Learning with Feedback Loops

DJ Thornton — UNSW Sydney

My website!

Motivation

Many learning processes have the potential problem that realizations of the data affect the data generating process itself.

This is true of all *recommendation platforms:*

- **Search engines** recommend high-traffic websites.
- **Social media platforms** promote content with high engagement.
- **E-commerce platforms** highlight bestselling products.
- **Streaming services** prioritize popular movies.
- **Large language models** generate answers based on information in the training data.

Key questions

Key question #1: *What are the implications for learning if the platform fails to account for the **feedback** in the data-generating process?*

Key question #2: *If learning fails, how sophisticated must the algorithm be for learning to occur?*

Recommendation algorithm

- A **recommendation algorithm** is a rule for generating recommendations as a function of:
  - The history of actions $\mathbf{A}_t$, and
  - The history of recommendations $\mathbf{R}_t$.
- I focus on algorithms for which the subjective belief is a sufficient statistic for the recommendation rule.
- A **naïve recommendation platform** is one whose subjective belief does not account for the effect of recommendations on learning from the data.

Model

- I develop a model of *intermediated social-learning*.
- There is an unknown state of the world $\theta \in \{L, H\}$.
- Players: A long-lived **platform** and short-lived **agents** $1,2,\ldots$
- In period t :
  - Platform provides a **recommendation** $R_t \in \{0,1\}$ to the agent, which is implemented by an **algorithm.**
  - Agent observes an exogenous signal $s_t \sim F^\theta$ and chooses an **action** $a_t \in \{0,1\}$.
  - The platform observes the action of the agent but not the signal.

Utility

- Agent t receives a payoff of 1 if their action matches the state, and 0 otherwise,

$$u_t(a_t) = \mathbb{I}[a_t = a^\theta],$$

- where $a^L = 0$ and $a^H = 1$.
- The platform's objective is such that it prefers for the algorithm to make **truthful recommendations.**

Equilibrium

- I characterise Bayesian Nash equilibria of the game.
- An equilibrium strategy for agents must satisfy:

$$a_t(R_t, s_t) \in \arg\max_a \mathbb{E}_t[u_t(a) \mid R_t, s_t] .$$

- Say that *asymptotic learning occurs* (in equilibrium) if

$$\liminf_{t\to\infty} \mathbb{P}(\theta = H \mid R_t, s_t) = 1.$$

and that ***learning fails*** otherwise.

Results

Proposition 1 (Failure of learning)

Under a naïve recommendation platform, learning fails.

Intuitively:

- Asymptotic learning can occur iff the platform eventually "learns" the true state with certainty.
- The platform does not properly correct for "contradictory" behaviour $\implies$ failures can persist **even when agents' information is arbitrarily good.**

Corollary 1 (Mass of mistakes)

If the platform recommends $R_\infty = 0$ and $\theta = H$, then the limiting fraction of agents who take the incorrect action is $\mathbb{P}(s_t \leq \pi) = F^H(\pi)$.

Intuitively:

- The fraction of agents who play the incorrect action are the fraction who observe signals which are "not convincing enough" given the platform's recommendation.

Proposition 2 (Sophisticated algorithms)

Under a Bayesian recommendation algorithm, learning occurs. Moreover, there are less sophisticated algorithms under which learning occurs.

Intuitively:

- If platform is fully Bayesian, learning occurs for "standard reasons".
- A full characterisation depends on what we allow the platform to communicate.