Reward innovation for long-term member satisfaction

Gary Tang, Jiangwei Pan, Henry Wang and Justin Basilico

NETFLIX

Goal

Create a **personalized homepage** to help members find content to watch and enjoy that

maximizes **long-term** member **satisfaction**.

Long-term satisfaction for Netflix

Member enjoys watching Netflix

So continues the subscription

and

tells their friends about it

NETFLIX

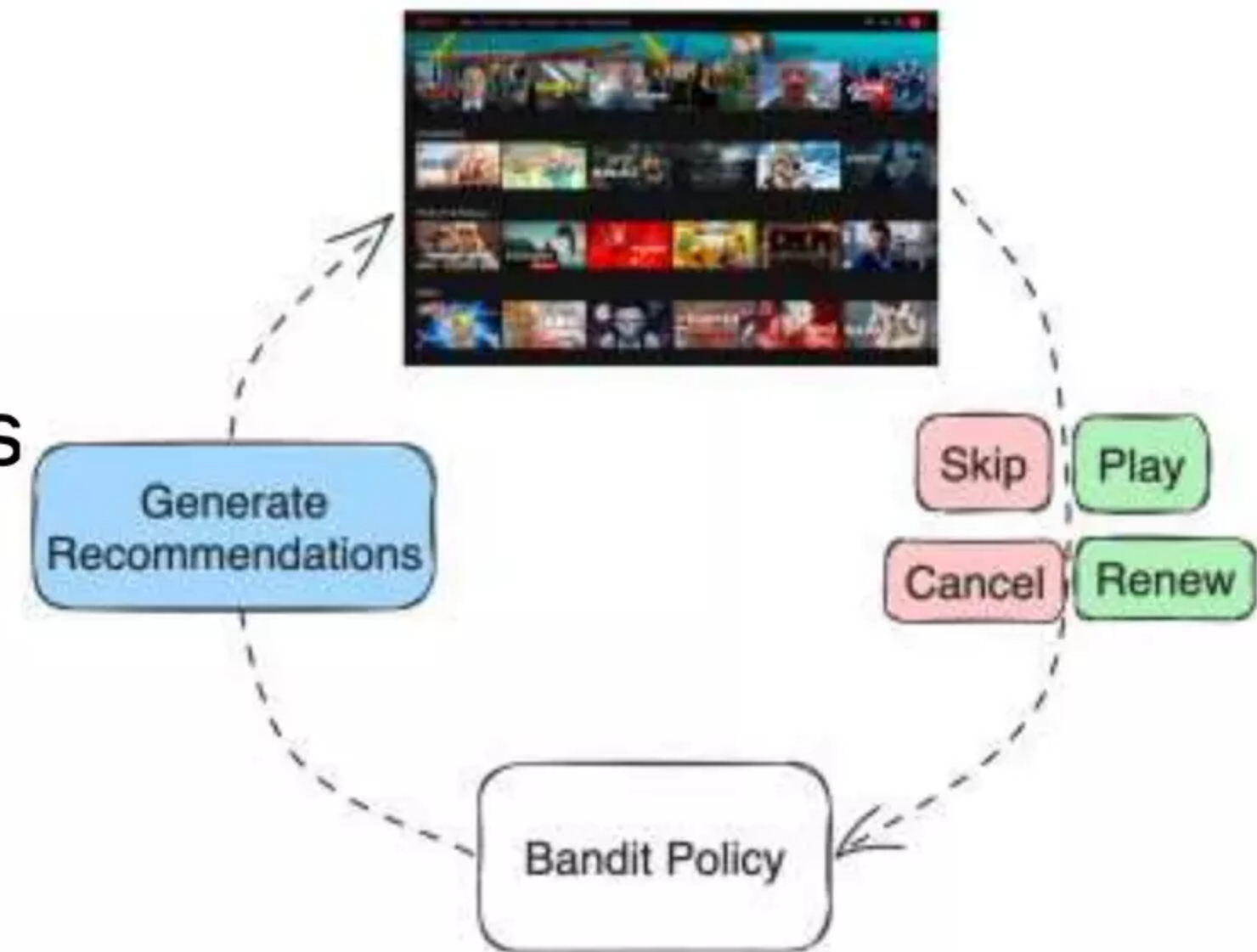# Batch learning from bandit feedback

Production policy

- show recommendations to the member

Member gives feedbacks on recommendations

- immediate: skip/play a show
- long-term: cancel/renew subscription

Goal

- train a policy to maximize the long-term reward

NETFLIX

Batch learning from bandit feedback

Production policy

- show recommendations to the me...

