SECOND PRESENTATION

16102284 Lee SungHo

INDEX

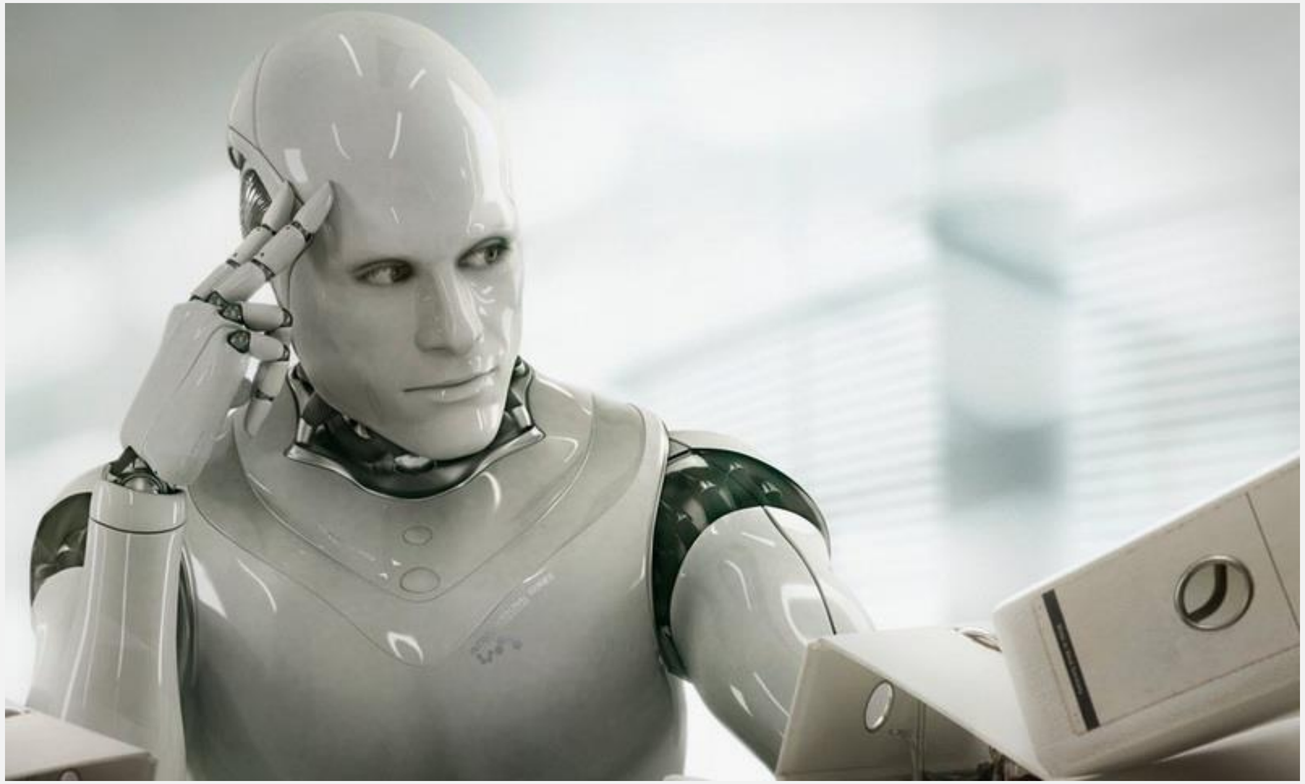
1. Introduction

Introduction

Introduction

Introduction

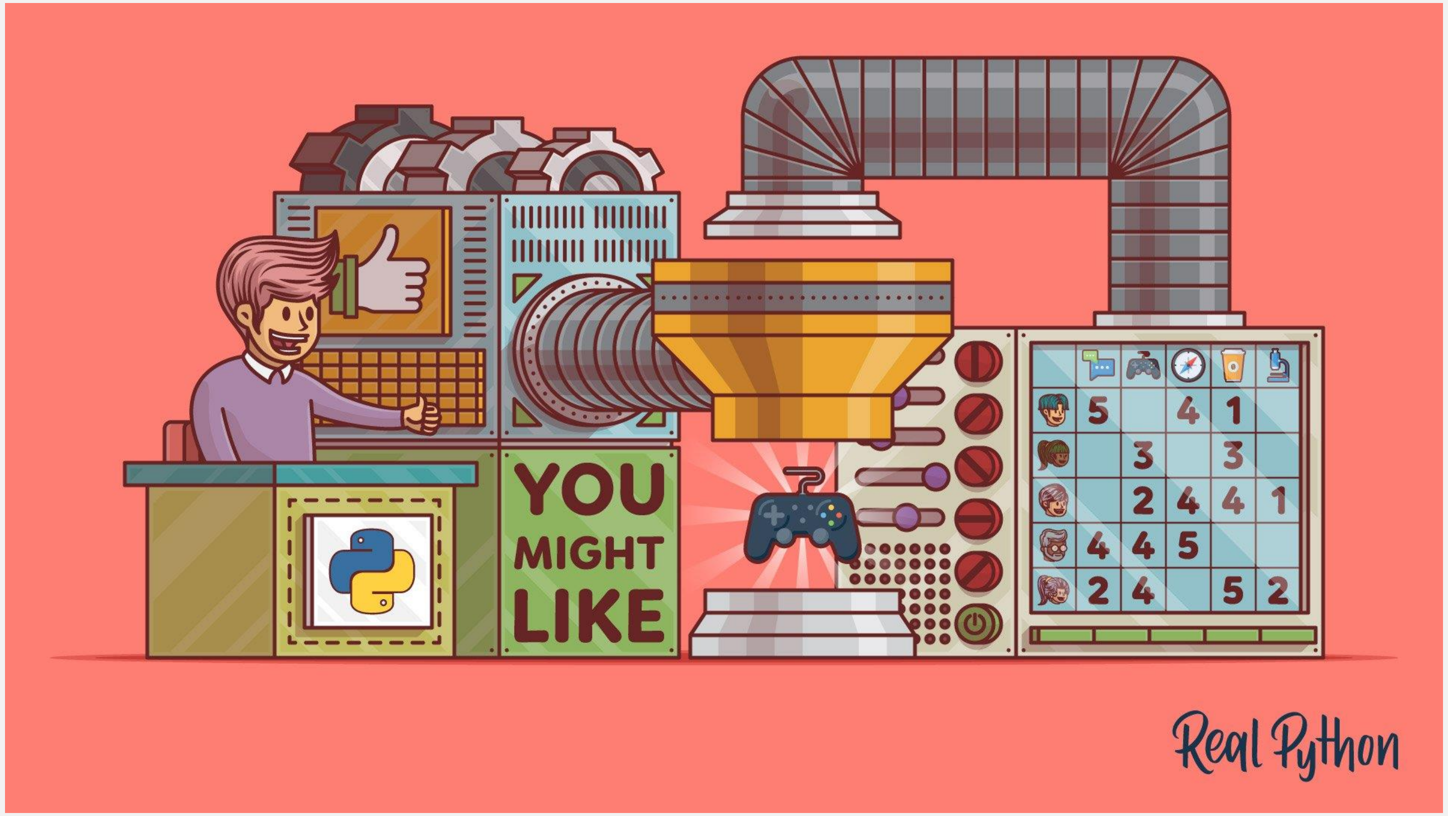
Algorithm: A procedure or rule set for solving a problem



Introduction

Algorithm? Recommend Algorithm!

