The closed loop between opinion formation and personalised recommendations

Wilbert Samuel Rossi[1], Jan Willem Polderman[2], Paolo Frasca[3]

[2]University of Twente (Netherlands)

[1]University of Groningen (Netherlands)

[3]CNRS, Grenoble (France)

Outline

1. Why this research (2015–): online opinion formation
2. Dynamical model: interconnecting user and recommender
   - User model: opinion dynamics & clicking behavior
   - Recommender model
3. Results on the closed-loop system
   - Types of trajectories
   - Simulations and analytical results
4. Conclusion

Nowadays, much social dynamics take place on **online social media**:

- online dynamics influence opinion formation and offline behaviours [Aral (2012)];
- online dynamics depend on *how* online platforms distribute information between their users [Hirst (2017)].

Online platforms are accused of producing "information disorders":

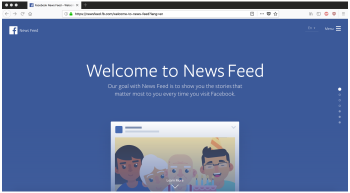
- the formation of *filter bubbles* [Pariser (2011)]
- the viral spreading of *fake news* [Venturini (2019)]
- unnatural accelerations of the *attention cycles* [Castaldo et al. (2022)]

The role of recommendation systems

Online platforms have to manage huge amounts of information. How do they do it?

Modern online platforms (Facebook, YouTube, Tiktok, Quora,. . .) use **recommendation systems** to optimize user experience and engagement.

Recommendation systems produce (possibly endless) "feeds" for users to scroll.



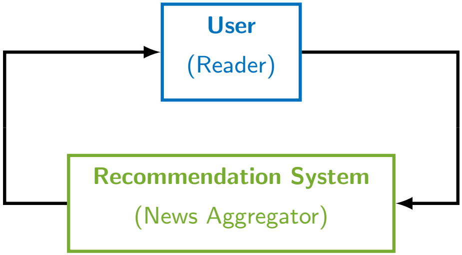
Personalized feeds have become the dominant way for users to access online content.

Starting points of this research:

- Online social dynamics cannot be understood without understanding how recommendation systems work and how they interact with the users.
- Being the recommendations personalized, users and recommendation systems form a closed loop

Dynamical model

Case study: a **news aggregator** that recommends news articles to a **reader**



User model: Opinion dynamics

User has a time-dependent continuous *opinion* $o_{usr}(t) \in [-1, 1]$ about an issue

At every (discrete) time t,

- the user receives one article that has binary *position* $p_{art}(t) \in \{-1, 1\}$
- the user updates her opinion by

$$o_{usr}(t+1) = \alpha\, o^0_{usr} + \beta\, o_{usr}(t) + \gamma\, p_{art}(t) \qquad t \in \mathbb{N}_0$$

where

$o^0_{usr} \in [-1, 1]$ is a *prejudice* that coincides with initial opinion (i.e. $o_{usr}(0) = o^0_{usr}$)

$\alpha, \beta, \gamma \geq 0$ and $\alpha + \beta + \gamma = 1$ are weights that describe the relative importance of prejudice, memory, and new information

This opinion formation model is based on Chaiken (1987); Friedkin and Johnsen (1990)

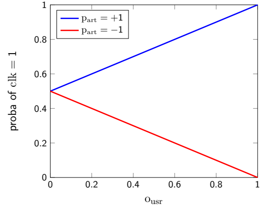
User Model: To click or not to click?

