Discussion Quality Estimation from Text

Abstract

Online discussions suffer from a lot of bad practices, like off-topic discussions, polarization and disrespectful language. Measuring the quality of the discussion can help in detecting and fixing these problems. In this project you will implement a theoretical framework for discussion quality using state-of-the-art Natural Language Processing (NLP) methods. The framework contains ratings on various separate aspects, like respect classification, argument mining and sentiment analysis. By representing each of these aspects as a unique NLP task, we can use specialized tools for extracting the rating.

Introduction

Discussions between individuals is central in the process of coming to important decisions, like how to challenge climate change or which lockdown measures to reduce. When the goal of such a discussion is to find arguments from a variety of perspectives, we talk about a **deliberation**. In a deliberation, it is important that participants adhere to principles like respect and active listening. Measuring how much they exhibit these principles opens up possibilities to detect and resolve conflict.

We will focus on measuring quality in online discussions through text. Earlier, scientists came up with the Discourse Quality Index (DQI, Steenbergen et al., 2003) as a way to measure the most important principles underlying deliberations. Recently, an attempt has been made to implement the quality index in a single program (Fournier-Tombs and Di Marzo Serugendo, 2019). The code is available online[1]. However, instead of focusing on a manually coded dataset, our aim is to use Natural Language Processing (NLP) to estimate scores for each of the categories in the DQI automatically.

In this project, there is a possibility to focus on one of the five categories of DQI: **respect**. In the DQI framework, respect is measured by three sub-indicators. Each of the three sub-indicators pose an interesting challenge for Text Mining systems. It is up to you to implement, evaluate and analyze the implementation of (neural) methods for estimating their values. Below, a description is given for each of the three indicators.

- **Statements about groups**: the amount of negative and positive statements about groups of persons.

- **Statements about demands**: the amount of negative and positive statements about demands made by participants.

- **Counterarguments** : the degree to which counterarguments are taken into account.

Some example key words and techniques you will be working with: **Sentiment analysis**, **Named Entity Recognition**, **WordNet**. The projects assumes you are familiar with Python programming, basic NLP techniques and a solid foundation of machine learning. Additional datasets will be used to train and evaluate the measurement of respect in isolation.

Contact

For more information, reach out to:

- Michiel van der Meer `m.t.vandermeer-1@tudelft.nl` (LIACS, Leiden University & Interactive Intelligence, TU Delft, Hybrid Intelligence)

- dr. Pradeep Murukannaiah `p.k.murukannaiah@tudelft.nl` (Interactive Intelligence, TU Delft, Hybrid Intelligence)

References

Marco R Steenbergen, André Bächtiger, Markus Spörndli, and Jürg Steiner. Measuring political deliberation: A discourse quality index. *Comparative European Politics*, 1(1):21–48, 2003. doi: 10.1057/palgrave.cep.6110002.

Eleonore Fournier-Tombs and Giovanna Di Marzo Serugendo. Delibanalysis: Understanding the quality of online political discourse with machine learning. *Journal of Information Science*, page 0165551519871828, 2019. doi: 10.1177/0165551519871828.

---

[1] `https://github.com/eleonoreft/DelibAnalysis`