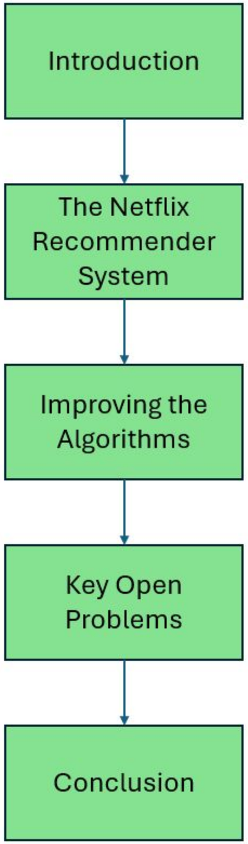
The Netflix Recommender System: Algorithms, Business Value, and Innovation

Authors: Carlos A. Gomez-Uribe and Neil Hunt

Presented by: Shelah Ameli, Thomas Fillers and William Parham
COSC 540: Advanced Software Engineering
March 18th, 2024

Outline

- Introduction
- The Netflix Recommender System
- Improving the Algorithms
- Key Open Problems
- Conclusion



Introduction

- Storytelling is a key factor of human interaction
- Television gave way to the widespread use of of video for storytelling
- Now, with the entrance of the internet, Netflix aims to be the bridge between the Internet and storytelling.
- With millions of users and countless more hours of streamed content, there is a constant need for innovation when recommending new content
- Netflix's recommender system consists of many unique algorithms for several use cases that all work together to provide the complete experience

The Netflix Recommender System

The Recommender Problem: Internet TV is about choice. Netflix defines the recommender problem as making sure their customers find something compelling to view within 60 to 90 seconds

The different algorithms that make up the Recommender System.

- **Personalized Video Ranker (PVR)**:
  - Orders the entire catalog of videos for each member's profile in a personalized way.
  - Works better when personalized signals are blended with a dose of unpersonalized popularity.
- **Top-N Video Ranker**:
  - Finds the best few personalized recommendations in the entire catalog for each member.
  - It focuses only on the head of the catalog ranking unlike PVR which ranks arbitrary subsets of the catalog.
  - Similarities with PVR includes combining personalization with popularity, incorporating trends over different time windows.

The Netflix Recommender System (Cont.)

- ❑ **Trending Now**:
  - ❑ Focuses on shorter-term temporal trends, combined with a level of personalization.
  - ❑ It usually identifies trends that repeat yearly (e.g Christmas movies), or one-off short-term events.
- ❑ **Continue Watching**:
  - ❑ Responsible for episodic content as well as non episodic content viewed in small bites.
  - ❑ Sorts the subset of recently seen videos based on an estimate of whether the member intends to resume watching or rewatch, or whether the member has abandoned the video.
- ❑ **Video-Video Similarity (Sims)**:
  - ❑ Uses the SIMS unpersonalized algorithm that computes a ranked list of videos - the similars - for every video in the catalog.
  - ❑ The "*Because You Watched* ... ' category is created using the SIMS algorithm with a touch of personalization.
- ❑ **Page Generation: Row Selection and Ranking**:
  - ❑ Makes use of the output of all previously discussed algorithms to construct every single page of recommendations. It takes into account the relevance of each row to the member as well as the diversity of the page.

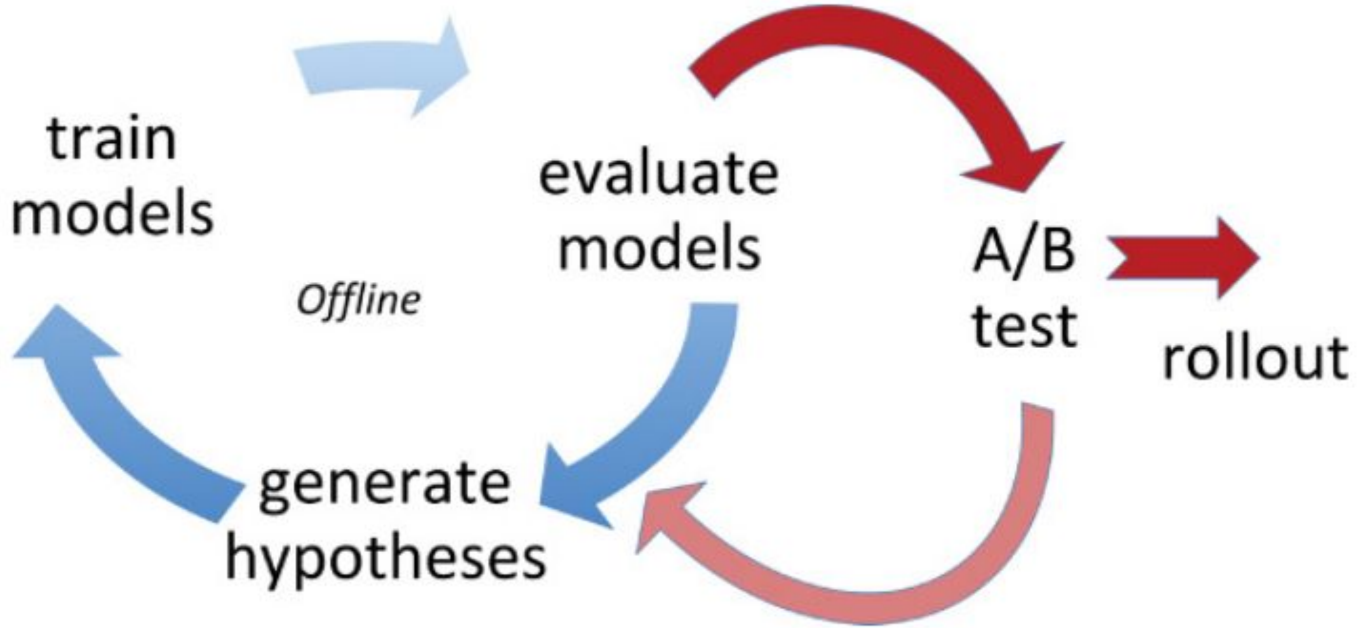
The Netflix Recommender System (Cont.)

