Using Artificial Intelligence and Algorithms for Natural Language Processing to Analyze the Emotions of Gender in Literature and Songs

Heiman, Alice (School: Viktor Rydbergs Gymnasium Odenplan)

Is it possible to measure the sentiments and emotions linked to words by analyzing their usage in different contexts? This experiment compares how feminine and masculine words are used in texts by utilizing computational techniques. Creating algorithms that can analyze the emotional associations with gender could give a quantitative tool to inspect how women and men have been perceived throughout history. The research questions covered were: (1) Is it possible to measure the sentiments and emotions associated with gender in literature and songs? (2) Are gender-specific terms used differently throughout history? (3) Can computers capture the emotional contents of words? A neural network was trained on pre-labeled data to classify sentences containing gendered words as positive or negative. Furthermore, an algorithm was created to translate words into numerical representations of their emotional usage, called sentiment vectors. Roughly 5,000 books published between 1830-1930 were then analyzed to examine trends over time. The results displayed that neural network classification could capture views on gender in certain songs. Sentiment vector analysis indicated a minimal difference in positivity between gendered words but found that masculine phrases were more often linked with trust. The developed techniques display the potential to highlight how natural language mirrors attitudes towards men and women and found that masculine and feminine words, in some cases, are associated with different emotions. Finally, sentiment vectors can be applied to examine any set of words, allowing applications in impartiality evaluation, measuring and quantifying societal change, and assisting literary research.

Awards Won:

Third Award of $1,000

NC State College of Engineering: Award to attend NC State Engineering Summer Camp