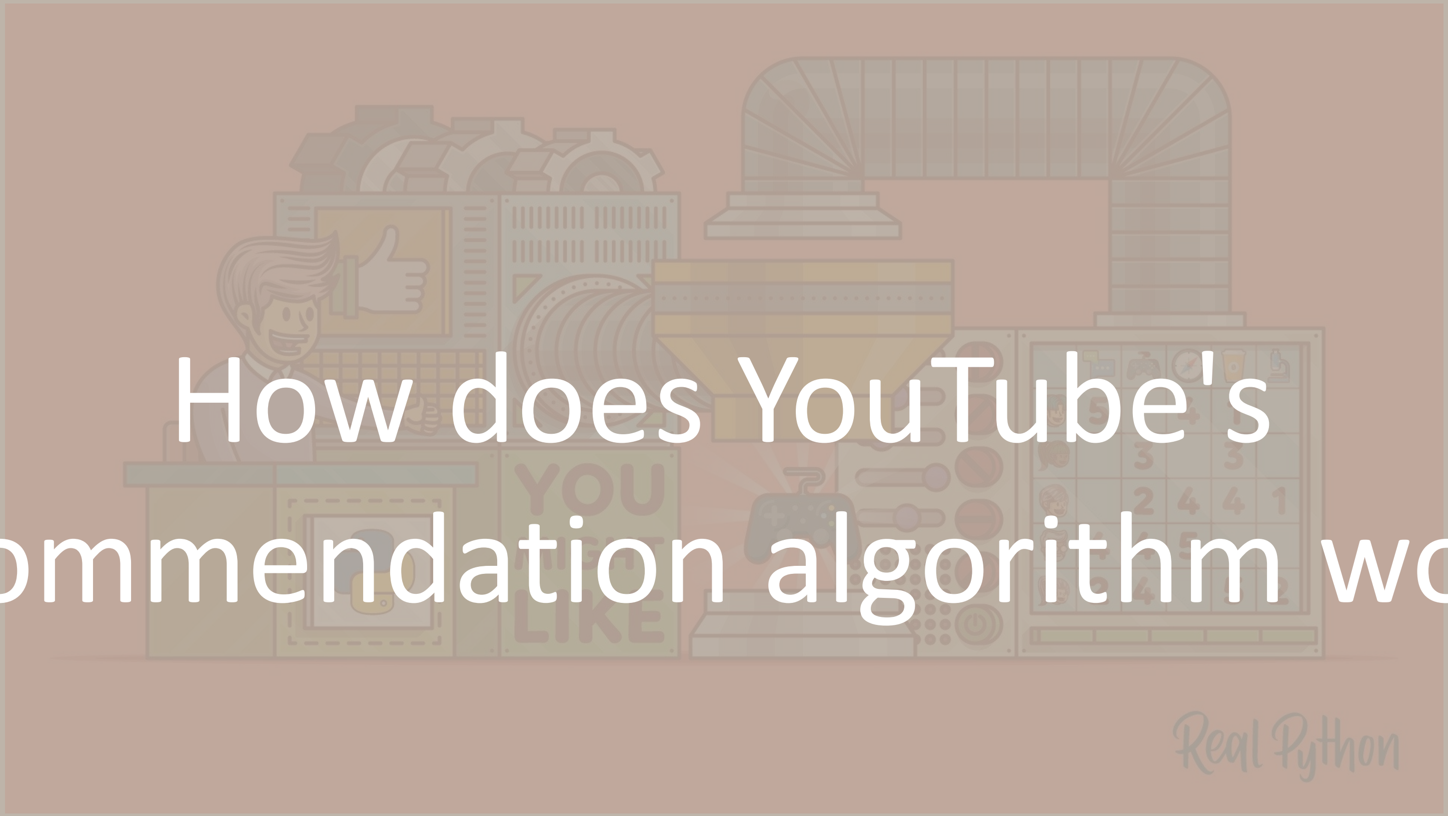
Introduction



Recommendation algorithm : It refers to predicting what users would prefer among items (ex. Movies) that they have not yet consumed through data.

