

Identifying Sub-Topics and Generating a Fair Summary of Migration-Related News Articles

Thousands of news are produced and consumed each day. These news are scattered across diverse set of topics such as business, economy, international relation, Migration etc. Migration is an issue currently faced by many countries across the globe. Different news channels cover various news about migration related activities. A close inspection reveals that these migration-related news is composed of different sub-topics such as human trafficking, immigration, organized crime etc. Identifying these sub-topics and summarizing them is an essential first step to get a coherent real-time view of the situation. The objective of this project is two fold:

1. Identifying migration-related sub-topics and classifying news sources into these categories.
2. Generating a report/summary of the scenario by combining information from multiple sources. The objective is to represent each sub-topic in a fair way.

An ideal candidate should have

1. Strong background in python and deep learning
2. motivation behind learning and exploring the data
3. interest in some human level annotation
4. knowledge about basic NLP concepts

Interested students are encouraged to email to Dr. Koustav Rudra at [rudra\(at\)l3s\(dot\)de](mailto:rudra@l3s.de) for scheduling a meeting.

References:

1. Summarizing User-generated Textual Content: Motivation and Methods for Fairness in Algorithmic Summaries [<https://arxiv.org/abs/1810.09147>]
2. Fair and Diverse DPP-based Data Summarization [<https://arxiv.org/abs/1802.04023>]