

Examining the Presence of Gender Bias in Customer Reviews Using Word Embedding

Arul Mishra, Himanshu Mishra, Shelly Rathee*

Abstract

Humans have entered the age of algorithms. Each minute, algorithms shape countless preferences from suggesting a product to a potential life partner. In the marketplace, algorithms are trained to learn consumer preferences from customer reviews because user-generated reviews are considered the voice of customers and a valuable source of information to firms. Insights, mined from reviews, play an indispensable role in several business activities ranging from product recommendations, targeted advertising, promotions, segmentation, etc. In this research, we question whether reviews might hold stereotypic gender bias that algorithms learn and propagate. Utilizing data from millions of observations and a word embedding approach, GloVe, we show that algorithms designed to learn from human language output, also learn gender bias. We also examine why such biases occur: whether the bias is caused because of a negative bias against females or a positive bias for males. We examine the impact of gender bias in reviews on choice and conclude with policy implications for female consumers, especially when they are unaware of the bias, and the ethical implications for the firms.

Keywords: Gender bias, natural language processing, customer reviews, text analysis, word embedding

*Authors are at David Eccles School of Business, Department of Marketing, University of Utah, Salt Lake City, Utah - 84112

Firms extensively use natural language processing algorithms to gather insights. Insights mined from reviews play an indispensable role in several business activities ranging from product recommendation, targeted advertising, promotions, segmentation, etc. For instance, algorithmic recommendations have been ubiquitous. They are used for guiding consumer decisions ranging from what brand to buy, places to visit, to even job offerings or a potential life partner. While there is plethora of research on the importance of reviews, there is no work that documents the presence of a bias in reviews, especially gender bias. In this research, we apply a recent word-embedding algorithm, GloVe (Global Vector), to more than 11 million reviews from Amazon and Yelp. Our findings indicate the presence of gender bias in reviews. Along with demonstrating the presence of gender bias, we examine whether gender bias exists because of a bias against women or a bias for men. Interventions can be successful only when the nature of the gender bias is known. Overall, we find that women are more associated with negative (e.g. fickle, impulsive, homebound) rather than positive attributes (e.g. loyal, sensible, industrious). However, men are more associated with positive rather than negative attributes. In order to address the question of whether gender bias in reviews have an influence on consumer decisions, we focus on a specific marketing domain, product recommendations, and demonstrate that consumer choice especially that of female consumers does get impacted. Firms routinely gather insights from customer reviews assuming them to be the voice-of-the-customer but if these reviews hold gender bias, then any decision made using those reviews has the power to harm customers. Biased input can result in biased recommendations, messages, or insensitive portrayals, and could result in learning and using consumer vulnerability against female consumers.

Insights mined from reviews play an indispensable role in several business activities ranging from product recommendation, targeted advertising, promotions, segmentation etc. For instance, algorithmic recommendations have become ubiquitous. They are used for guiding consumer decisions ranging from which brand to buy, books to read, clothes to wear, foods to eat, places to visit, to even job offerings or a potential life partner (Jannach et al. 2010; Schrage 2014). Their ubiquity comes from the fact that consumer preferences can be nudged and shaped by these recommendations (Hubl and Murray 2006). While some recommendations rely on prior purchase behavior, the bulk of recommendations are made for products that a consumer has never purchased before. Consider travel planning - for 63% of consumers who make travel plans after searching the Internet, booking a hotel follows a standard pattern online (Rheem 2012). They search for the destination based on their vacation dates and see numerous booking sites that Google shows them. The travel site has very little information about the consumers prior travel history or preferences except for some demographic information like gender and prior browsing history that Google shares with the travel site. In such situations, the recommendations by the travel site are invariably based on how similar a consumer might be to available consumer profiles, based on data shared by Google. How are these profiles created? Algorithms create these profiles by analyzing millions of reviews generated by other consumers. Analysis of prior reviews enables a site to make recommendations to consumers with no prior

information about their preferences (Levi et al. 2012).

This marketplace example highlights the fundamental role textual analysis of consumer reviews holds in creating profiles of potential consumers (Adomavicius and Tuzhilin 2015; Hariri et al. 2011). Customer generated reviews are critical for their role in providing customized recommendations to consumers (Dong et al. 2015; Levi et al. 2012; Sun et al. 2009). Since, reviews reflect consumers post-purchase consumption experience, they play an indispensable role in several business activities ranging from targeted advertising, promotions, segmentation etc. (Lee 2009; Netzer et al. 2012; Sun et al. 2009). As existing marketing literature documents, it makes logical sense for businesses to use reviews in different business activities because consumers themselves rely on other consumers reviews to make decisions (Aciar et al. 2007; Jakob et al. 2009). Reviews are considered the voice of customers by firms (Griffin and Hauser 1993) because they inform firms about consumer opinions, sentiments, needs, (Humphreys and Wang 2017) and can be used as a voluntary source of market research data (Lee and Bradlow 2011). Reviews are so important that companies such as Amazon and Yelp constantly moderate the content and act promptly if there is any tampering with the reviews¹.