How does YouTube's recommendation algorithm work?

Introduction



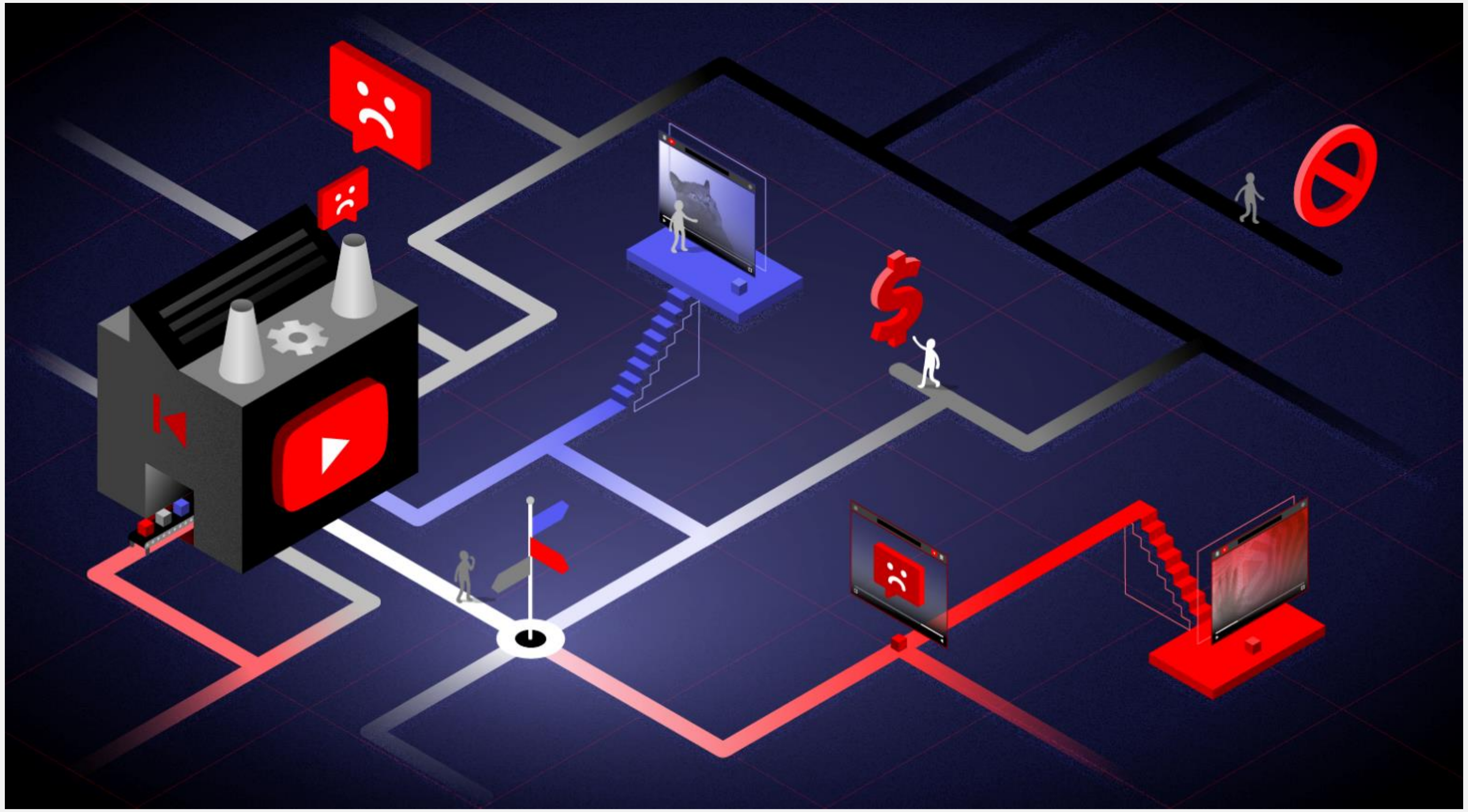
Recommendation algorithm : It refers to predicting what users would prefer among items (ex. Movies) that they have not yet consumed through data.

