

COSC 247 Assignment 4.

Paper reviewed:

Thelwall, Mike. "Gender bias in sentiment analysis." *Online Information Review* (2018).

<https://www-proquest-com.silk.library.umass.edu/docview/2031409065/fulltextPDF/A74D873E505744DBPQ/1?accountid=14572>

Summary:

The purpose of the paper is to show that online sentiment analysis algorithms that use some form of machine learning have a gender bias. They attempt to prove this by designing an experiment in which reviews were extracted from TripAdvisor.com and segregating the users to different genders based on their name, then they train two different classifiers Naïve Bayes and Support vector Machines on the data, and using those models to predict test case reviews. The paper concluded that the experiments gave clear evidence of bias in machine learning based sentiment analysis algorithms, in particular the paper concluded that such programs tend to favor the opinion of women more than the opinion of men.

Nature of social impact:

Sentiment analysis algorithms are used for many purposes. They are used by politicians to determine government policy, stock market trading, and by companies to find the right target market for their products. The finding that these learning models are biased means that the end users of these algorithms are getting flawed results.

Clarifying Question:

The author was pretty vague about the design of support vector machine and other learning models used in this experiment. I think providing the code for the training models used and the actual data used for training would make the paper more clear. The author should also explore the possibility that the reason they are getting skewed results could be because of how his particular algorithm was designed and not because of how machine learning algorithms are designed in general.

Substantive critique:

The experiment designed here is sort of flawed in how they collect data. They tried to group people into different genders based on their online username and this can lead to the data not including people with a gender ambiguous username. Also, I think it is a bit of stretch to say that all or most sentiment analysis programs are "better able to detect the sentiments of females than of males" based on a single experiment, with data from a single source and only two algorithms tested.