The use of reviews in any marketplace decision e.g. to make recommendation to consumers, segmentation etc. relies on the premise that reviews provide a window into consumers mind that can help the firm better satisfy consumer needs by customizing the offerings and hence maximize profits. While there are hundreds of academic articles written on the importance of reviews, to the best of our knowledge there is virtually no work that documents the presence of any bias in reviews especially any unfair difference in the way women and men are portrayed.

Specifically, in this research we demonstrate that user-generated reviews hold stereotypic gender bias. Even on websites that monitor reviews for offensive content, we find the existence of gender bias. Our analysis of more than 11 million reviews from Amazon and Yelp, using a recent word-embedding algorithm, GloVe (Global Vector), indicates that the profile of women is more associated with negative (e.g. fickle, impulsive, lazy, homebound) rather than positive attributes (e.g. loyal, sensible, industrious, career oriented).

We not only investigate whether customer reviews may hold stereotypic bias, but also examine the nature of the bias that might affect firm decisions. Interventions can be successful only if we know the nature of the bias. Past research has shown that one form of bias that exists against women is because they are perceived to be more associated with family and less so with career (Greenwald, McGhee, and Schwartz 1998). Therefore in our research, we examine whether gender bias in customer reviews exists because of a bias against women (considering them less suitable for a career but more suitable for home) or a bias for men (considering them more suitable for a career but less suitable for home). Reviews are consequential for firms because they are used for many business activities ranging from product design, promotions, segmentation, targeting, recommendations etc. To examine what harm can be caused if reviews hold gender bias, we further focus on a specific use of reviews: product recommendations.

¹<https://www.amazon.com/gp/help/customer/display.html?nodeId=201929730>

Mimicking how gender bias could influence recommendations, we first identify products that are associated with positive and negative labels (e.g. impulsive/vain versus sensible/determined) and then recommend these labeled products to both male and female consumers. If consumers were more likely to choose the recommended product, it would demonstrate the harm caused by the gender bias in reviews, which are used to create the recommendations.

Ethical Consequences

The presence of bias in reviews may seem unsurprisingly consistent with the portrayal of women in media and advertising. For decades, research has documented that media portrays women as the more dependent gender (Ford et al. 1998). Since reviews are not written in a vacuum, they reflect the cultural milieu. However, one can argue that since product reviews are considered the voice-of-the-customer, it doesn't matter whether they hold gender bias. Even with biases, reviews do provide insights into the minds of the consumer. The problem with such an argument is that if a firm uses reviews, containing gender bias to design products and promotions, recommend products to consumers, determine how to portray individuals in a commercial, segment consumers etc. then it could lead to biased recommendations, messages, and insensitive portrayals. Business ethics relate to ensuring that firms know the consequences of their decision and individuals are not harmed by the decisions of a firm. For instance, if recommendations are designed based on biased reviews, the consequence can be unethical: using consumer vulnerability (e.g. impulsivity) to persuade consumers towards a certain choice against them. If the reviews contain bias, then the recommendations are biased and unethical, and could result in legal ramifications against the firm. Biased reviews, which are used to make product recommendations and to make any business decisions, can acquire a troubling distinction when we consider the following two aspects unique to algorithms, influencing society.

First, algorithms learn, amplify, and propagate. An example is Google translate: entering a gender-neutral sentence in Hindi such as *Vah ek doktor hai* (That is a doctor) gets translated in English as *He is a doctor*. Similarly, *Vah ek nurse hai* (That is a nurse) gets translated in English as *She is a nurse*. While, it is true that there are more male doctors than female doctors due to rates of burnout (Poorman 2018), the algorithm is interpreting it as a stronger association of males with doctors and females with nurses. Thus, the stereotype that doctors are male and nurses are female is perpetuated. Another example is when YouTube's autocomplete algorithm, which relies on prior textual searches by other users to predict what is being typed, began providing pedophilic suggestions (Graham 2017). The impact of propagating false stereotypes is emerging in job search algorithms as well. Incorrectly learning a social disparity that women are overrepresented in lower paying jobs as women's preference for lower paying jobs, the algorithms recommend lower paying jobs to women than to men (Bolukbasi et al. 2016). This leads to an unavoidable feedback loop that would only lead to higher female representation in lower paying jobs. Recent research underlines such a disturbing trend when it finds that online posts on a