

Text Technologies for Data Science Assignment 1

Mark Nemec

1 Introduction

The purpose of this report is to outline decisions made in the implementations of search algorithms for information retrieval and discuss the performance on provided datasets and techniques used.

2 Overlap algorithm

The average precision of this search algorithm on `qrys.txt` and `docs.txt` was 0.1527.

3 tf.idf algorithm

The average precision of this search algorithm on `qrys.txt` and `docs.txt` was 0.3230.