Loïs Bilat, Yves Rychener, Nicolas Zimmermann

More at:
ada-lyn.github.io

Bias Correction for Amazon Reviews

In the past, when buying an item, one had to trust reviews in newspapers or from friends. In today's age, with online shopping, we have to tap into the minds of thousands of people who have purchased the product we are thinking about. With the help of Amazon reviews and their star-system, we can easily analyse how good the product likely is. But while a newspapers or professional reviewer is generally working hard for consistency and unbiasedness, these facts are not given for a general public reviewer writing a comment. With the help of the Amazon review dataset, we tried to find bias in the reviews, and came up with a method to reduce this bias.

Application and Limitations

The method can be helpful for vendors to better analyse their received ratings, as it helps to correct reviews of generally pessimistic/optimistic users.

However, if the results were also shown, the risk of manipulation is substantial: one could influence the correction by artificially lowering/increasing the average rating and therefore the correction we do.

Idea

We want locate and correct bias in Amazon reviews. As bias we classify patterns in reviews that are not due to product quality. Two main biases were observed that will be corrected:

- Non stationarity of reviews over time
- Users with different levels of average ratings

Other biases, such as Herding behavior (Overall trends in ranting influenced by the first review) were noticed but not corrected.

Results

The images of time-bias show clearly, that this effect has been greatly reduced. The biggest effect was observed in the year the review was made, but the other effects were also statistically significant (using 95% confidence bands deduced from the Central Limit Theorem). Substantial differences were also observed between users. The ratings given by users were corrected in the following way:

- Decrease and increase the rating if the user has an average rating higher and lower, respectively, than the average over all users
- Trust the review more (i.e. don't change as much the rating) if many users classified it as «helpful»
- Trust the review more if it is a verified purchase

Time Bias
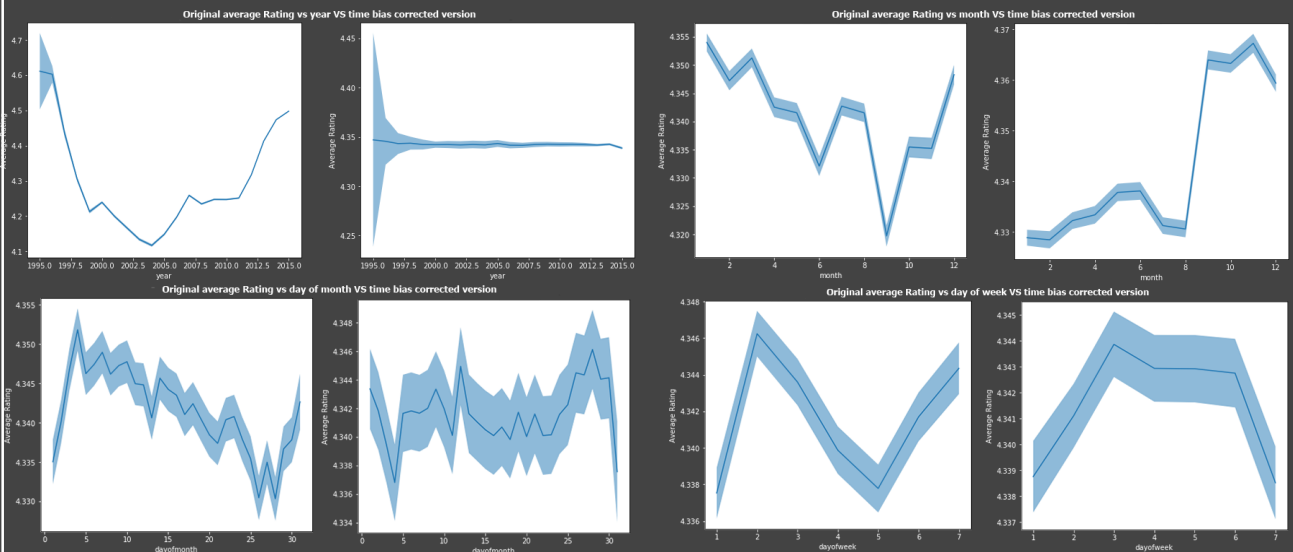
User bias

Title	Original Rating	After time bias correction	After time & user bias correction
ShowTime Popular: Level 2A (Showtime Piano) O	4.800	4.711	4.485
Mad for Decades: 50 Years of Forgettable Humor from MAD Magazine	4.250	4.195	4.252
Pearson Nurse's Drug Guide 2011	4.450	4.484	4.570
Not a Suicide Pact: The Constitution in a Time of National Emergency (Inalienable Rights)	3.750	3.851	4.289
A Darwinian Left: Politics, Evolution, and Cooperation	3.700	3.816	4.049