9^{th} International Conference on Computational Social Science IC^2S^2
July 17-20, 2023, Copenhagen, Denmark

How do recommendation algorithms learn and leverage political preferences of users?

Keywords: Recommender systems, Political attitudes, Algorithm Transparency, Explainability, Social Media

Extended Abstract

The last few decades have seen a growing interest in the impact of social media [7, 8]. One of the reasons for this is their increasing impact on social and political dynamics [4], which raises concerns about several hypothetical phenomena [2]: echo chambers, polarization, filter bubbles, extremism, etc. Furthermore, research indicates that the recommendation algorithms might play a role in some of those phenomena [10, 11].

While past research has focused on the outputs of the recommendation algorithms [1], our work aims to focus on the algorithm itself and on the process of computation used to create the recommendations. By doing this we integrate our work in the growing literature around algorithmic interpretability and explainability [9]. And we put the basis for further work on algorithmic audit and transparency highly researched by the Europeans public institutions [3].

In this work we focus on User-Content recommendation algorithms (algorithm predicting which content will fit for a given user) because of their vast use in the market. Content recommendation algorithms from the state of the art mainly work in 2 steps [5]. The first step consists in estimating from the interaction between users and the content an embedding (of the users and the content) which takes into account those interactions. The second step (usually computationally simpler than the first) estimates, from the embedding, the probability that a given user likes a given content. We argue that this embedding is trained to be a meaningful dimension reduction from the original data that give us information on which aspects of the data are leveraged by the algorithms to build recommendations.

Recent work by Ramaciotti Morales et al. [12] has shown that it is possible to estimate the political attitudes of twitter users from the members of the parliament (MPs) they follow. We extract users from the french twitter sphere, and estimate their political attitudes on 2 axes : left-right and attitude toward elite (see fig 1).

Taking those users and the URLs they shared as data, we develop a recommendation algorithm able to predict which URLs will be shared. We used a standard method of non-negative matrix factorisation [6], based on collaborative filtering methods (which assumes that similar users will like similar content). We trained the algorithm on the twitter data and evaluated it using the classical *Hits@10* metric.

We then compare the algorithmic embedding of the users with their political attitudes and their socio-demographic profile. We finally show that there is a statistical relationship between those quantities. In particular we see (fig 2) that some algorithmic dimensions are sorting people by political attitude. Furthermore we show (fig 3) that the political attitudes have a clear impact on the algorithm embedding. We can confirm our hypothesis by seeing (tab 1) that the contents associated with those particular dimensions are commonly known in France as left-wing (or right-wing) medias.

Our work aims to propose a proof of concept for a new method of explanation for recommendation algorithms and open the possibility to a new branch of studies where suspected social media issues (as polarization or filter bubbles) can be studied directly by looking at the recommendation mechanism. Furthermore our work gives a new light to the sociological debate around algorithmic feedback loops and machine habitus.

References

[1] E. Bakshy, S. Messing, and L. A. Adamic. Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6239):1130–1132, June 2015.

[2] Y. Benkler, R. Faris, and H. Roberts. *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*. Oxford University Press, Sept. 2018.

[3] L. C. d'État. S'engager dans l'intelligence artificielle pour un meilleur service public. https://www.conseil-etat.fr/actualites/s-engager-dans-l-intelligence-artificielle-pour-un-meilleur-service-public, Aug. 2022.

[4] N. Gaumont, M. Panahi, and D. Chavalarias. Reconstruction of the socio-semantic dynamics of political activist Twitter networks—Method and application to the 2017 French presidential election. *PLOS ONE*, 13(9):e0201879, Sept. 2018.

[5] D. Kluver, M. D. Ekstrand, and J. A. Konstan. Rating-Based Collaborative Filtering: Algorithms and Evaluation. In P. Brusilovsky and D. He, editors, *Social Information Access: Systems and Technologies*, Lecture Notes in Computer Science, pages 344–390. Springer International Publishing, Cham, 2018.

[6] Y. Koren, R. Bell, and C. Volinsky. Matrix Factorization Techniques for Recommender Systems. *Computer*, 42(8):30–37, Aug. 2009.

[7] D. Lazer, A. Pentland, L. Adamic, S. Aral, A.-L. Barabási, D. Brewer, N. Christakis, N. Contractor, J. Fowler, M. Gutmann, T. Jebara, G. King, M. Macy, D. Roy, and M. Van Alstyne. Computational Social Science. *Science*, 323(5915):721–723, Feb. 2009.

