COM3110 Information Retrieval Assignment Report

Kiran Da Costa | November 2023

# Implementation

I’ve chosen to break up query processing into four stages: retrieval of relevant documents from the index, vectorisation of the query, vectorisation of the documents, and the ranking of results using cosine similarity. The first three of these have been turned into their own functions both to improve code readability and break the program up into logical, sequential steps.

The vectorisation functions are where different weighting metrics are accounted for – this keeps the main for\_query() function clean and readable, as well as

# Results

The overall results of my implementation are presented below. The best and worst results are italicised:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | **Stop-Listing** | | **No Stop-Listing** | |
| **Binary** | **Stemming** | Precision: | 0.16 | Precision: | 0.12 |
| Recall: | 0.13 | Recall: | 0.10 |
| F-Measure: | 0.14 | F-Measure: | 0.11 |
| **No Stemming** | Precision: | 0.15 | Precision: | 0.12 |
| Recall: | 0.12 | Recall: | 0.09 |
| F-Measure: | 0.13 | F-Measure: | 0.10 |
| **TF** | **Stemming** | Precision: | 0.16 | Precision: | 0.13 |
| Recall: | 0.13 | Recall: | 0.10 |
| F-Measure: | 0.14 | F-Measure: | 0.11 |
| **No Stemming** | Precision: | 0.14 | *Precision:* | *0.10* |
| Recall: | 0.11 | *Recall:* | *0.08* |
| F-Measure: | 0.13 | *F-Measure:* | *0.09* |
| **TFIDF** | **Stemming** | *Precision:* | *0.18* | Precision: | 0.16 |
| *Recall:* | *0.15* | Recall: | 0.13 |
| *F-Measure:* | *0.16* | F-Measure: | 0.14 |
| **No Stemming** | Precision: | 0.17 | Precision: | 0.14 |
| Recall: | 0.14 | Recall: | 0.11 |
| F-Measure: | 0.15 | F-Measure: | 0.12 |

## Binary

The binary term weighting metric is *generally* the worst-performing of the three. This is unsurprising given that it is simply a metric of whether the terms are in a document or not. However, I do not know if it is supposed to be so similar to the TF result set, and perhaps something is off in my method of calculation.

## TF

Using a term frequency weighting scheme, the results are virtually identical to the ones gathered using a binary weighting scheme. With no stemming, TF seems to perform worse than binary without stemming and even more surprisingly, the worst overall result in across all schemes was achieved using a TF weighting scheme with no stop-listing or stemming.

## TFIDF

Lorem

# Conclusion