Determining country-specific engagement in climate change policies[1]

Alexandra Butoi, Mian Zhong
Big Data for Public Policy, ETHZ, Spring Semester 2021

Motivation

The Paris Agreement is an international treaty which aims to limit the effects of climate change. It works in 5-year cycles, before which the participating countries submit the climate policies that they are going to implement in the following cycle. These policies are submitted in a Nationally Determined Contributions (NDC) document and are publicly available on the United Nations website. Given the flexible nature of these documents, as they contain both binding and voluntary policies, it is expected that the countries adjust their NDCs according to their specific circumstances and needs. Therefore, this leaves plenty of room for analysis of the content of these documents, in particular the countries' main topics of interest, which might differ depending on the country's wealth, development or pollution level.

In this project we performed topic modeling on the NDC documents submitted under the Paris Agreement for the 2020 deadline, and developed a web application which shows our main results using the LDA model. The web application displays the topic intensity by country for each of the main topics found in the NDCs and can be easily used to make comparisons between countries' level of engagement in certain topics (e.g. natural disasters, finance, carbon emissions, natural resources, etc.). Additionally, the web application contains a visualization tool for the topics, showing the most salient words for each of them, as well as topic similarity.

Challenges

One of the main challenges that this project posed was data acquisition i.e., collecting the PDF documents and scraping them in order to convert them to a suitable format for topic modeling. As there is no standard format for the documents (each country can choose freely what the document looks like and what it contains), they required careful scraping and further post-processing before feeding the data to the topic model. Another challenge was posed by the limited number of documents. In order to create more data, we split each NDC into paragraphs and generated a topic distribution for each of them, then averaged them to obtain an aggregated topic distribution by country.

Future Work

---

[1] https://github.com/butoialexandra/Determining-country-specific-engagement-in-climate-change-policies

Due to time constraints, we computed a topic distribution for each country by averaging the paragraph-level topic distributions. However, it might be worth experimenting with other methods (e.g. using only the dominant topic in each paragraph to determine which topics are present in a certain NDC).

Another potential area for research would be gathering some country-specific statistics (such as the World Bank's development indicators[2]) and using them for predicting the topics found in the NDCs in a multi-label classification fashion. Then the feature importances can be computed and used to determine which country-specific indicators best explain the country's main climate policy interests.

[2] https://databank.worldbank.org/source/world-development-indicators#