6	R137	82828	51.8966088
7	R137	78455	51.34521113
8	R137	78456	48.27029219
9	R137	80039	47.91213897
10	R137	78483	46.67075597
11	R137	46874	46.00054136
12	R137	85108	45.16749116
13	R137	81444	43.05182561
14	R137	21644	41.9416184
15	R137	4650	41.74962435
1	R138	77971	52.12874827
2	R138	20899	49.28852847
3	R138	2552	49.14811692
4	R138	49872	47.17524627
5	R138	37732	38.66899717
6	R138	50734	35.3270566
7	R138	56318	34.0021972
8	R138	22407	33.7945294
9	R138	65378	33.70293101
10	R138	40341	33.49366607

Rank	Query ID	Document ID	Score
11	R138	47714	32.514572
12	R138	21088	31.71077348
13	R138	62620	31.67486674
14	R138	74775	31.03676499
15	R138	56045	30.44821673
1	R139	71102	153.8620258
2	R139	67204	143.0404695
3	R139	62586	105.8697682
4	R139	58507	96.44419613
5	R139	41231	74.1007868
6	R139	50076	72.95380981
7	R139	9544	68.23555814
8	R139	54844	66.42626887
9	R139	79525	57.6586926
10	R139	81694	55.44979852
11	R139	65692	52.71976999
12	R139	37425	52.6966834
13	R139	48566	52.55979669
14	R139	35853	50.76776713
15	R139	79033	42.43787771
1	R140	44400	49.3696469
2	R140	6857	47.84674798
3	R140	22321	45.87457792
4	R140	58343	43.05068786
5	R140	25063	29.95955047
6	R140	45396	29.0035863
7	R140	18393	27.82681292
8	R140	18151	27.44520742
9	R140	43139	25.74331844
10	R140	11825	24.36707341
11	R140	35785	12.49440375
12	R140	16137	11.27566406
13	R140	38058	10.60960118
14	R140	13885	10.35356944
15	R140	55086	10.30802212
1	R141	60155	77.84396867
2	R141	60221	75.56763693
3	R141	62292	74.23485558
4	R141	65284	73.80592563
5	R141	55848	59.67308504
6	R141	58503	54.44037902
7	R141	52215	48.32314414
8	R141	62212	43.19622156
9	R141	75561	41.89422638
10	R141	62213	37.65052014
11	R141	39304	37.08798644
12	R141	83965	35.33100218

Rank	Query ID	Document ID	Score
13	R141	78720	34.90672276
14	R141	75471	34.87605286
15	R141	75641	34.83859254
1	R142	26011	51.60395799
2	R142	83869	50.59176514
3	R142	54126	49.16051118
4	R142	68798	48.96060599
5	R142	54675	41.25377291
6	R142	53188	39.64268965
7	R142	15462	0
8	R142	19908	0
9	R142	26511	0
10	R142	28354	0
11	R142	34410	0
12	R142	38696	0
13	R142	39258	0
14	R142	39840	0
15	R142	41614	0
1	R143	64896	92.98323421
2	R143	74605	88.68547417
3	R143	72134	87.32357429
4	R143	75564	85.5564109
5	R143	63379	78.81274896
6	R143	3666	68.80315136
7	R143	21265	61.2122915
8	R143	11744	59.24144455
9	R143	23698	58.15462437
10	R143	11852	56.23136961
11	R143	72032	51.74372904
12	R143	76550	51.74372904
13	R143	30864	45.83408986
14	R143	44232	45.05571554
15	R143	7041	43.62029992
1	R144	27525	79.42870021
2	R144	69511	67.76701754
3	R144	36318	61.63860696
4	R144	35258	61.54720157
5	R144	27535	60.91785654
6	R144	26257	55.01501568
7	R144	30016	50.80993381
8	R144	19626	46.37513655
9	R144	12658	46.34399794
10	R144	12579	45.79349031
11	R144	43899	43.56044497
12	R144	71904	42.45217647
13	R144	3364	41.41458267
14	R144	19705	40.95989704

Rank	Query ID	Document ID	Score
15	R144	12855	40.69798092
1	R145	4300	0.738212021
2	R145	11106	0.735456992
3	R145	32472	0.679109264
4	R145	21077	0.659515606
5	R145	32232	0.583767315
6	R145	79732	0.55174953
7	R145	22691	0.531536875
8	R145	28025	0.511475021
9	R145	41461	0.511188163
10	R145	54291	0.499015306
11	R145	11606	0.493424554
12	R145	40278	0.480377608
13	R145	17689	0.480030397
14	R145	35017	0.476615567
15	R145	8169	0.469859719
1	R146	58000	40.84713061
2	R146	38821	37.02422019
3	R146	24287	33.21821651
4	R146	12196	29.26433165
5	R146	12396	27.25176219
6	R146	54065	25.09444213
7	R146	70836	23.00903954
8	R146	12035	19.72844003
9	R146	11944	0
10	R146	12476	0
11	R146	14541	0
12	R146	15478	0
13	R146	17369	0
14	R146	19416	0
15	R146	19557	0
1	R147	56691	96.52984843
2	R147	27441	89.81269862
3	R147	64153	74.88663159
4	R147	73146	74.47289075
5	R147	34675	61.50531684
6	R147	27911	58.6061578
7	R147	59813	57.51690666
8	R147	41481	56.87046372
9	R147	27845	54.4828975
10	R147	27921	50.27586386
11	R147	75754	49.86650363
12	R147	76089	49.63613822
13	R147	76092	49.63613822
14	R147	23293	48.99243409
15	R147	81325	48.70103075
1	R148	31619	42.33235828

Rank	Query ID	Document ID	Score
2	R148	8622	41.7750463
3	R148	9133	41.69028922
4	R148	84829	40.68798354
5	R148	52489	36.97215188
6	R148	51260	36.83325535
7	R148	51453	36.40620653
8	R148	58009	35.83509019
9	R148	24044	35.23784799
10	R148	60293	30.77249603
11	R148	75205	29.94085563
12	R148	85344	29.02203228
13	R148	85663	27.29245381
14	R148	55609	25.61870216
15	R148	52661	25.13711865
1	R149	41538	143.6406493
2	R149	39661	123.4231243
3	R149	39926	122.3640253
4	R149	14196	115.2645064
5	R149	32400	73.17080606
6	R149	3936	61.78875286
7	R149	62392	59.49621133
8	R149	25529	44.94932604
9	R149	72985	41.43768926
10	R149	17390	37.87513699
11	R149	17360	37.49104993
12	R149	25220	27.84218667
13	R149	16941	0
14	R149	18090	0
15	R149	30390	0
1	R150	19409	72.2150916
2	R150	19411	67.758228
3	R150	19420	64.45291411
4	R150	53361	56.87280788
5	R150	64482	55.67070572
6	R150	19885	52.1348554
7	R150	3148	51.21467195
8	R150	50614	47.2060643
9	R150	54538	35.19280662
10	R150	39907	34.00699365
11	R150	53269	32.54138228
12	R150	85997	27.78810267
13	R150	23361	23.92536859
14	R150	26648	23.0986711
15	R150	75429	21.47109199