2. Youtube Recommend DL Algorithm

2. Youtube Recommend DL Algorithm

※ A scene from the presentation of a paper taken by Professor Peter Brusilovsky of the University of Pittsburgh at ACM RecSys

2. Youtube Recommend DL Algorithm



“”YouTube uses the information of the video you are currently watching to find similar videos and analyzes data from various people to find out what people like me might want to see at this time.””

2. Youtube Recommend DL Algorithm

Factors used in recommendations

방송인 유재석이 아들 지호와 유튜브 알고리즘으로 다투고 있다고 밝혔다.

1. ***Data related to images and users or contexts.***

2. ***Data that can be set as goals to predict user behavior***

2. Youtube Recommend DL Algorithm

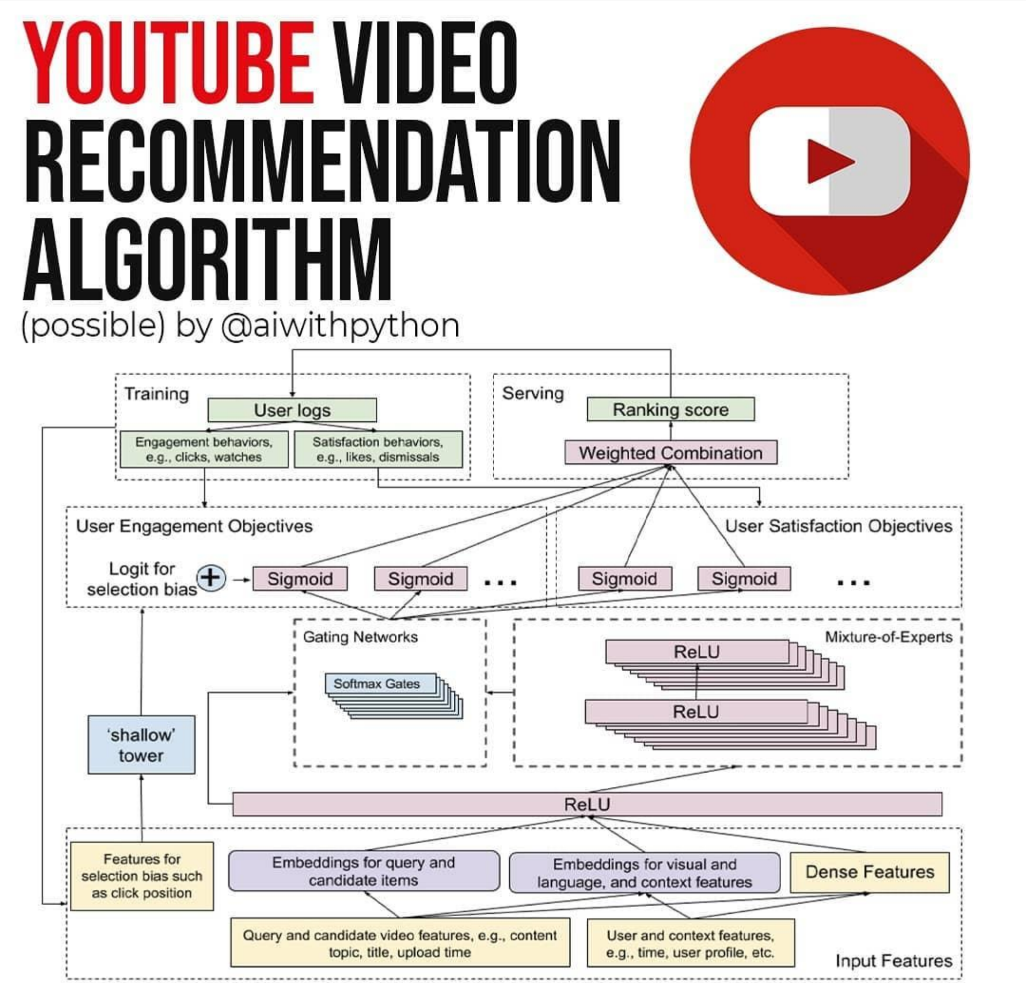


Fig. Complete architecture of the model

2. Youtube Recommend DL Algorithm

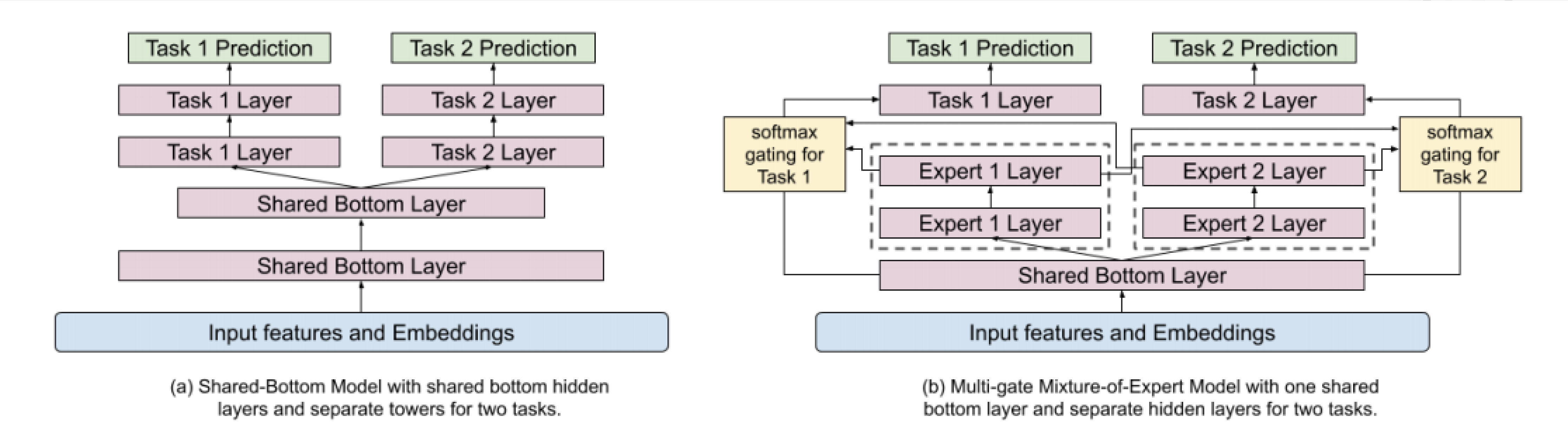


Fig. Replacing shared-bottom layers with MMoE

2. Recommend Algorithm - Youtube

How to troubleshoot an existing recommendation system

Check point 1 : Recommendation Goal Setting

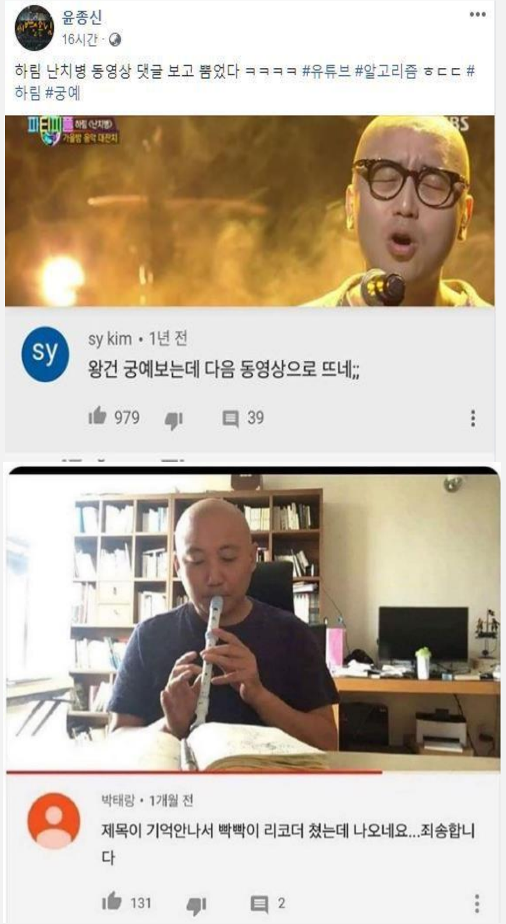
- YouTube categorizes goals into engagement and satisfaction, and uses machine learning to predict multiple goals without conflict.

Check point 2 : Bias for top recommendations

- YouTube calculates and predicts the degree to which bias affects the popularity of a video.

