Home Work 3

Information Retrieval

Report

A simple statistical relevance model has been implemented in the retrieval system based on the vector relevance model, using the inverted list index that was implemented as part of homework 2. The system will retrieve the documents based on the ranked search model for each query. The vector representations of documents and queries are ranked based on the two weighting functions.

The relevance model reads a query, extracts the tokens, eliminates the stop words and each token is lemmatized.

Q1 and Q2 Refer the output files:

Query\_W1.txt: output of query, relevant documents its title, content, External Document ID and its rank for weighting function W1.

Query\_W2.txt: output of query, relevant documents its title, content, External Document ID and its rank for weighting function W2.

Query\_W1\_Vector.txt: Vector representation of queries for weighting function W1.

Query\_W2\_Vector.txt: Vector representation of queries for weighting function W2.

Doc\_W1\_Vector.txt: Vector representation of top 5 relevant documents for given queries for weighting function W1.

Doc\_W2\_Vector.txt: Vector representation of top 5 relevant documents for given queries for weighting function W2.

Results:

1. **Query:** what similarity laws must be obeyed when constructing aeroelastic models of heated high speed aircraft.

**Indexed Query:** similarity law must obey construct aeroelastic model heated high speed aircraft

**W1 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | No |
| 2 | Somewhat |
| 3 | Yes |
| 4 | Somewhat |
| 5 | Yes |

**W2 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | No |
| 3 | Yes |
| 4 | Yes |
| 5 | Yes |

2. **Query:** what are the structural and aeroelastic problems associated with flight of high speed aircraft

**Indexed Query** : structural aeroelastic problem associate flight high speed aircraft.

**W1 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | No |
| 2 | Yes |
| 3 | No |
| 4 | Yes |
| 5 | Yes |

**W2 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | Yes |
| 3 | No |
| 4 | Yes |
| 5 | Yes |

3. **Query:** what problems of heat conduction in composite slabs have been solved so far.

**Indexed Query** : problem heat conduction composite slab solve so far

**W1 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | Yes |
| 3 | Yes |
| 4 | No |
| 5 | Yes |

**W2 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | No |
| 3 | Yes |
| 4 | Yes |
| 5 | No |

4. **Query:** can a criterion be developed to show empirically the validity of flow solutions for chemically reacting gas mixtures based on the simplifying assumption of instantaneous local chemical equilibrium

**Indexed Query**: can criterion develop to show empirically validity flow solution chemically react gas mixture base simplify assumption instantaneous local chemical equilibrium

**W1 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | Yes |
| 3 | No |
| 4 | Yes |
| 5 | No |

**W2 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | Yes |
| 3 | Yes |
| 4 | No |
| 5 | No |

5. Query: what chemical kinetic system is applicable to hypersonic aerodynamic problems

Indexed Query: chemical kinetic system applicable to hypersonic aerodynamic problem

**W1 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | No |
| 3 | No |
| 4 | Yes |
| 5 | No |

**W2 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | No |
| 3 | Yes |
| 4 | Yes |
| 5 | Yes |

6. Query: what theoretical and experimental guides do we have as to turbulent couette flow behavior.

Indexed Query: theoretical experimental guide do we to turbulent couette flow behavior.

**W1 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | No |
| 2 | No |
| 3 | Yes |
| 4 | Yes |
| 5 | No |

**W2 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | Yes |
| 3 | No |
| 4 | Yes |
| 5 | Yes |

7. Query: is it possible to relate the available pressure distributions for an ogive forebody at zero angle of attack to the lower surface pressures of an equivalent ogive forebody at angle of attack

Indexed Query: possible to relate available pressure distribution ogive forebody at zero angle attack lower surface equivalent

**W1 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | Yes |
| 3 | No |
| 4 | Yes |
| 5 | Yes |

**W2 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | No |
| 3 | No |
| 4 | Yes |
| 5 | No |

8. Query: what methods -dash exact or approximate -dash are presently available for predicting body pressures at angle of attack

Indexed Query: method dash exact approximate presently available predict body pressure at angle attack

**W1 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | No |
| 2 | Yes |
| 3 | Yes |
| 4 | No |
| 5 | No |

**W2 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | No |
| 2 | No |
| 3 | Yes |
| 4 | Yes |
| 5 | No |

9. Query: papers on internal /slip flow/ heat transfer studies

Indexed Query: papers internal slip flow heat transfer study

**W1 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | No |
| 3 | No |
| 4 | Yes |
| 5 | Yes |

**W2 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | Yes |
| 3 | No |
| 4 | Yes |
| 5 | No |

10. Query: are real-gas transport properties for air available over a wide range of enthalpies and densities

Indexed Query: real gas transport property air available over wide range enthalpy density