At time t, the user also decides whether to *read the recommended article or not*

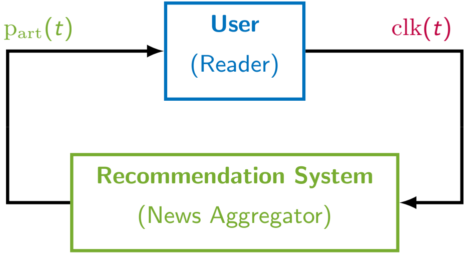
We assume that the user is subject to a **confirmation bias** [Nickerson (1998)]: she prefers contents that are consistent with her opinion $\mathrm{o_{usr}}$

To model this confirmation bias, each *click decision* $\mathrm{clk} \in \{0, 1\}$ follows the **stochastic** law [Dandekar et al. (2013)]:

$$\mathrm{clk}(t) = \begin{cases} 1 & \text{with probability } \frac{1}{2} + \frac{1}{2}\,\mathrm{o_{usr}}(t)\,\mathrm{p_{art}}(t) \\ 0 & \text{with probability } \frac{1}{2} - \frac{1}{2}\,\mathrm{o_{usr}}(t)\,\mathrm{p_{art}}(t) \end{cases}$$



The recommender's perspective



The recommendation system:

- measures and records whether the user clicks or not on each article
- has the **purpose of maximizing clicks**: more precisely, the recommender's purpose is the online optimization of the *click-through rate* $\mathrm{ctr}(t) = \frac{1}{t}\sum_{s=0}^{t-1} \mathrm{clk}(s)$

Recommender model: solving a bandit problem

In our very simple setting, the recommender only has two options and has to choose one at each time step.

Therefore, the online optimization of the click-through rate is equivalent to taking the user as a **"two-armed bandit"**, where the two actions (arms) are -1 and +1 and the reward is the click action

This is a one-armed bandit...

- The recommender faces the **exploration-exploitation dilemma** of sequential decision problems between staying with the most successful option so far (i.e. exploitation) and testing the other option (i.e. exploration), which might become better in the future [Bubeck and Cesa-Bianchi (2012); Li et al. (2010)]
- The user is a bandit with non-stationary rewards: the reward depends on the current state (opinion), which in turn depends on the previous actions

ϵ-greedy recommender model

We assume that the recommender *balances exploration and exploitation by an* **ϵ-greedy algorithm**

$$\mathrm{p_{art}}(t) = \begin{cases} \textit{exploitation} & \text{with probability } 1-\epsilon \\ \textit{exploration} & \text{with probability } \epsilon \end{cases}$$

To this purpose, the recommender just needs to compute the **most successful "arm"**, by the following algorithm

- Define counters that track
  - recommendations $\mathrm{r}_+(t)$, $\mathrm{r}_-(t)$

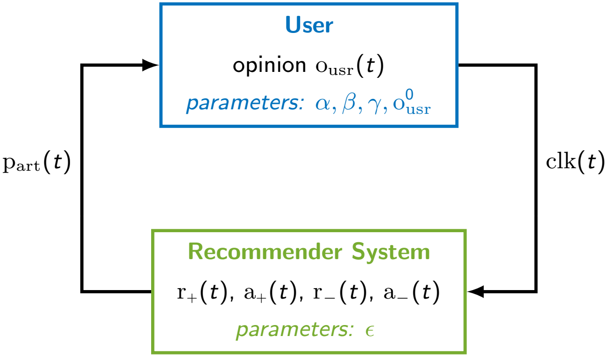
$$T_+(t) = \{s : 0 \le s \le t-1 \text{ and } \mathrm{p_{art}}(s) = +1\} \qquad \mathrm{r}_+(t) = \# T_+$$
$$T_-(t) = \{s : 0 \le s \le t-1 \text{ and } \mathrm{p_{art}}(s) = -1\} \qquad \mathrm{r}_-(t) = \# T_-$$

  - 'successes' $\mathrm{a}_+(t)$, $\mathrm{a}_-(t)$: $\quad \mathrm{a}_+(t) = \sum_{s \in T_+(t)} \mathrm{clk}(s)\,, \qquad \mathrm{a}_-(t) = \sum_{s \in T_-(t)} \mathrm{clk}(s)$

