RACKATHON

Education Technology

- Zerone Girls

Problem Statement

1. Taking notes during online learning/lectures, helps students to stay engaged and also helps students to revise concepts

2. Each student's notes may vary i.e. one student's notes may contain points that are not present in the notes taken down by his friend.

3. There is always a chance that one student among a few students may have picked up something vital that the others have missed or might have come up with a good way of simplifying a complex topic.

4. By sharing notes, one can reprocess the information from others' notes and gain a better understanding of the given concept

Solution

Multi-Document Summarization

We plan on summarizing lecture notes of a group of students by extracting information from multiple texts/documents (student's notes) written about the same topic and generate a summary with the information from all the input documents. This way, students can have access to a well-formulated notes/copy that they can refer to before their exams and get a quick summary of the concepts taught during the classroom lectures.

By combining notes of different students, we can create a comprehensive summary with different explanations put together and outlined, and with every topic described from multiple perspectives - all in a single document !

Implementation

Document Processing

The documents are concatenated into a single document and minimal text processing is applied.The processing steps can be conducted using SpaCy NLP modules

Graph Clustering

Apply graph clustering techniques such as spectral clustering on the sentence graph to get partitions within the graph.

Sentence Graph Construction

A sentence graph is constructed where the nodes correspond to the sentences generated at the document processing step and edges are drawn based on lexical and deep semantic relations between sentences.

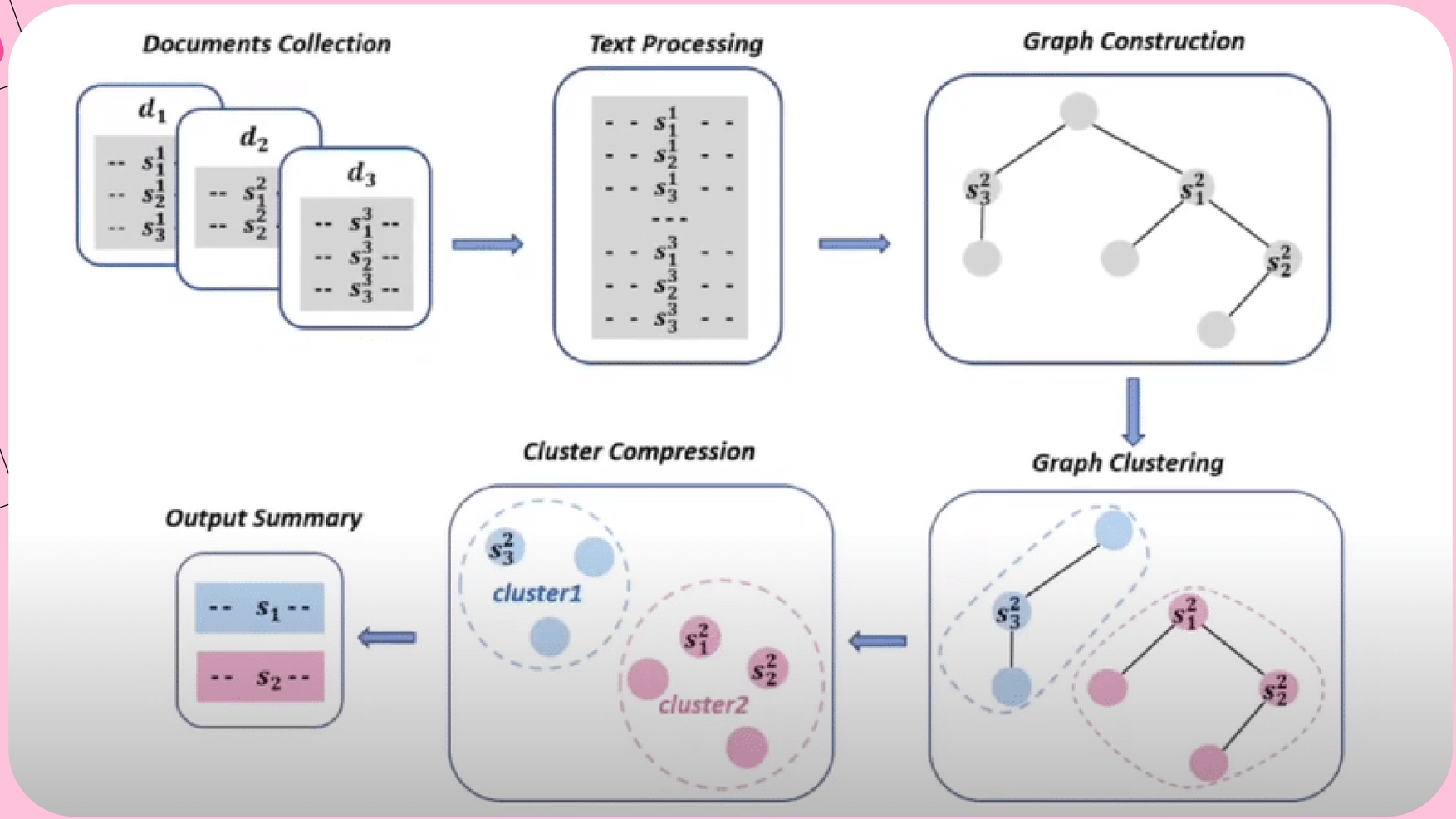
Generate summary texts

Use Multi-sentence compression to genrate a single summary text from the extracted sub-graphs

Pipeline for Unsupervised Multi-Document Summarization

Here, the input is multiple text documents/lecture notes

Output is the multi document summary

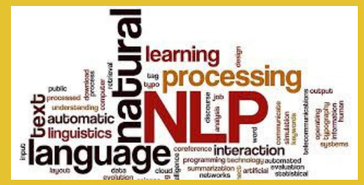


React JS

Framework

mongoDB

Database



NLP module

sklearn python library for Spectral Clustering

Technology Stack

HTML CSS JS

Front-end

Express JS

Backend

Client/server and routing

netlify

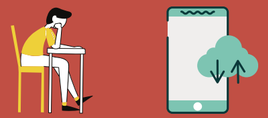
Hosting Application

WORK FLOW OF THE APPLICATION

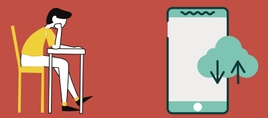
UPLOAD

Upload different text lecture notes via application

Student 1



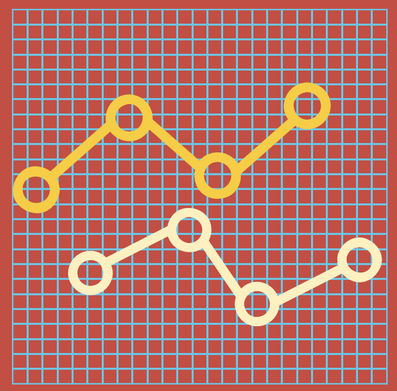
Student 2



Student 3

PROCESSING AND SUMMARIZATION

Apply Multi-Document summarization technique using sentence graph compression



OUTPUT SUMMARY

Output summarized document is sent to all the 3 students

Impact

- Helps in generating a single document/copy containing concepts explained in multiple perspectives, hence enhancing the understanding of concepts
- Helps in effective preparation during examination.
- Presents information that is organized around the key aspects to represent diverse views and produces an overview of a given topic.
- Helps in getting a quick recap of the lectures in less time.
- Combines the main themes in the documents with completeness, readability and concision.

Difficulties/Technological Challenges

1. Multi-Document Summarizer is a complex task
2. Text within sections should be divided into meaningful paragraphs
3. Good readability
4. The summary should have a clear structure, including an outline of the main content
5. There should be no semantic redundancy or information noise