2. Youtube Recommend DL Algorithm

Remaining problems with recommender systems



2. Youtube Recommend DL Algorithm

Remaining problems with recommender systems

Problem 1 : Difficult to handle due to dynamic change

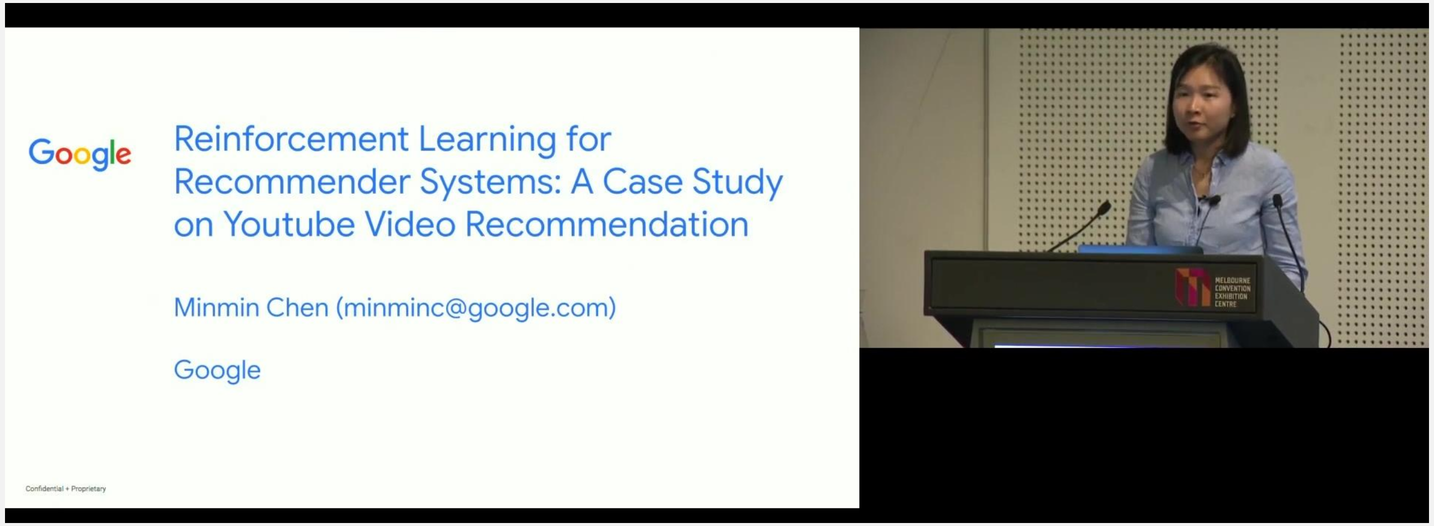
- YouTube has over a billion users and billions of videos are played and upload every day. YouTube has to use so much data to do machine learning and use it to provide recommended videos to users in real time.

Problem 2 : Myopic Recommendation(unable to see distant objects clearly)

- The duration of popularity of a video varies from video to video. Also, the user's interest changes with time. Also there is a tendency to recommend similar items to users, and this tendency reduces users' interest in similar topics.

3. Youtube Recommend RL Algorithm

2. Youtube Recommend RL Algorithm



https://www.youtube.com/watch?v=HEqQ2_1XRTs

User Response Models to Improve a REINFORCE Recommender System

Minmin Chen, Bo Chang, Can Xu, Ed H. Chi
Google, Inc.
Mountain View, CA
minminc,bochang,canxu,edchi@google.com

2. Youtube Recommend RL Algorithm

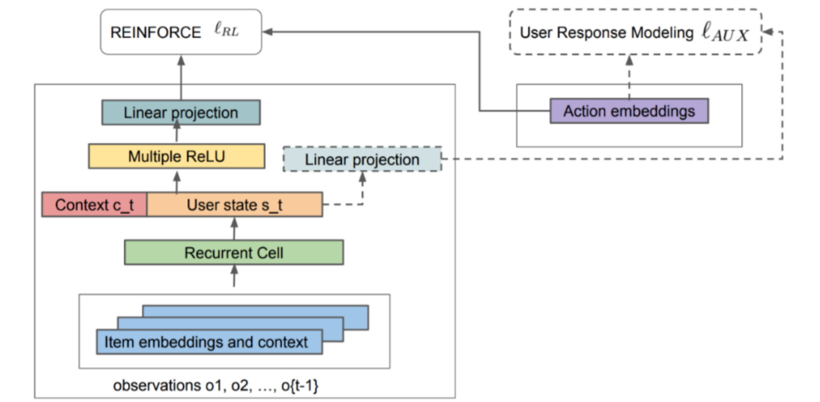
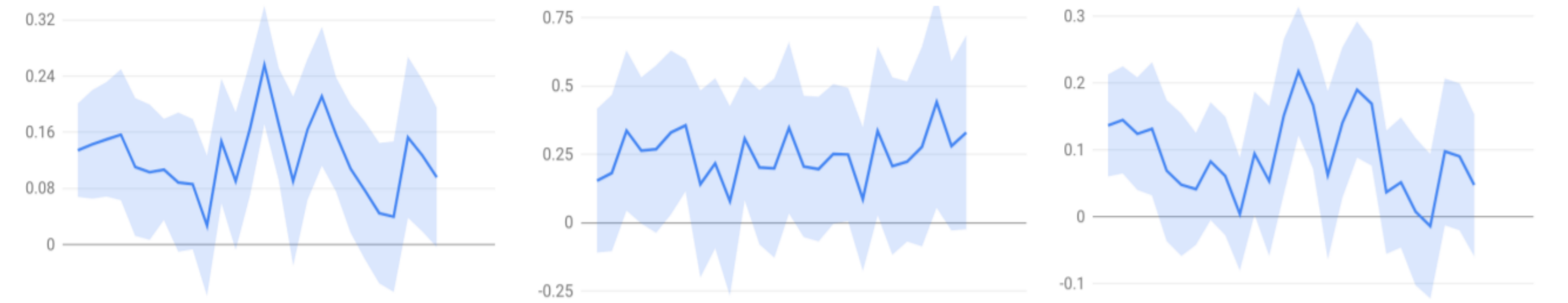


Figure 2: User response models to augment the training of the REINFORCE agent.



(a) Global user base (b) Low-activity users (c) High-activity users

Figure 9: Comparison of the base REINFORCE agent and URL in a live system.