[8] D. M. J. Lazer, A. Pentland, D. J. Watts, S. Aral, S. Athey, N. Contractor, D. Freelon, S. Gonzalez-Bailon, G. King, H. Margetts, A. Nelson, M. J. Salganik, M. Strohmaier, A. Vespignani, and C. Wagner. Computational social science: Obstacles and opportunities. *Science*, 369(6507):1060–1062, Aug. 2020.

[9] R. Marcinkevičs and J. E. Vogt. Interpretability and Explainability: A Machine Learning Zoo Mini-tour, Dec. 2020.

[10] T. T. Nguyen, P.-M. Hui, F. M. Harper, L. Terveen, and J. A. Konstan. Exploring the filter bubble: The effect of using recommender systems on content diversity. In *Proceedings of the 23rd International Conference on World Wide Web*, WWW '14, pages 677–686, New York, NY, USA, Apr. 2014. Association for Computing Machinery.

[11] E. Noordeh, R. Levin, R. Jiang, and H. Shadmany. Echo Chambers in Collaborative Filtering Based Recommendation Systems. *arXiv:2011.03890 [cs]*, Nov. 2020.

[12] P. Ramaciotti Morales, J.-P. Cointet, G. Muñoz Zolotoochin, A. Fernández Peralta, G. Iñiguez, and A. Pournaki. Inferring Attitudinal Spaces in Social Networks. Feb. 2022.

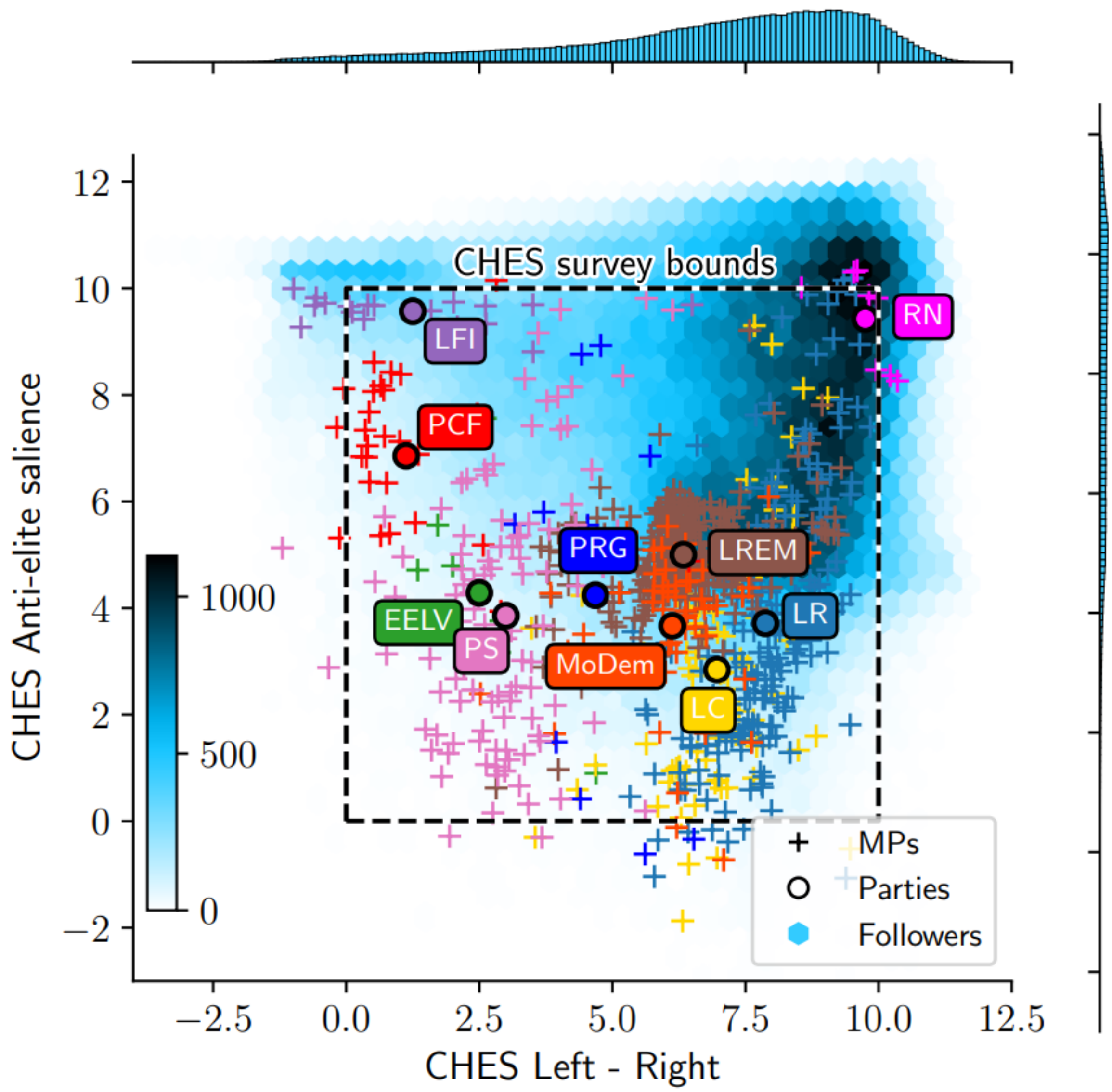


Figure 1: Distribution of our user sample on the 2 political attitudes axes. Left-Right refers to the standard political measure, Anti-elite Salience expresses a negative attitude toward elite. (image from [12])

