Ideas for master projects

Malvina Nissim
`m.nissim@rug.nl`
room 1311.421

IK meeting, 22 November 2019

Project 1: Authorship Attribution with Twin Data

Twin AA

- collect twin and non-twin writings (we can ask the Twin Registry in Amsterdam)
- use and extend existing AA model (GLAD)
- test whether distinguishing twin writing is harder than non-twin writings
- analyse which aspects (features) are relevant

Project 2: Bias, Stereotypes, Defaults

Bias, Stereotypes, Defaults

- explore methods to observe bias and stereotypes avoiding the problem of confirmation bias
- practical tool: (mainly) language models
- theoretical tool: categorisation theory
- requirements: creativity, linguistic interest

Project 3: Style-aware generation

Style-aware generation

- see this as a machine translation problem
- leverage existing approaches
- evaluation strategy
- pick some interesting "style"
- work on Dutch

Project 4: Gender Prediction with self-reported data

Gender Prediction with self-reported data

- collect data with self-reported gender
- use existing data with gender labels
- leverage existing gender prediction models
- English, Dutch, choose your language
- use prediction models to study gender across datasets

Project 5: Hyperpartisan project: extensions