Introduction Assignment 2

Niels Scholten | JM2050-M-6 Natural Language Processing

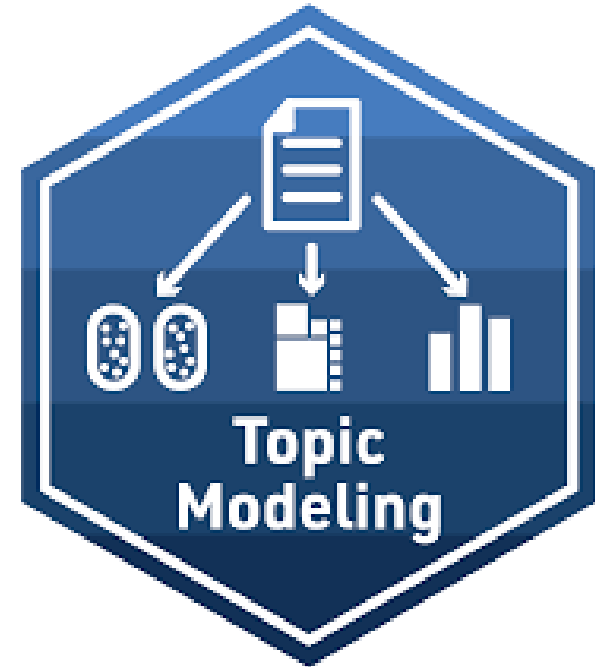JADS is powered by | TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY | TILBURG UNIVERSITY

JADS Jheronimus Academy of Data Science

Assignment 2

- Topic Modeling

- Three Topic Models
  - Latent Dirichlet Allocation
  - FLSA-W
  - BERTopic

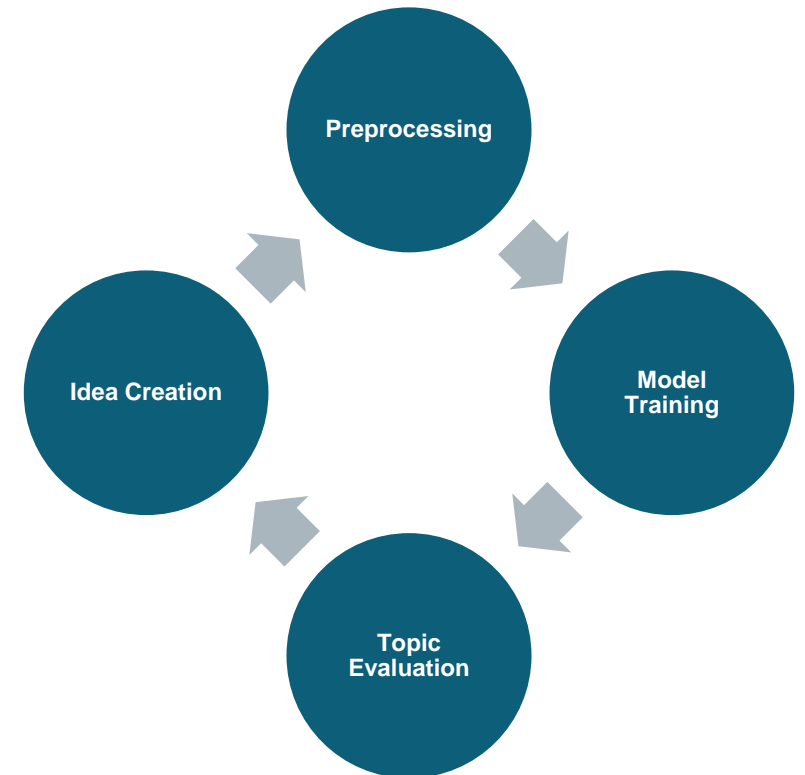- Groups of 3 students

- Deadline 07-11-2024

Nvidia Stock Market News Dataset

- The same dataset as assignment 1

- Some preprocessing steps might be different for topic modeling

Go Through Multiple Iterations!

- Don't settle on your first iteration of your model.

- For each iteration, mention:
    - The changes made.
    - The motivation of the changes.
    - The effect of the changes.

For your Report

- The focus should be technical.
  - You should not spend a lot of time describing insights found from the topic modeling.
  - You should spend a lot of time on the quality of your topics, the content of the topics and what influences the quality.

Timeline

- 10-10-2024 Introduction Assignment 2

- 17-10-2024 Tutorial on Topic Modeling

- 07-11-2024 Deadline Assignment 2

- Good luck!