

## Gender Bias in Natural Language Processing: Machines Reflect Misogyny in Society

**Authors :** Irene Yi

**Abstract :** Machine learning, natural language processing, and neural network models of language are becoming more and more prevalent in the fields of technology and linguistics today. Training data for machines are at best, large corpora of human literature and at worst, a reflection of the ugliness in society. Machines have been trained on millions of human books, only to find that in the course of human history, derogatory and sexist adjectives are used significantly more frequently when describing females in history and literature than when describing males. This is extremely problematic, both as training data, and as the outcome of natural language processing. As machines start to handle more responsibilities, it is crucial to ensure that they do not take with them historical sexist and misogynistic notions. This paper gathers data and algorithms from neural network models of language having to deal with syntax, semantics, sociolinguistics, and text classification. Results are significant in showing the existing intentional and unintentional misogynistic notions used to train machines, as well as in developing better technologies that take into account the semantics and syntax of text to be more mindful and reflect gender equality. Further, this paper deals with the idea of non-binary gender pronouns and how machines can process these pronouns correctly, given its semantic and syntactic context. This paper also delves into the implications of gendered grammar and its effect, cross-linguistically, on natural language processing. Languages such as French or Spanish not only have rigid gendered grammar rules, but also historically patriarchal societies. The progression of society comes hand in hand with not only its language, but how machines process those natural languages. These ideas are all extremely vital to the development of natural language models in technology, and they must be taken into account immediately.

**Keywords :** gendered grammar, misogynistic language, natural language processing, neural networks

**Conference Title :** ICNLPAG 2020 : International Conference on Natural Language Processing and Affix Grammars

**Conference Location :** Paris, France

**Conference Dates :** March 26-27, 2020