OpenNRE Evaluation Notes

Dataset: Wikigender

Model: pcnn (piecewise CNN) + attention model

4ish Inputs

* Name anonymization: “NA” / “NoNA”
* Equalized Gender Mentions: “Eq” / “NoEq”
* Gender Swap Dataset: “GS” / “NoGS”
  + Name Swap: “NS” / “NoNS”
* Debiased Embeddings: “DE” / NoDE”
* in theory, with all 5 inputs, we should have 25 = 32 input combinations
* and then 32 \* 2 = 64 models b/c of male and female versions
  + but right now we have 28

Not using equalized gender mentions because the datapoints that get upsampled by the female gender just get removed

* Actually we may try downsampling

6 different input combinations (all NoEq, two of each [M/F])

1. NA\_NoEq\_GS\_NoNS\_DE
   1. Name anonymized, gender-swapped, no name-swap, debiased embeddings
2. NA\_NoEq\_GS\_NoNS\_NoDE
   1. Name anonymized, gender-swapped, no name-swap, no debiased embeddings
3. NA\_NoEq\_NoGS\_DE
   1. Name anonymized, no gender-swap, name-swapped (presumably), debiased embeddings
4. NA\_NoEq\_NoGS\_NoDE
   1. Name anonymized, no gender-swap, name-swapped (presumably), no debiased embeddings
5. NoNA\_NoEq\_GS\_NoNS\_DE
   1. No name anonymization, gender-swap, no name-swap, debiased embeddings
6. NoNA\_NoEq\_GS\_NoNS\_NoDE
   1. No name anonymization, gender-swap, no name-swap, debiased embeddings

Relative F1\_Score

Relation Observations

* birthdate: mixed results, some men, 1 woman
* birthplace: mixed results, very slightly toward man
* spouse is largely favored toward men (close to 10% more)
* hypernym: negligible

Inputs

* Name anonymization: mostly male bias
* Gender swap: seems to do a good job of mitigating bias except for spouse relation
* Debiased embeddings: doesn’t seem to have significant impact

Interesting combinations

* NA\_NoEq\_NoGS\_NoDE
  + much better for female birthdate relation
  + 2nd highest male bias for spouse
* NA\_NoEq\_GS\_NoNS\_NoDE
  + same as above, but with gender-swapping
  + female birthdate relation bias goes down
  + male bias for spouse goes up

Relative Precision

Relation Observations

* Male bias almost across the board
* Spouse > birthdate >> hypernym > birthplace

Inputs:

* Name anonymization: seems to have less difference across all relations when you do not name anonymize
* Gender swap: doesn’t seem to be the pivotal factor
  + have the difference results on both ends of the spectrum, no gender swap has difference results that is the average
* Debiased embeddings: does not seem to be significant

Interesting combinations

* NA\_NoEq\_GS\_NoNS\_DE, NA\_NoEq\_GS\_NoNS\_NoDE, NA\_NoEq\_NoGS\_DE
  + these 3 exhibit most difference
  + What’s in common? NA\_NoEq
  + ⅔ of these are gender swapped, but ⅔ of the 3 lowest are also gender-swapped

Relative Recall

Relation Observations

* Male bias for spouse relation and birthplace relation
* Slightly weaker female bias for birthdate and hypernym relation

Inputs:

* name anonymization: seems to correlate with stronger male spouse bias, but not much otherwise
* gender-swap: ⅔ cases of male birthdate bias
  + ⅔ cases for female birthdate bias, but not significantly
* debiased embeddings: seems to reduce bias across the board

Interesting combinations

* NA\_NoEq\_NoGS\_NoDE
  + female bias for birthdate and hypernym
  + male bias for spouse