SEUPD@CLEF Task 1: Information retrieval for English and French documents: Team JIHUMING

Jesús Moncada Ramírez[1], Isil Atabek[1], Huimin Chen[1], Michele Canale[1], Nicolò Santini[1] and Giovanni Zago[1]

Abstract

Our group will propose an original and efficient information retrieval system for Longitudinal Evaluation of Model Performance (LongEval) by CLEF2023[1]. Focus is on short term and long term temporal persistence of the systems' performance, for both English and French documents. The aim is to find a model giving good results for longitudinal evolving benchmarks, for the subject Search Engines, University of Padova.

Keywords

CLEF 2023, Information retrieval, LongEval, English, French, Search Engines

1. Introduction

This report aims at providing a brief explanation of the Information Retrieval system built as a team project during the Search Engine course 22/23 of the master's degree in Computer Engineering and Data Science at University of Padua, Italy. Task chosen by the group is CLEF LongEval: Longitudinal Evaluation of Model Performance.

The paper is organized as follows: Section 2 describes our approach; Section 3 explains our experimental setup; Section 4 discusses our main findings; finally, Section 5 draws some conclusions and outlooks for future work.

2. Methodology

Describe the methodology you have adopted, the architecture of your system, your workflow, etc.

"Search Engines", course at the master degree in "Computer Engineering", Department of Information Engineering, and at the master degree in "Data Science", Department of Mathematics "Tullio Levi-Civita", University of Padua, Italy. Academic Year 2022/2023

jesus.moncadaramirez@studenti.unipd.it (J. M. Ramírez); isil.atabek@studenti.unipd.it (I. Atabek); huimin.chen@studenti.unipd.it (H. Chen); michele.canale.1@studenti.unipd.it (M. Canale); nicolo.santini.1@studenti.unipd.it (N. Santini); giovanni.zago.3@studenti.unipd.it (G. Zago)

CEUR Workshop Proceedings (CEUR-WS.org)

3. Experimental Setup

Describe the experimental setup, i.e.

- used collections
- evaluation measures
- url to git repository and its organization
- hardware used for experiments
- ...

4. Results and Discussion

Provide a summary of the performance on the previous year dataset.
Discuss the results and any relevant issues.

5. Conclusions and Future Work

Provide a summary of what are the main achievements and findings.
Discuss future work, e.g. what you may try next and/or how your approach could be further developed.

References

[1] CLEF2023, LongEval CLEF 2023 Lab, https://clef-longeval.github.io/, 2023.