

Project: Key Word and Key Sentence Extraction

Purpose

This project constitutes the UPG1 part of the course TNA010 Matrix Methods for AI. The purpose of this work is to consider an application problem, implement a computer program that solves the problem in question using mathematical theory from the course literature Eldén [1]. The project is conducted in groups with two to three members.

Application problem and implementation

Consider the automatic key word and key sentence extraction problem in Eldén [1, § 14]. Implement extraction methods based on both the Saliency Score and the rank- k matrix approximation.

Data set is available in the Projects folder at the Lisam course webpage.

Written report

Write a report with the following content:

1. Introduce the project with a clearly defined problem.
2. Give a detailed description of the data set used for the problem.
3. Present the theoretical foundation and mathematical models that are used to address the problem.
4. Describe and discuss computational and algorithmic aspects of your implementation. Present your algorithm with pseudocode in the style of e.g. the algorithm in [1, §7.2.6, p. 90].
5. Describe conducted experiments and present test results.
6. Discussion and conclusion.
7. References and literature sources.

The report should be written as a technical report, with the usual formality requirements and, e.g. the ability to reproduce the work. The length of the report should not be too long, about five pages should suffice. Consider using \LaTeX .

Presentation

The project will be presented formally at the last scheduled seminar. A practice presentation will be conducted a week before the formal presentation. 45 minutes are allocated for the presentation

with 30–35 minutes of prepared content and 10–15 minutes for discussion and questions from the audience. The prepared content should have an emphasis on mathematical and computational aspects but also on the experimental results.

Each group from the audience must read the reports of the other groups and prepare 2–4 questions for discussion.

Referenser

- [1] L. Eldén. *Matrix Methods in Data Mining and Pattern Recognition*. SIAM, 2 edition, 2019.