Final Project Topic

Analyzing UN Speeches' Transcripts (1971 – 2018)

Ritesh Kumar

2024SP_MS_DSP_453-DL_SEC61: Natural Language Processing

Module 4

Project P.1

Nethra Sambamoorthi and Sudha BG

May 30, 2024

Project: The objective of this project is to analyze the evolution of topics in United Nations General Assembly (UNGA) speeches from 1971 to 2018. By applying Latent Dirichlet Allocation (LDA), a widely-used topic modeling technique, we aim to uncover the underlying thematic structure within the speeches. This analysis will reveal the dominant topics discussed over nearly five decades and track their prominence over time.

Methodology: Latent Dirichlet Allocation (LDA) treats each document as a mixture of topics and each topic as a mixture of words. By applying LDA to the UNGA speeches, we can identify the hidden thematic structure within the dataset. This method allows us to uncover the dominant topics and understand how these topics have evolved over time.

Significance of the Project: Analyzing the evolution of topics in UNGA speeches using LDA offers a comprehensive view of how global priorities and geopolitical landscapes have changed over time. This approach not only highlights the emergence and decline of specific issues but also provides a deeper understanding of the international community's collective concerns and responses to global challenges.

Expected Outcomes:

The expected outcomes of this project include:

- Identification of dominant topics in UNGA speeches from 1971 to 2018.
- Visualization of the evolution and prominence of these topics over time.
- Insights into significant trends and shifts in international discourse.

Research Papers for Reference:

1. Blei, David & Ng, Andrew & Jordan, Michael. (2001). Latent Dirichlet Allocation. The Journal of Machine Learning Research. 3. 601-608. (Link)

2. Rahul Kumar Gupta, Ritu Agarwalla, Bukya Hemanth Naik, Joythish Reddy Evuri, Apil Thapa, Thoudam Doren Singh, Prediction of research trends using LDA based topic modeling. (Link)
3. A. Goyal and I. Kashyap, "Latent Dirichlet Allocation - An approach for topic discovery," (Link)