ACADEMIA | Letters

Artificial Intelligence in Social Networks Produces Polarization

Yaakov Stein, Tel Aviv University

Social networks have become an essential element in society. They have replaced traditional journalism as a major factor in determining the opinions of a large part of the population. And social networks use Artificial Intelligence to advertise products, and more importantly they use AI to keep people "on platform", that is to keep them "hooked".

These two uses of AI are similar – they employ a user's previous behavior to predict his or her interests, and then serve up content tailored to interest that specific user. That is why social networks are such strong advertising tools – they can predict what the user desires before the user even knows of the desire. And that is also why social networks are so addictive.

A necessary byproduct of this is that AI decides not only what advertising the user sees, but what posts, blogs, and news items are presented to him or her. And this means that a person with rightist views will never see posts somewhat left of center, but *will* see views more or less to the right (you can replace *rightist* with *leftist*, *Trump-supporter*, *green*, *belief in climate change*, *pro-vaccination*, etc.).

Over time this exposure to views exclusively belonging to one camp leads a person, even one initially only slightly off-center, to believe not only that these views make sense, but that these are the only possible views. More than the two opposing camps being separated geographically, they are separated socially. And the distance between the camps widens over time.

The fact that social networks are capable of shaping opinions and can hence contribute to polarization of society has been observed before [1,2,3]. I believe that the present analysis is the first to directly demonstrate how the prevalent use of AI-based algorithms inevitably lead to polarization of views.

Academia Letters, August 2021

Corresponding Author: Yaakov Stein, author@dspcsp.com
Citation: Stein, Y. (2021). Artificial Intelligence in Social Networks Produces Polarization. *Academia Letters*, Article 2919.

To study this effect I ran a series of simulations. I started with a population with views obeying a Gaussian distribution around the center of the spectrum. I then let random people air their views (an action we shall call "tweeting"), but only exposed people in the same half of the spectrum to these tweets. Each tweet slightly influences the views of those exposed to it, moving them in the direction of the tweeter's stance.

The simulations model a population of N=10,000 with single dimensional views x initially distributed as a Gaussian with unity standard deviation. An epoch is a sequence of N tweets, where the identity of each tweeter is randomly selected. The effect of each tweet by tweeter **i** on the view of reader **n** is zero unless sign(x**i**) = sign(x**n**), and in that case is x**n** ← x**n** + **λ** (x**i** -x**n**) where we used a step size λ=0.005 (one half of one percent).

Corresponding Author: Yaakov Stein, author@dspcsp.com
Citation: Stein, Y. (2021). Artificial Intelligence in Social Networks Produces Polarization. *Academia Letters*, Article 2919.