- Apply the randomized decision rule (with small $\epsilon > 0$):

$$\begin{cases} \text{if } \frac{\mathrm{a}_+(t)}{\mathrm{r}_+(t)} > \frac{\mathrm{a}_-(t)}{\mathrm{r}_-(t)} \text{ then } & \mathbb{P}(\mathrm{p_{art}}(t) = 1) = 1-\epsilon, & \mathbb{P}(\mathrm{p_{art}}(t) = -1) = \epsilon \\ \text{if } \frac{\mathrm{a}_+(t)}{\mathrm{r}_+(t)} = \frac{\mathrm{a}_-(t)}{\mathrm{r}_-(t)} \text{ then } & \mathbb{P}(\mathrm{p_{art}}(t) = 1) = 0.5, & \mathbb{P}(\mathrm{p_{art}}(t) = -1) = 0.5 \\ \text{if } \frac{\mathrm{a}_+(t)}{\mathrm{r}_+(t)} < \frac{\mathrm{a}_-(t)}{\mathrm{r}_-(t)} \text{ then } & \mathbb{P}(\mathrm{p_{art}}(t) = 1) = \epsilon, & \mathbb{P}(\mathrm{p_{art}}(t) = -1) = 1-\epsilon \end{cases}$$

Summary: feedback interconnection



The closed-loop system

The closed loop systems with state $\mathbf{x}(t) = [\mathrm{r}_+(t), \mathrm{r}_-(t), \mathrm{a}_+(t), \mathrm{a}_-(t), \mathrm{o}_{\mathrm{usr}}(t)]^\top$ follows the stochastic dynamics

$$x(t+1) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \beta \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \alpha \mathrm{o}^0_{\mathrm{usr}} \end{bmatrix} + \mathbf{f}(\mathbf{x})$$

with initial datum $\mathbf{x}(0) = [0, 0, 0, 0, \mathrm{o}^0_{\mathrm{usr}}]^\top$.

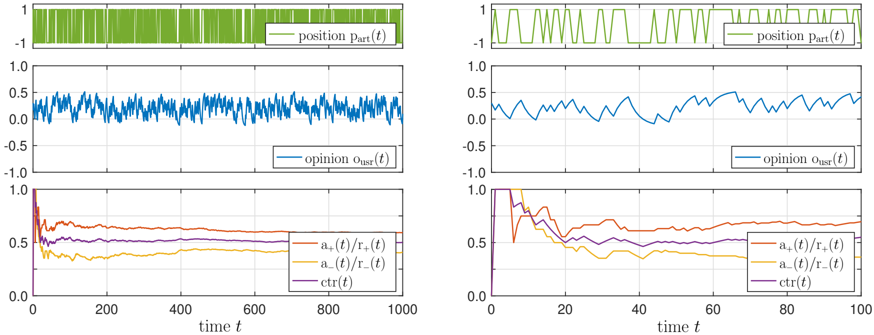
The random vector $\mathbf{f}(\mathbf{x}(t))$ can take on four values, corresponding to the cases "position $+1$, no click", "position $+1$, get click", "position -1, no click" and "position -1, get click":

$$\mathbf{f}(\mathbf{x}) = \begin{cases} [1, 0, 1, 0, +\gamma]^\top \text{ with probability } h_\epsilon(\Delta)\left(\frac{1}{2} + \frac{1}{2}\mathrm{o}_{\mathrm{usr}}\right) \\ [1, 0, 0, 0, +\gamma]^\top \text{ with probability } h_\epsilon(\Delta)\left(\frac{1}{2} - \frac{1}{2}\mathrm{o}_{\mathrm{usr}}\right) \\ [0, 1, 0, 1, -\gamma]^\top \text{ with probability } \left(1 - h_\epsilon(\Delta)\right)\left(\frac{1}{2} - \frac{1}{2}\mathrm{o}_{\mathrm{usr}}\right) \\ [0, 1, 0, 0, -\gamma]^\top \text{ with probability } \left(1 - h_\epsilon(\Delta)\right)\left(\frac{1}{2} + \frac{1}{2}\mathrm{o}_{\mathrm{usr}}\right), \end{cases}$$