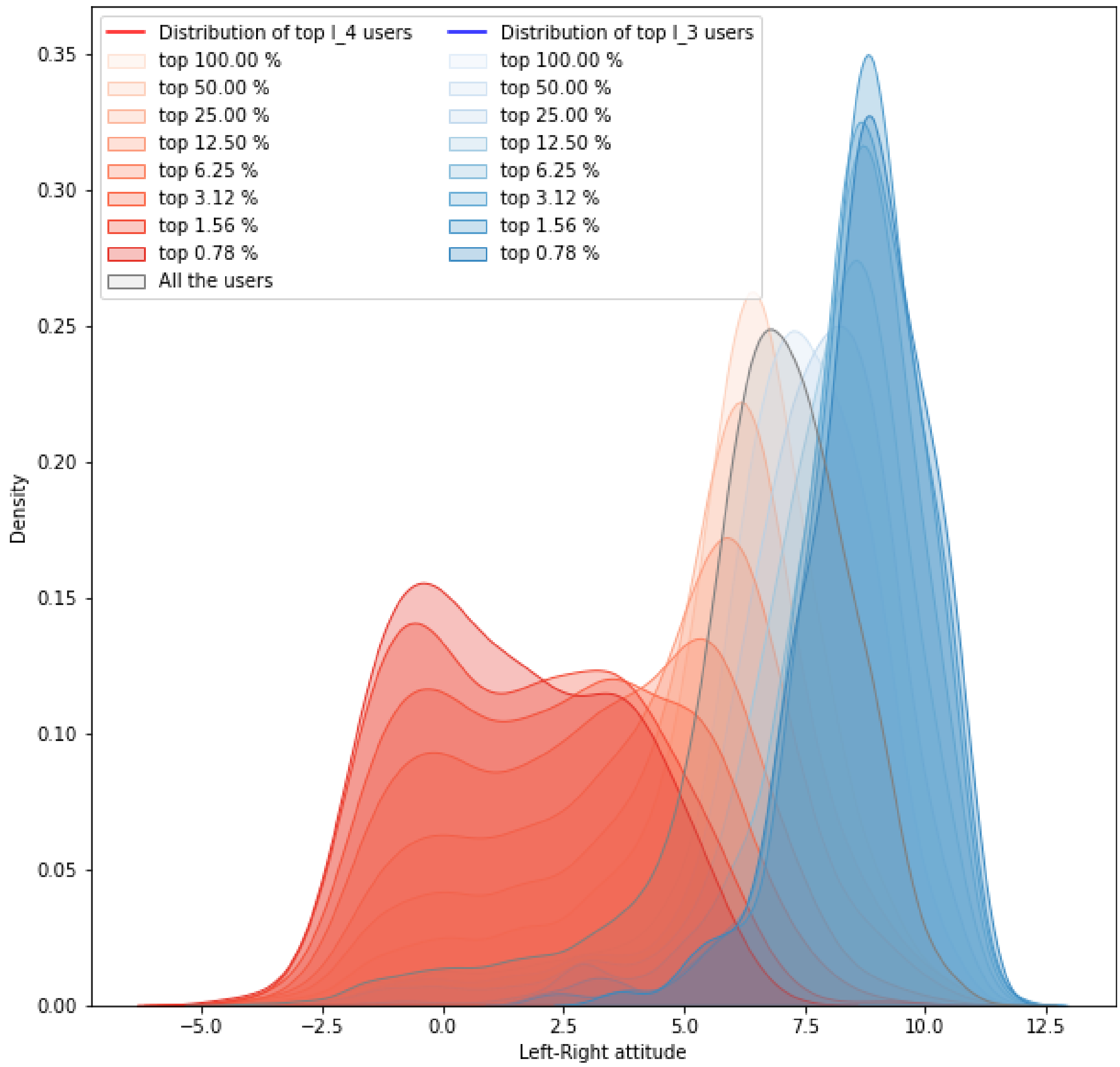
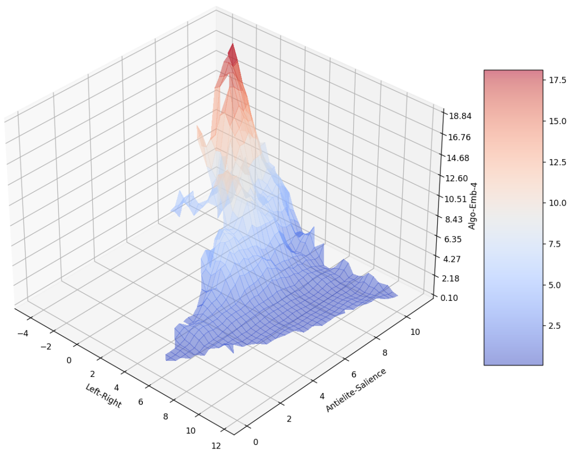
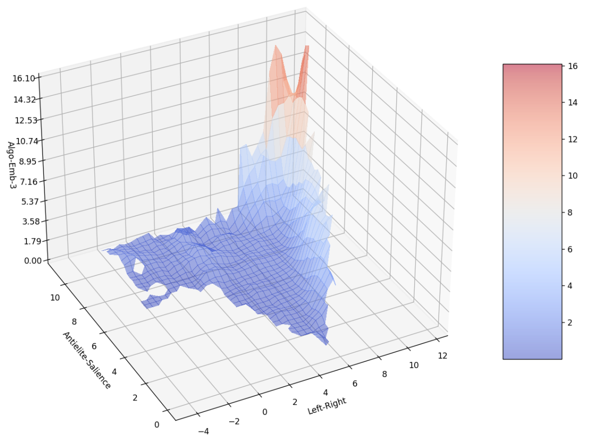