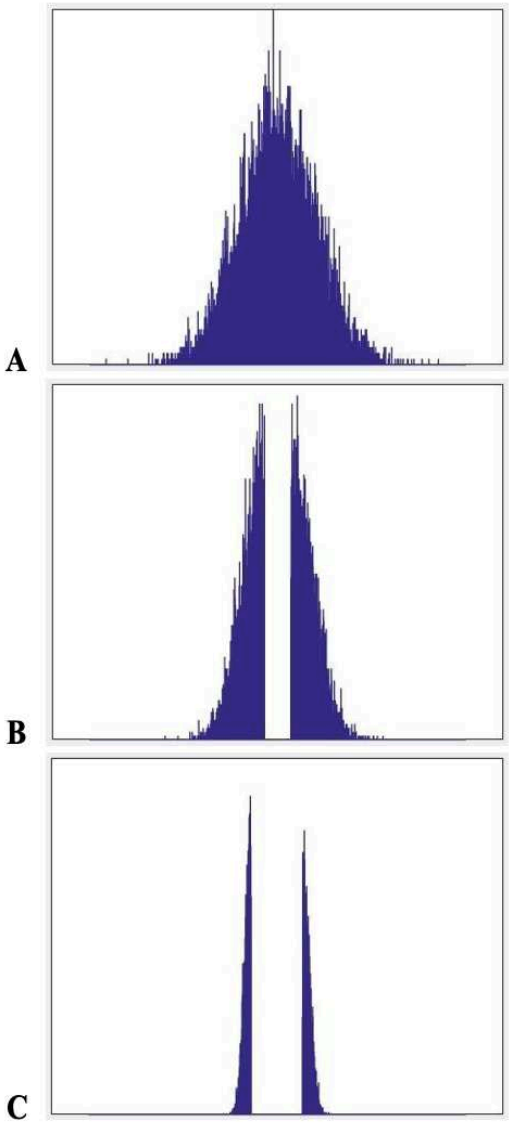

Figure Distribution of views A. Initial Gaussian distribution B. Distribution after 15 epochs C. Distribution after 50 epochs

The results are dramatic, as can be seen in the figure. Almost immediately the center of the political spectrum disappears creating two separate camps. The gap between the two camps widens over time. No matter how little each tweet influences those viewing it, the accumulated

Corresponding Author: Yaakov Stein, author@dspcsp.com
Citation: Stein, Y. (2021). Artificial Intelligence in Social Networks Produces Polarization. *Academia Letters*, Article 2919.

effect is inexorable. After enough time has elapsed the camps become worlds apart, with no possibility of anyone switching camps.

In the simulations depicted in the figure we observe that the extremes also shrink due to the mass of people closer to center influencing those further off-center. However, slightly different assumptions (e.g., assuming people with extreme views tweet more frequently or are more persuasive, or exposing people only to tweets more radical than their own views) change this aspect of the conclusions.

Another observation is that the loss of the center is so rapid that the relative size of the two groups remains close to its initial value. So, starting with an unbiased Gaussian distribution the two camps will end up of equal size. When there is an initial bias, the relative camp sizes reflect this bias.

The sensitivity of these conclusions was investigated by modifying various aspects of the simulations. Occasional exposure to tweets of the opposing camp may lead to occasional jumps from camp to camp, but does not avert polarization.

The most effective approach to avoiding polarization is to expose people to a broader spectrum of views. Simulating the influence of balanced main-stream media by adding tweets distributed according to the initial Gaussian distribution can depolarize, but only if its broadcasts are perceived as more reputable (i.e., larger λ) and applied early enough.

An alternative mechanism of avoiding polarization is to turn off the AI (i.e., to expose everyone to all tweets). But social networks have become so addicted to AI that there is little chance of that happening any time soon.

References

1. Yardi, Sarita, and Danah Boyd. "Dynamic debates: An analysis of group polarization over time on twitter." *Bulletin of science, technology & society* 30.5 (2010): 316-327.

2. Conover, Michael, et al. "Political polarization on twitter." *Proceedings of the International AAAI Conference on Web and Social Media*. Vol. 5. No. 1. 2011.

3. Conover, Michael D., et al. "Partisan asymmetries in online political activity." *EPJ Data science* 1.1 (2012): 1-19.

4. Messing, Solomon, and Sean J. Westwood. "Selective exposure in the age of social media: Endorsements trump partisan source affiliation when selecting news

online." *Communication research* 41.8 (2014): 1042-1063.

Corresponding Author: Yaakov Stein, author@dspcsp.com
Citation: Stein, Y. (2021). Artificial Intelligence in Social Networks Produces Polarization. *Academia Letters*, Article 2919.

5. Bakshy, Eytan, Solomon Messing, and Lada A. Adamic. "Exposure to ideologically diverse news and opinion on Facebook." *Science* 348.6239 (2015): 1130-1132.

6. Taylor, Cameron E., Alexander V. Mantzaris, and Ivan Garibay. "Exploring how homophily and accessibility can facilitate polarization in social networks." *Information* 9.12 (2018): 325.

Corresponding Author: Yaakov Stein, author@dspcsp.com
Citation: Stein, Y. (2021). Artificial Intelligence in Social Networks Produces Polarization. *Academia Letters*, Article 2919.