where $h_\epsilon(s) := \begin{cases} 1-\epsilon & \text{if } s > 0 \\ \frac{1}{2} & \text{if } s = 0 \\ \epsilon & \text{if } s < 0 \end{cases}$ and $\Delta(\mathbf{x}(t)) := \frac{\mathrm{a}_+(t)}{\mathrm{r}_+(t)} - \frac{\mathrm{a}_-(t)}{\mathrm{r}_-(t)}$.

Clearly, $\Delta(\mathbf{x}(t)) > 0$ if and only if $+1$ is the most successful recommendation so far.

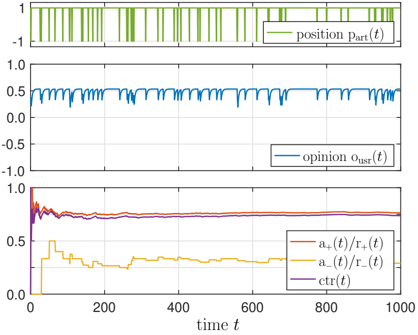
Results: closed-loop behavior

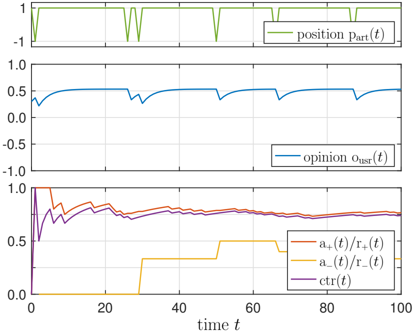


Left: up to time $t_{\max} = 1000$. Right: zooming into the first 100 steps

Parameters: $\alpha = 0.15$, $\beta = 0.70$, $\gamma = 0.15$, $\mathrm{o}^0_{\mathrm{usr}} = 0.30$ and $\epsilon = 0.50$

Sample trajectories: Non-random recommendations



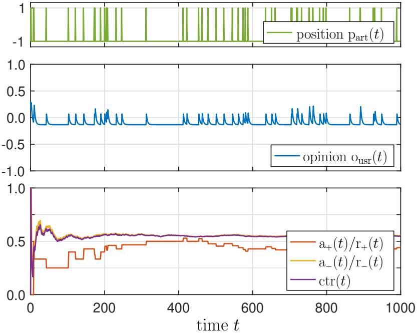


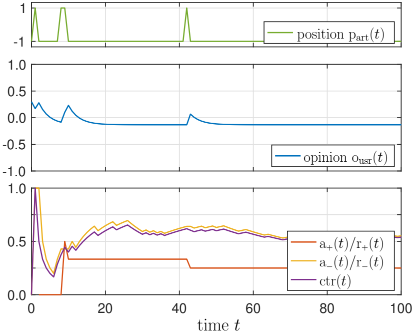
Left: up to time $t_{max} = 1000$.

Right: zooming into the first 100 steps

Parameters: $\alpha = 0.15$, $\beta = 0.70$, $\gamma = 0.15$, $o^0_{usr} = 0.30$ and $\epsilon = 0.05$

Note: Here the most recommended position is $+1$





Left: up to time $t_{max} = 1000$. Right: zooming into the first 100 steps

Parameters: $\alpha = 0.15$, $\beta = 0.70$, $\gamma = 0.15$, $o^0_{usr} = 0.30$ and $\epsilon = 0.05$

Note: Here the most recommended position is -1

We could study the average dynamics $\mathbb{E}[\mathbf{x}(t)]$, but we don't do that, because:

1. the dynamics of $\mathbb{E}[\mathbf{x}(t)]$ is impractical to write due to the **nonlinearities** and **statistical dependences** between the variables
2. since there are **two kinds of trajectories** (very different from each other), the average would be a poor description of either

Our approach:

1. Focus on **conditional expectations**, depending on the type of trajectory:

$$\mathbb{E}^{+}[\mathbf{x}(t)] := \mathbb{E}[\mathbf{x}(t) \mid +1 \text{ is more likely}]$$
$$\mathbb{E}^{-}[\mathbf{x}(t)] := \mathbb{E}[\mathbf{x}(t) \mid -1 \text{ is more likely}]$$

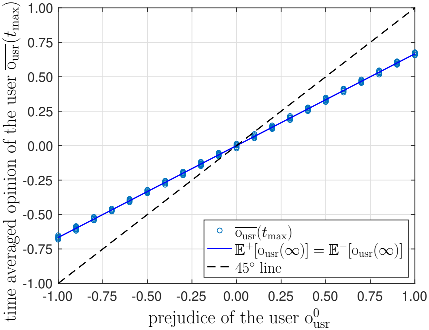
2. Write and solve the **linear** dynamics for $\mathbb{E}^{\pm}[\mathbf{x}(t)]$

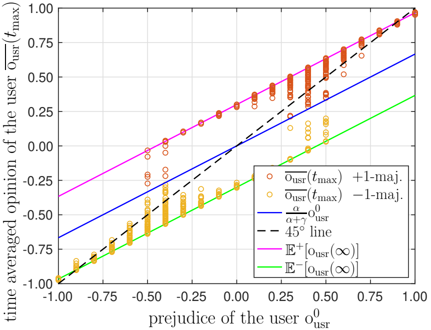
$$\mathbb{E}^{+}[\mathbf{x}(t+1)] = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \frac{1}{2}(1-\epsilon) \\ 0 & 0 & 0 & 1 & -\frac{1}{2}\epsilon \\ 0 & 0 & 0 & 0 & \beta \end{bmatrix} \mathbb{E}^{+}[\mathbf{x}(t)] + \begin{bmatrix} 1-\epsilon \\ \epsilon \\ \frac{1}{2}(1-\epsilon) \\ \frac{1}{2}\epsilon \\ \alpha o^{0}_{\mathrm{usr}} + \gamma(1-2\epsilon) \end{bmatrix}$$

3. Compare analytical $\mathbb{E}^{\pm}[\mathbf{x}(t)]$ with simulated **time-average** $\bar{\mathbf{x}}(t) = \frac{1}{t}\sum_{s=0}^{t-1} \mathbf{x}(s)$

Results (matching analysis with simulations)

Opinion trajectories split into $+1$-trajectories and -1-trajectories, concentrating around the conditional expectations

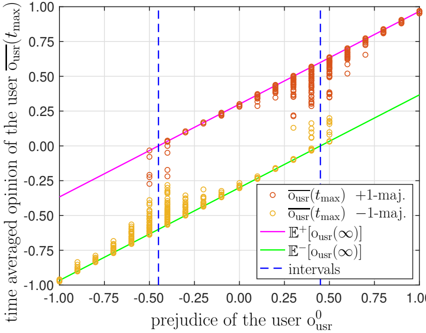


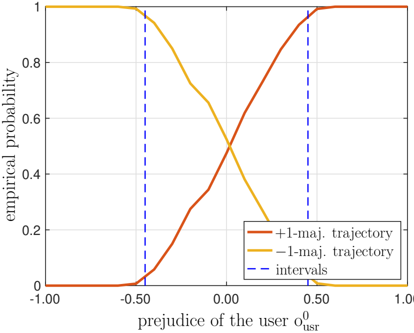


Parameters: $\alpha = 0.15$, $\beta = 0.70$, $\gamma = 0.15$. Left: $\epsilon = 0.50$ (random). Right: $\epsilon = 0.05$

Expected opinions (for large time): $\lim_{t\to\infty} \mathbb{E}^{\pm}[o_{usr}(t)] = \dfrac{\alpha o_{usr}^0 \pm \gamma(1-2\epsilon)}{\alpha+\gamma}$