4. Bing news Recommend

RL Algorithm

4. Bing news Recommend RL Algorithm

Bing news

“” In Microsoft Bing, a recommendation algorithm was implemented using DQN for news recommendation..””

4. Bing news Recommend RL Algorithm

DRN: A Deep Reinforcement Learning Framework for News Recommendation

Guanjie Zheng[†], Fuzheng Zhang[§], Zihan Zheng[§], Yang Xiang[§]

Nicholas Jing Yuan[§], Xing Xie[§], Zhenhui Li[†]

Pennsylvania State University[†], Microsoft Research Asia[§]
University Park, USA[†], Beijing, China[§]
gjz5038@ist.psu.edu,{fuzzhang,v-zihanzhe,yaxian,nicholas.yuan,xingx}@microsoft.com,jessieli@ist.psu.edu

ABSTRACT

In this paper, we propose a novel Deep Reinforcement Learning framework for news recommendation. Online personalized news recommendation is a highly challenging problem due to the dynamic nature of news features and user preferences. Although some online recommendation models have been proposed to address the dynamic nature of news recommendation, these methods have three major issues. First, they only try to model current reward (e.g., Click Through Rate). Second, very few studies consider to use user feedback other than click / no click labels (e.g., how frequent user returns) to help improve recommendation. Third, these methods tend to keep recommending similar news to users, which may cause users to get bored. Therefore, to address the aforementioned challenges, we propose a Deep Q-Learning based recommendation framework, which can model future reward explicitly. We further consider user return pattern as a supplement to click / no click label in order to capture more user feedback information. In addition,

34], and hybrid methods [12, 24, 25]. Recently, as an extension and integration of previous methods, deep learning models [8, 45, 52] have become the new state-of-art methods due to its capability of modeling complex user item (i.e., news) interactions. However, these methods can not effectively address the following three challenges in news recommendation.

First, the dynamic changes in news recommendations are difficult to handle. The dynamic change of news recommendation can be shown in two folds. First, news become outdated very fast. In our dataset, the average time between the time that one piece of news is published and the time of its last click is 4.1 hours. Therefore, news features and news candidate set are changing rapidly. Second, users' interest on different news might evolve during time. For instance, Figure 1 displays the categories of news that one user has read in 10 weeks. During the first few weeks, this user prefers to read about "Politics" (green bar in Figure 1), but his interest gradually moves to "Entertainment" (purple bar in Figure 1) and "Technology"

4. Bing news Recommend RL Algorithm

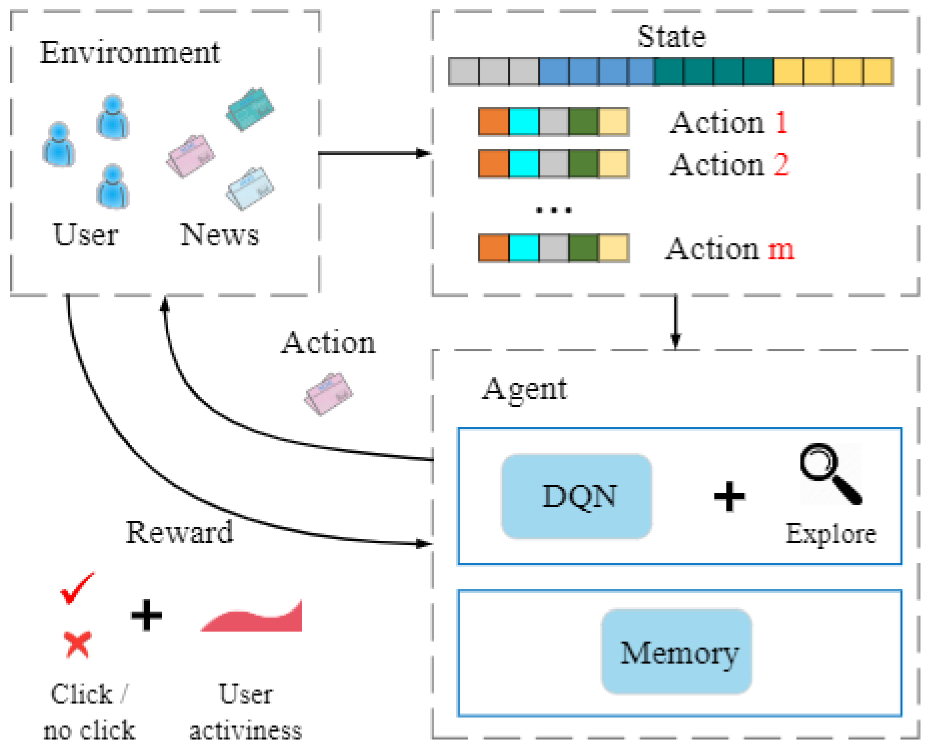


Figure 2: Deep Reinforcement Recommendation System

1. *The present and future rewards are considered simultaneously using the 'DQN (Deep Q-Learning)' framework.*

2. *In the existing click/no click reward method, feedback called 'activeness' is added that considers how often users return to the main page after receiving one recommendation.*

