GENDER BIAS BETWEEN LANGUAGES
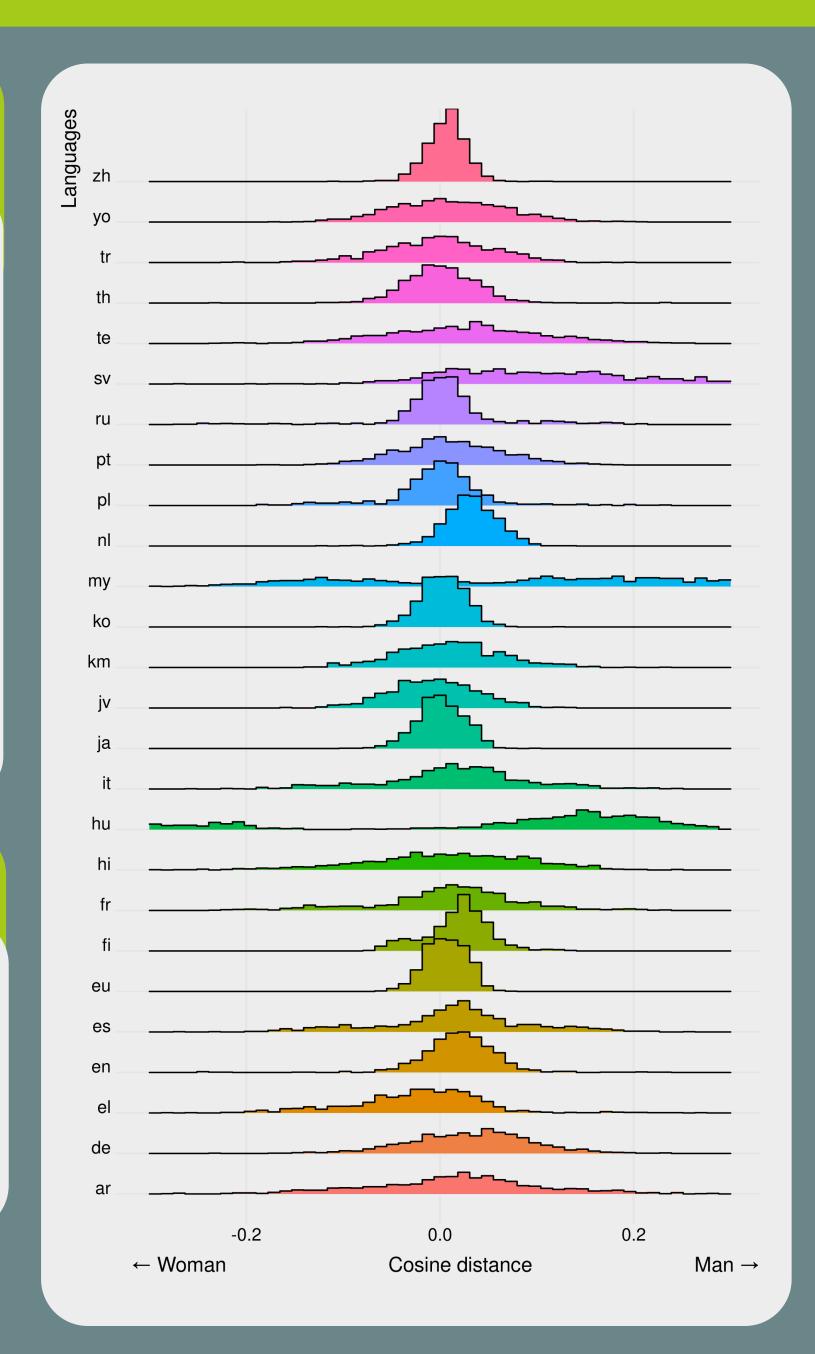
THIJS RAYMAKERS - CSE3000 - 2020-05-20

RESEARCH QUESTION

Do word embeddings of different languages differ in gender bias?

Word embedding are an important tool in natural language processing, but should be used with caution because they can contain human biases. This research will look at the gender differences in a word embedding between the words *man* and *woman* in different languages to assess possible gender bias.

METHODS

- 26 languages of 16 language families
- Cosine distance
- Measure relation between attribute words (*man* and *woman*) and target words (1000 most used words)
- Statistics used to compare languages

PRELIMINARY RESULTS

- Languages do differ in gender bias, i.e. some languages have more bias than others
- Not yet known how much bias there is or which languages are more biased than others

PLAN

- Find a statistic that can help to determine which languages are more biased than others
- Difficult because languages have different distributions and different variances

CONTACT INFO

Thijs Raymakers - t.raijmakers@student.tudelft.nl
Professors: Marco Loog, David Tax
Supervised by: Stavros Makrodimitris, Arman Naseri Jahfari, Tom Viering

TU Delft