Strong prejudices lead to recommendations that adhere to the prejudices

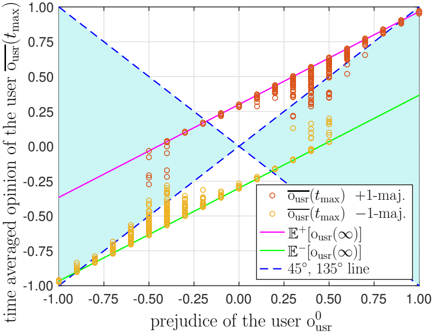




Parameters: $\alpha = 0.20$, $\beta = 0.70$, $\gamma = 0.10$, $\epsilon = 0.05$.

Dashed blue lines have abscissas $-\frac{\gamma}{\alpha}(1-2\epsilon)$ and $\frac{\gamma}{\alpha}(1-2\epsilon)$: between the lines, both $+1$ and -1-trajectories are theoretically allowed in the long run

Most trajectories produce final opinions that are more extreme than the initial ones (polarization)

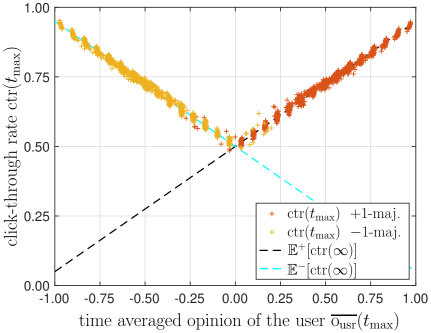


Parameters: $\alpha = 0.20$, $\beta = 0.70$, $\epsilon = 0.05$

In shaded areas, the final time-averaged opinion $\overline{o_{usr}}(t_{max})$ is *less extreme* than prejudice o^0_{usr}; in white areas, it is *more extreme*. Most samples fall into white areas.

Combined effects on opinions and click-through rate

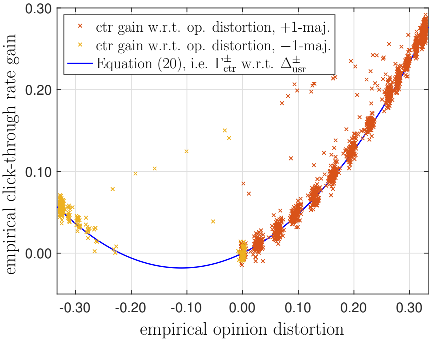
Recommendations are more effective when opinions are more extreme



Parameters: $\alpha = 0.20$, $\beta = 0.70$, $\epsilon = 0.05$

$$\mathbb{E}^{\pm}[\mathrm{ctr}(\infty)] = \frac{1}{2} \pm \frac{1}{2}(1 - 2\epsilon)\frac{\alpha o_{usr}^0 \pm \gamma(1-2\epsilon)}{\alpha+\gamma}$$

The randomness parameter ϵ controls the trade-off between impact on the opinions and achievable click-through rate



Opinion distortion $\Delta^{\pm}_{\mathrm{usr}} := \mathbb{E}^{\pm}[\mathrm{o}_{\mathrm{usr}}(\infty);\epsilon] - \mathbb{E}^{\pm}[\mathrm{o}_{\mathrm{usr}}(\infty);\epsilon = 0.5] = \pm\frac{\gamma}{\alpha+\gamma}(1-2\epsilon)$

Click-through rate gain

$$\Gamma^{\pm}_{\mathrm{ctr}} := \mathbb{E}^{\pm}[\mathrm{ctr}(\infty);\epsilon] - \mathbb{E}^{\pm}[\mathrm{ctr}(\infty);\epsilon = 0.5] = \pm\frac{1}{2}\frac{\alpha}{\alpha+\gamma}\mathrm{o}^{0}_{\mathrm{usr}}(1-2\epsilon) + \frac{1}{2}\frac{\gamma}{\alpha+\gamma}(1-2\epsilon)^2$$

blue line is $\Gamma^{\pm}_{\mathrm{ctr}} = \frac{1}{2}\frac{\alpha}{\gamma}\mathrm{o}^{0}_{\mathrm{usr}}\Delta^{\pm}_{\mathrm{usr}} + \frac{1}{2}\frac{\alpha+\gamma}{\gamma}(\Delta^{\pm}_{\mathrm{usr}})^2$

click-through rate gain is a function of opinion distortion!