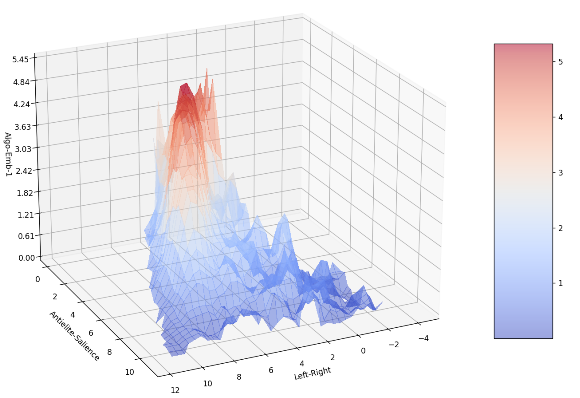
Figure 2: We show that some dimensions of the algorithm embedding select users only from some part of the political spectrum. Here we see the variation of the distribution on the Left-Right ax of the top users for 2 algorithm embedding dimensions (l_4, l_3). Being classified as "high l_4" by the algorithm highly augments the probability that a user is left wing (or right wing for l_3).



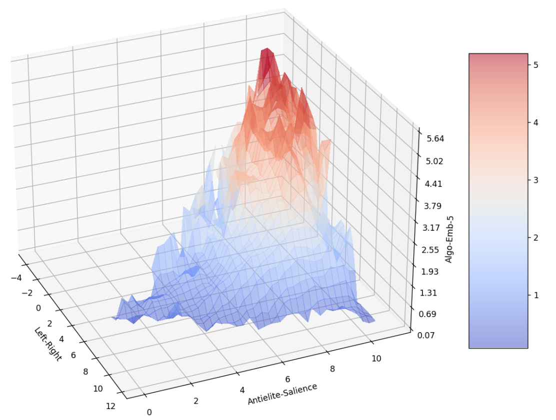
(a) Average l_4 (algorithm embedding dimension 4) as a function of political attitude.



(b) Average l_3 as a function of political attitude.



(c) Average l_1 as a function of political attitude.



(d) Average l_5 as a function of political attitude.

Figure 3: We can see from these graphs that the political attitudes of users influence the expected position in the algorithmic embedding for some specific dimensions.

	Top content : l_4	**Top content : l_3**
1	politis.fr	tvlibertes.com
2	revolutionpermanente.fr	adoxa.info
3	alternatives-economiques.fr	lefigaro.fr
4	monde-diplomatique.fr	causeur.fr
5	change.org	actu17.fr
6	bastamag.net	bvoltaire.fr
7	mediapart.fr	fr.sputniknews.com
8	blogs.mediapart.fr	francais.rt.com
9	humanite.fr	fdesouche.com
10	reporterre.net	valeursactuelles.com

Table 1: Here we can see the domains with the highest values of l_4 (expected left-wing) and l_5 (expected right-wing) in the algorithmic embedding.