**W1 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | No |
| 3 | Yes |
| 4 | Yes |
| 5 | No |

**W2 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | No |
| 2 | Yes |
| 3 | Yes |
| 4 | Yes |
| 5 | No |

11. Query: is it possible to find an analytical, similar solution of the strong blast wave problem in the newtonian approximation

Indexed Query: possible find analytical similar solution strong blast wave problem newtonian approximation

**W1 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | Yes |
| 3 | No |
| 4 | Yes |
| 5 | No |

**W2 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | Yes |
| 3 | Yes |
| 4 | Yes |
| 5 | No |

12. Query: how can the aerodynamic performance of channel flow ground effect machines be calculated

Indexed Query: can aerodynamic performance channel flow ground effect machine calculate

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | No |
| 3 | No |
| 4 | Yes |
| 5 | No |

**W2 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | No |
| 3 | Yes |
| 4 | Yes |
| 5 | No |

13. Query: what is the basic mechanism of the transonic aileron buzz

Indexed Query: basic mechanism transonic aileron buzz

**W1 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | No |
| 2 | Yes |
| 3 | Yes |
| 4 | No |
| 5 | Yes |

**W2 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | Yes |
| 3 | No |
| 4 | Yes |
| 5 | No |

14. Query: papers on shock-sound wave interaction

Indexed Query: papers shock sound wave interaction

**W1 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | No |
| 2 | Yes |
| 3 | Yes |
| 4 | Yes |
| 5 | No |

**W2 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | No |
| 3 | Yes |
| 4 | Yes |
| 5 | No |

15. Query: material properties of photoelastic materials

Indexed Query: material property photoelastic

**W1 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | No |
| 3 | No |
| 4 | Yes |
| 5 | No |

**W2 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | Yes |
| 3 | No |
| 4 | Yes |
| 5 | Yes |

16. Query: can the transverse potential flow about a body of revolution be calculated efficiently by an electronic computer

Indexed Query: can transverse potential flow about body revolution calculate efficiently electronic computer

**W1 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | Yes |
| 3 | No |
| 4 | Yes |
| 5 | Yes |

**W2 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | No |
| 3 | No |
| 4 | Yes |
| 5 | No |

17. Query: can the three-dimensional problem of a transverse potential flow about a body of revolution be reduced to a two-dimensional problem

Indexed Query: can three dimensional problem transverse potential flow about body revolution reduce to two

**W1 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | No |
| 3 | No |
| 4 | No |
| 5 | Yes |

**W2 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | Yes |
| 3 | No |
| 4 | Yes |
| 5 | No |

18. Query: are experimental pressure distributions on bodies of revolution at angle of attack available

Indexed Query: experimental pressure distribution body revolution at angle attack available

**W1 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | Yes |
| 3 | Yes |
| 4 | No |
| 5 | No |

**W2 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | Yes |
| 3 | Yes |
| 4 | No |
| 5 | No |

19. does there exist a good basic treatment of the dynamics of re-entry combining consideration of realistic effects with relative simplicity of results

**W1 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | No |
| 2 | Yes |
| 3 | Yes |
| 4 | Yes |
| 5 | No |

**W2 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | Yes |
| 3 | Yes |
| 4 | Yes |
| 5 | No |

20. has anyone formally determined the influence of joule heating, producedby the induced current, in magnetohydrodynamic free convection flows under general conditions

**W1 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | No |
| 3 | Yes |
| 4 | No |
| 5 | No |

**W2 Function:**

|  |  |
| --- | --- |
| Rank | Relevant? |
| 1 | Yes |
| 2 | Yes |
| 3 | Yes |
| 4 | No |
| 5 | No |

4. Describe why the top-ranked non-relevant document for each query did not get a lower score.

The top-ranked non-relevant document did not get a lower score because a single most-frequently occurring word in a document has a higher W1 value and that value being added to the Scores[N] gives a relatively higher value than a relevant document having one word in the term.

5. Briefly discuss the different effects you notice with the two weighting schemes, either on a query-by-query basis or overall, whichever is most illuminating. For example, you can point out that the weighting scheme seems to be working for this query as well as a list of other queries, but not for some other queries you have noticed. Try to explain why it works and why it does not work

W1: Does not take into account the document length and the average document length but takes into account the maximum term frequency.

W2: Takes into account the length of the document. Does not take into account the maximum term frequency.

W1 works better for few cases:

When a most important word is repeated many times in the document. The weight of the term is improved and hence a more accurate relevant model can be built.

W2:

When all documents are of equal length and average and doclen tends to become 1. Then this relevance model fails. Else, this relevance model takes into account the size of the document which plays the most important role in the relevance model.

6. Describe the design decisions you made in building your ranking system. (5 points)

Design decisions:

Created an Index of the queries.

The index of the documents are fetched.

Cosine similarity of every query is computed and normalized using weight functions provided and Scores[] is sent to find the Top5 relevant terms in the document.

When computing the relevance the vector representations of the documents are saved.