Conclusion

Summary & Discussion

Summary

- We presented an analytical model of user-recommender interaction, motivated by news aggregators and constructed from "prime principles"
- The model predicts a strong connection between personalized recommendations and distorted opinion evolution
- The model identifies a trade-off between opinion distorsion and effectiveness of the recommendations, controlled by the randomness level of the recommendations

  W. S. Rossi, J. W. Polderman and P. Frasca, "The closed loop between opinion formation and personalized recommendations," in IEEE Transactions on Control of Network Systems, *9(3): 1092-1103, Sept. 2022*

Discussion & Validation

The current model is a "conceptual" model of the user-recommender interconnection, which focuses on the role of the personalization:
All its "ingredients" are sound, but the model remains hard to validate and identify from data, because

- users, recommendation systems, and generally the platforms, are much more complex than this model;
- good data (about the whole feedback loop!) is scarcely available

Open problems

Refinements

- Refine the probabilistic insights and justify the assumptions made in the current analysis
- Design optimal recommender algorithms for the closed-loop dynamics

Relevant extensions

- Multiple users (possibly, connected through a social network)
- Non-binary recommendations (more than two choices, perhaps a continuum of choices, are possible)
- Recommendations via collaborative filtering
- Multiple recommendations (several, sorted items are proposed together)

Validation:

- Identify suitable dataset
- Match (via appropriate abstractions) this coarse model with the data

Some references

Aral, S. (2012). Social science: Poked to vote. *Nature*, 489(7415):212–214.

Bressan, M., Leucci, S., Panconesi, A., Raghavan, P., and Terolli, E. (2016). The limits of popularity-based recommendations, and the role of social ties. In *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, pages 745–754, New York, NY, USA.

Bubeck, S. and Cesa-Bianchi, N. (2012). Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends® in Machine Learning*, 5(1):1–122.

Castaldo, M., Venturini, T., Frasca, P., and Gargiulo, F. (2022). Junk news bubbles modelling the rise and fall of attention in online arenas. *New Media & Society*, 24(9):2027–2045.

Chaiken, S. (1987). The heuristic model of persuasion. In *Social influence: the Ontario symposium*, volume 5, pages 3–39. Psychology Press, New York, US.

Dandekar, P., Goel, A., and Lee, D. (2013). Biased assimilation, homophily, and the dynamics of polarization. *Proceedings of the National Academy of Sciences*, 110(15):5791–5796.

Friedkin, N. E. and Johnsen, E. C. (1990). Social influence and opinions. *Journal of Mathematical Sociology*, 15(3-4):193–206.

Hirst, M. (2017). Towards a political economy of fake news. *The Political Economy of Communication*, 5(2).

Jiang, R., Chiappa, S., Lattimore, T., György, A., and Kohli, P. (2019). Degenerate feedback loops in recommender systems. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, AIES '19, pages 383–390, New York, NY, USA. Association for Computing Machinery.

Li, L.and Chu, W., Langford, J., and Schapire, R. E. (2010). A contextual-bandit approach to personalized news article recommendation. In *Proceedings of the 19th International Conference on World Wide Web*, WWW'10, pages 661–670, New York, NY, USA.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2):175–220.

Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. Penguin UK.

Venturini, T. (2019). From fake to junk news, the data politics of online virality. In Bigo, D., Isin, E., and Ruppert, E., editors, *Data Politics: Worlds, Subjects, Rights*. Routledge, London.