

# Эффекты самоорганизации в рекомендательных системах

Дементьев Сергей

Московский физико-технический институт

17 апреля, 2025

- Если не учитывать эффекты самоорганизации в рекомендательных системах, то можно получить деградацию модели
- Также можно получить смещение распределения пользователей – останутся только те, кому нравится именно этот рекомендательный алгоритм. (петля обратной связи)
- Но если мы знаем какие параметры влияют на возникновение петли и каков характер этой связи, то **мы можем контролировать появление самоорганизации в системе.**

- **Существующие компоненты системы:**

- **Алгоритм рекомендаций ( $a_{rec}$ ):** отвечает за формирование персональных рекомендаций товаров пользователям.
- **Алгоритм выбора пользователя ( $a_{choice}$ ):** моделирует, на какие из предложенных рекомендаций пользователь вероятно отреагирует.
- **Генераторы новых пользователей ( $userGAN$ ) и новых товаров ( $itemGAN$ ):** алгоритмы, способный изменять распределения объектов: как удаляют, так и добавляют новых.

- **Динамика распределений:**

- В процессе взаимодействия происходит изменение распределения по некоторому закону, который мы и хотим исследовать.
  - $\mathbf{D}$  – оператор эволюции распределения.
  - Распределение пользователей:  $f_u^t$  и  $\mathbf{D}f_u^t = f_u^{t+1}$
  - Распределение товаров:  $f_i^t$  и  $\mathbf{D}f_i^t = f_i^{t+1}$

На возникновение петли обратной связи в рекомендательной системе влияют следующие ключевые факторы:

- **Характеристики данных:**

- **Размерность данных:** чем выше размерность, тем сложнее может быть обнаружение истинных зависимостей.
- **Наличие шума:** высокий уровень шума может исказить сигналы и приводить к неверным выводам.

- **Функция удовлетворенности:**

- **Тип зависимости:** характер изменения удовлетворенности пользователей от рекомендаций с течением времени (например, линейная, убывающая, насыщающая).

- **Информационная обеспеченность алгоритма:**

- **Объем доступной информации:** количество данных о пользователях, товарах и их взаимодействиях, которое использует алгоритм рекомендаций.

# Как определить, что петля есть?

## Определение:

Пусть  $A_t$  – рекомендации системы, а  $R_t$  – оценки пользователей на шаге  $t$ . Если:

$$P(\{R_s\}_{s=1}^t \mid \{A_s\}_{s=1}^t) \neq \prod_{s=1}^t P(R_s \mid A_s)$$

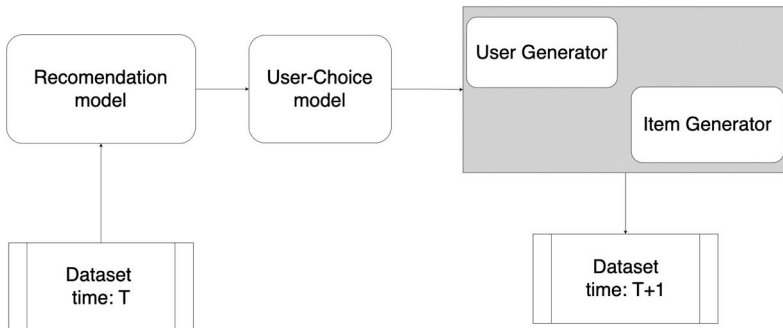
Тогда мы считаем, что в системе возникает петля обратной связи.

## Численный критерий возникновения петли:

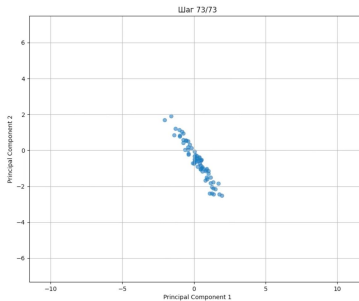
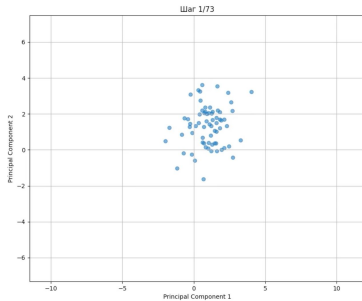
Пусть  $F, \mathcal{L}$  – функционал качества и функция Лосса

$\exists t_0, \Delta \in \mathbb{N} : \quad \mathcal{L}(R_{t+1}, A_{t+1}) < \mathcal{L}(R_t, A_t)$ , а также  
 $F(R_{t+1}, A_{t+1}) < F(R_t, A_t) \quad \forall t \in \overline{t_0, \dots, t_0 + \Delta - 1}$

# Модель

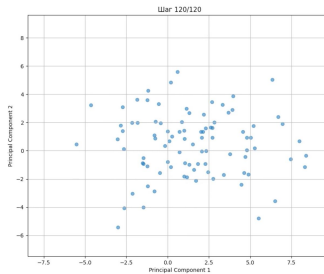
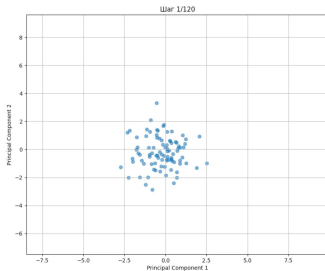


# Эксперимент и его результаты



Скрытая петля положительной обратной связи

# Эксперимент и его результаты



Скрытая петля отрицательной обратной связи



# Анализируем результаты

- В ходе экспериментов была получена явная зависимость от времени, которую если не учитывать, то система будет либо сходиться к стационарной, либо будет происходить data drift только малой части данных
- От количества информации в данных зависит возникновение петли. Это легко получить, используя разные размеры эмбедингов для модели рекомендации, а также изменяя преобразования эмбедингов

- Мы получили качественное правило появления петли обратной связи
- Смогли получить петлю на реальных данных
- в дальнейшем будем исследовать сходимость многомерных распределений, а также рассматривать  $KL$ -дивергенцию между соседними распределениями

# Список литературы



Wenlong Sun et al. *Debiasing the Human-Recommender System Feedback Loop in Collaborative Filtering* , 2019.



Karl Krauth et al. *Breaking Feedback Loops in Recommender Systems with Causal Inference*, 2022.



Anton Khritankov. *Positive feedback loops lead to concept drift in machine learning systems*, 2023.