- **Evidence Selection**:
  - Works together with the recommendation algorithms to define the Netflix experience.
  - Evidence includes information such as the predicted star rating; the synopsis; facts about the video (any awards, cast, or other metadata); and the images.
  - Evidence selection algorithms evaluate all the possible evidence items that can be displayed for every recommendation, to select the few that will be most helpful to the member viewing the recommendation.
- **Search**:
  - The search experience is built on several algorithms.
    - One algorithm attempts to find the videos that match a given query.
    - Another algorithm predicts interest in a concept given a partial query.
    - A third algorithm finds video recommendations for a given concept.
  - The search algorithms combine play data, search data, and metadata to arrive at the results and recommendations that we offer.

Improving the Algorithms

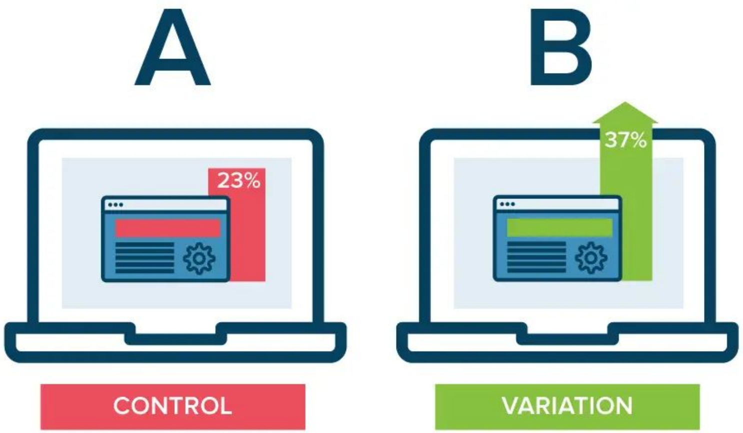
- **A/B Testing:**
  - Randomized, controlled experiments which compare retention rates among different algorithmic variants
  - While results are usually informative, certain results may require additional validation or testing due to some random variation
- **Alternative Metrics**
  - Time to first play, sessions without plays, days with a play, abandoned plays, etc
- **Test Audience**
  - Existing members
    - Large sample size
    - Behavior can be influenced by previous experiences
  - New members
    - Have not experienced previous versions
    - Fewer new members + one-month-free trial leads to less cancellations
- **Offline Experimentation**
  - Experimentation based on historical data
  - Allows for quick iteration and candidate pruning
  - Major drawback: assumes members would behave the same way (playing same content), if the new algorithm had been used to generate recommendations
  - Need to interpret how different the tested algorithm is from the production algorithm
- **Estimating Word-of-Mouth Effects**
  - Hard to measure
  - If a change leads to retaining more members in a period of time, it may be expected to generate a comparable magnitude of new member via word-of-mouth

Improving the Algorithms



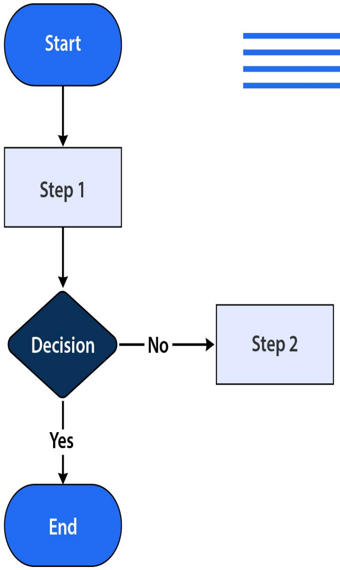
Key Open Problems

- A/B testing
  - Better experimentation protocols
    - Develop new offline experiment metrics
  - General improvement to A/B testing methods
    - Effective variance reduction methods
  - New A/B engagement metrics that are more correlated to retention rates
    - How can we balance long and short form content?



Key Open Problems (Cont.)

- Recommendation Algorithms
  - Global Algorithms
    - How can we improve suggestions for smaller countries while not affecting larger countri
      - How can we generalize mathematical tools and techniques to reflect different u accessing different catalogues.
    - How can we improve suggestions with language in mind?
      - We should not recommend an item in which a user has no common language ar translation exists.
  - Controlling Presentation Bias
    - How do we find member groups that respond similarly to different recommendations?
    - How can we introduce randomness to recommendations in order to learn better models?
  - Page Construction
    - How can we construct an optimal recommendations page that engages with users?
  - Member Coldstarting
    - What is the best way to recommend media to new users?
  - Account Sharing
    - What needs to change to accurately recommend videos on shared accounts?
    - Children’s viewing habits and content are different than that of adults.
  - Choosing the Best Evidence to Support Each Recommendation
    - What is the best format to present our recommendations?
      - Show actors, synopsis, similar media, etc.



Conclusion

- The Netflix Recommender system is a collection of algorithms.
- Humans face an ever increasing amount of choices making all presented information more valuable thus an optimal recommendation system is becoming increasingly valuable.
- The field of recommender systems will continue to play a pivotal role as the wealth of data increases.
- Recommender systems can make long-tail products, services, and information more accessible for everyone.