**Using ChatGPT to Generate Gendered Language**

**Datasets for Gendered Language generated using ChatGPT**

**Files description:**

Three datasets(set 1, set 2, set 3) generated using Gaucher lexicon are:

1. Gaucher\_dataset\_set\_1.csv
2. Gaucher\_dataset\_set\_2.csv
3. Gaucher\_dataset\_set\_3.csv

Three datasets(set 1, set 2, set 3) generated using Cryan lexicon are:

1. Cryan\_dataset\_set\_1.csv
2. Cryan\_dataset\_set\_2.csv
3. Cryan\_dataset\_set\_3.csv

Each CSV file has the following columns:

1. Index: Index of the sentence.
2. Sentences: List of sentences.
3. Labels: Labels assigned to the sentences. The labels are 'consistent' and 'contradictory'.
4. about\_Male: Has two values, either 0 or 1. 1 indicates that the sentence is about male, and 0 indicates it's about female.
5. Masc\_terms: Has two values, either 0 or 1. 1 indicates that the sentence is generated using male-coded words from the lexicon, and 0 indicates it's generated using female-coded words from the lexicon.

If you would like to use the dataset, please cite the paper mentioned below:

| @INPROCEEDINGS{10470830,  author={Soundararajan, Shweta and Jeyaraj, Manuela Nayantara and Delany, Sarah Jane},  booktitle={2023 31st Irish Conference on Artificial Intelligence and Cognitive Science (AICS)},  title={Using ChatGPT to Generate Gendered Language},  year={2023},  volume={},  number={},  pages={1-8},  keywords={Natural languages;Chatbots;Cognitive science;Artificial intelligence;natural language processing;machine learning;large language models;ChatGPT;gendered language;prompt engineering;zero-shot prompting},  doi={10.1109/AICS60730.2023.10470830}} |
| --- |