Computational Social Science

Mid-Term Presentation

Sameer Singh Rawat
(sameer.rawat@campus.lmu.de)

Matrikel Nr.- 12691600

Newyork Times Bestseller Analysis

1. Data Collection - Collecting the data of newyork times bestseller books through NYT api, and their respective summary from google books api.

2. Sentiment Analysis - To analyze the sentiment of each of the NYT bestseller's summary, and see the stories of which sentiment are prefered by majority of audience in general. BERT or hugging face models can be used for the purpose of sentiment analysis.

3. Topic Modelling - To analyze around which topics most of the NYT bestseller stories revolve. It can give us an idea about the type of books people are currently liking to read. BERTopic or Latent Dirichlet Allocation(LDA) can be used for this purpose.

4. Recommender System - In the end we would make an interactive interface for the user to input the book they have already read and give the most similar book from the books that has featured in Newyork Times Bestsellers.We will make use of cosine similarity to get the most similar NYT bestseller book to the inputed book.

5. Finally deploy our models online for public use.