

The Social Impact of Natural Language Processing

Vivek Gupta
NLP Seminar Spring 2019

Outline

1. Does NLP need an ethics discussion
2. The Social Impact of NLP research
 - a. Exclusion: Side effect of Data (e.g. Demographic Bias)
 - b. Overgeneralization: Side effect of Model
 - c. Overexposure and Underexposure: Effect of research design
 - d. Dual-use problem: intensional moral issues

Does NLP need an ethics discussion?

Less discussion on ethics in NLP research as human subjects are not involved (Medicine)

Earlier sources of NLP data not strongly linked to any particular human source (Social Media)

Not many language based technologies, so less ethical implications

More and more widely used language based technologies as NLP research improve

Concern started with unknowingly revealing the sensitive attribute of source (privacy)

Concern started with unequal opportunity for individual or groups e.g. minorities (fairness)

Language, Individual and Society are mutually related and impact society (e.g. Vision)

Exclusion

1. Dataset carries a **demographic bias**, i.e., carry information about demographics
2. Arises due to i.i.d assumption, assume equal representation in training dataset
3. Lead to exclusion or demographic misrepresentation of individuals or group
4. E.g. Standard language technology work better for some groups (fairness issue)
5. **Solution:** Methods which address overfitting or the imbalanced data problem
6. However, many models unknowingly and implicitly amplify demographic bias in dataset (data side effect)

Overgeneralization

1. Modeling side effect due to automatic inference of user attributes for personalization
2. Deals with the question whether **false answer be any better than no answer**
3. **Solution:** Learning with rejection, introducing dummy class variables, modeling the regularization, cost sensitive learning , varying of confidence thresholds.

The Problem of Exposure (Over/Under)

1. Topic overexposure originate from research design rather data or model choices
2. **Overexposure:** Some research topics receive more mainstream attention than others
3. **Availability Heuristic:** things I am familiar with seem much more important than unfamiliar stuff. E.g. people overestimate size of their own cities over unknown cities
4. In Ethics, it becomes stereotypes a.k.a associating events with particular group
5. **Underexposure:** Problem of less research in low resource language compared to English (Tree bank for only 11 of 31 frequent languages, less typological variety)
6. E.g. In India, there are very few services/tools which built in local language (commercial aspects : bigger market)

Dual-use problems

1. Text classification can detect slang and hidden messages, but has potential to be used for censorship.
2. NLP techniques can be used to detect fake news, but also can generate them in the first place (similar to images)
3. Schools in favour/against military-related funding or AI use military purpose for moral reasons
4. Thus, NLP enable morally questionable practices and corrective measures for this reason

Thank You!