Member gives feedbacks on r... ...ations

- immediate: skip/pl...
- long-term: ...ew subscription

Goal

- train a policy to maximize the long-term reward

Easy: (1) reward = retention (2) = profit 💰!

Generate Recommendations → Skip | Play | Cancel | Renew → Bandit Policy
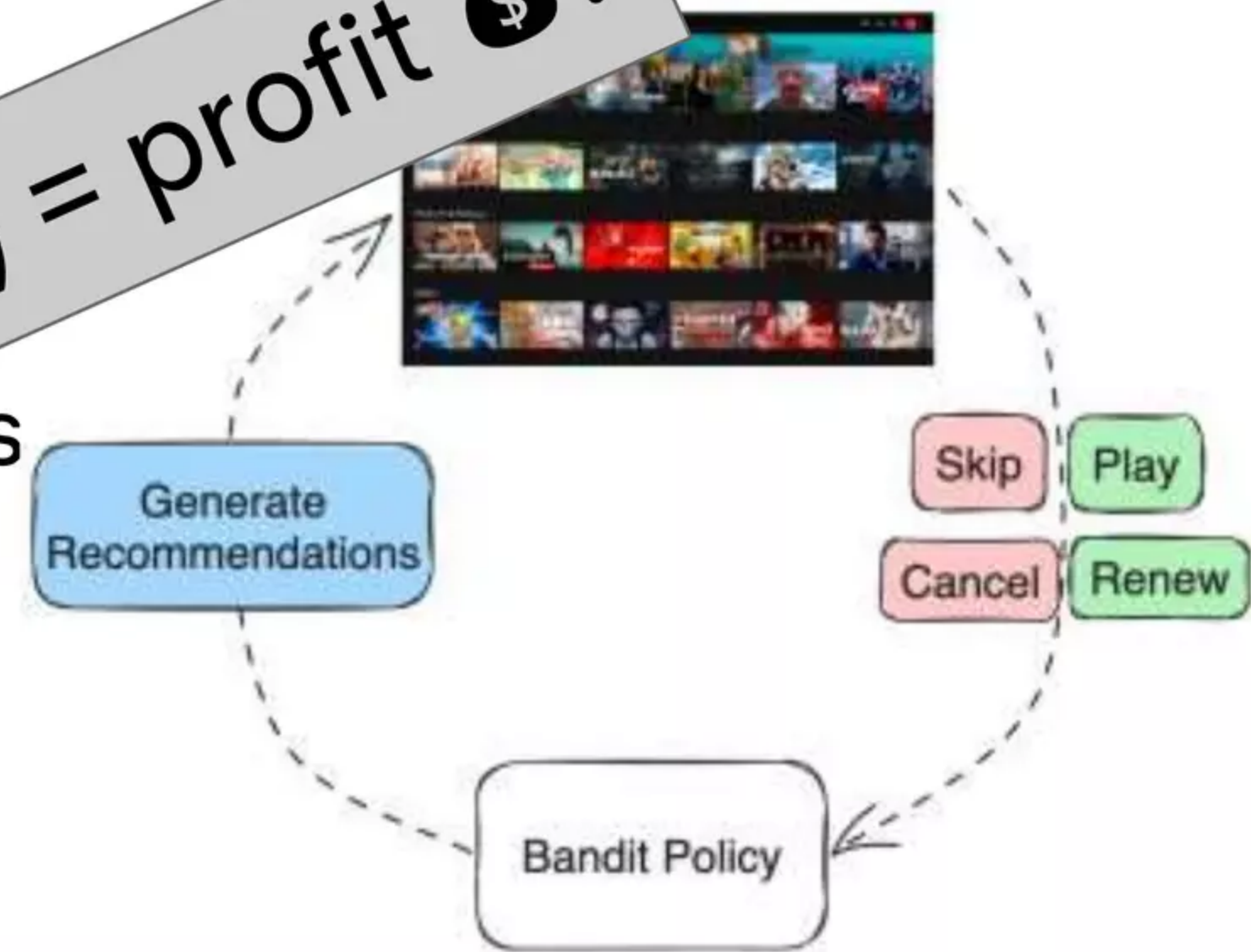
Challenges with long-term retention

Noisy signal	Nonsensitive signal	Delayed signal
Influenced by external factors	Only sensitive for "borderline" members	Need to wait a long time & hard to attribute