3. *An exploration strategy called 'Dueling Bandit Gradient Descent' is applied to the model to improve the diversity of recommendations.*

4. Bing news Recommend RL Algorithm

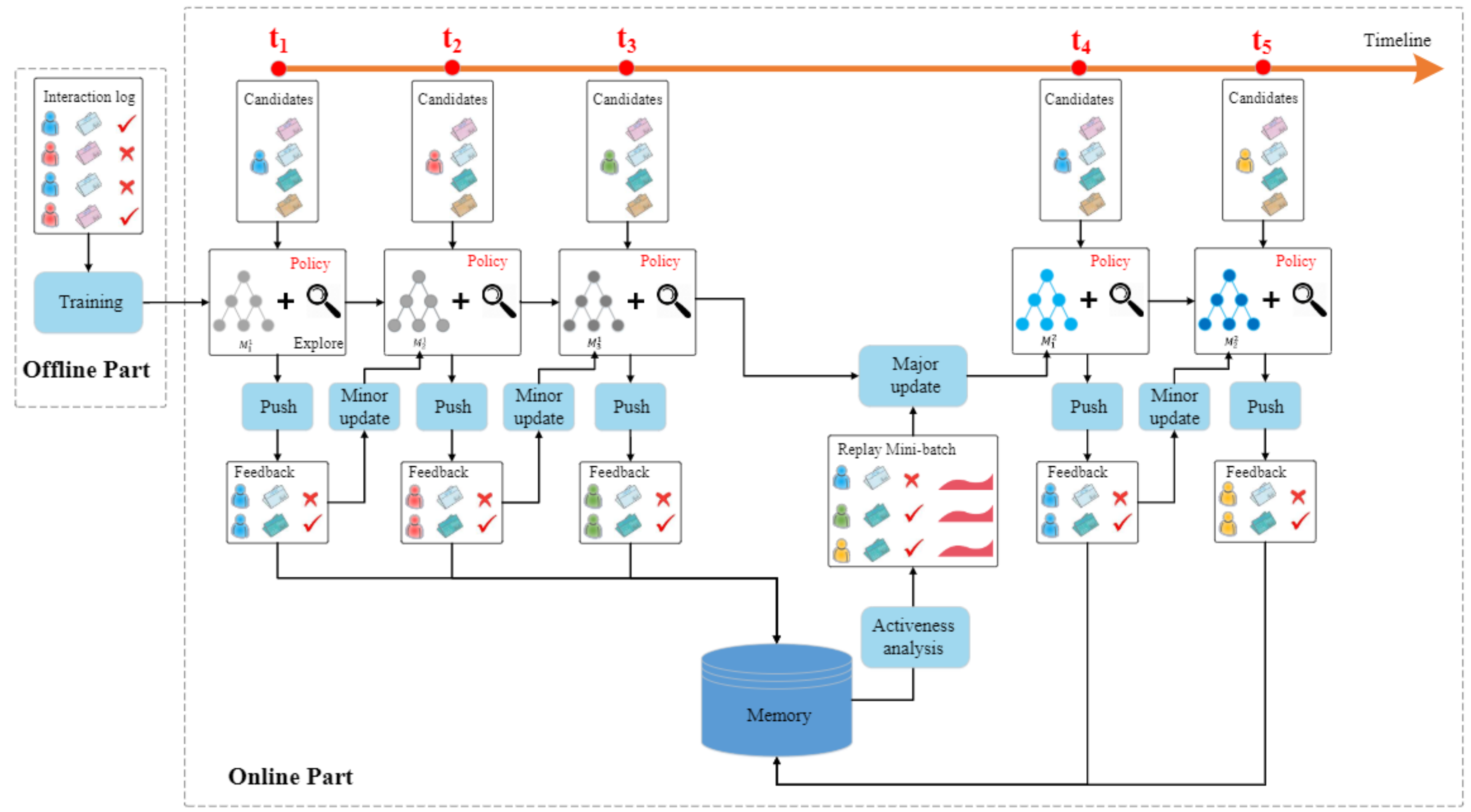


Figure 3: Model framework

4. Bing news Recommend RL Algorithm

- **Offline part(pretrained)**
  - The study learned from the offline part with user-news click logs and extracted features of four types (News, User, User News, Context).
    - Learn based on log information clicked/no by the user.
    - The offline part only deals with static data. Therefore, not only is there a limitation in expanding, but it is also difficult to grasp user activity over time.

- **Online Part(Real Time)**
  - **PUSH**: With the value entered by the agent, the current user's features and news candidates (articles with high similarity to the currently recommended list are randomly selected) and top-k articles are recommended. (Article list created considering both exploitation and exploration L)
  - **FEEDBACK**: The user clicks or no clicks each item in the article list (L)
  - **MINOR UPDATE**: At each timestamp, if the exploration network performs better, we update the current network of the next policy to explore, and if the exploitation network performs better, we leave it alone.
  - **MAJOR UPDATE**: After a certain amount of time has elapsed, user feedback and user activeness stored in memory are added. (agent keeps record of recent clicks and activeness)

5. Conclusion

5. Conclusion