Weekly Report   October 25, 2021

Ali Jandaghi Alaee   Report No. 13

alialaee98@gmail.com   00/08/02- 00/07/26

Abstract

- Reading the RL paper [2]
- Reading the HCRF paper [1]
- Getting information about AOL dataset

Description

- This paper [2] proposes a novel video recommendation system for YouTube based on RL. The environment is based on the Markov decision process. The goal is to extract about hundreds of videos according to the user's actions and states out of the large corpus of videos (billions of videos). The major contribution of this paper can be listed as below;

    - Deploying a massive RL model for video recommendation among billions of videos
    - Addressing the off-privacy recommendation issue; removing the effect of imposed materials that have been suggested to user by other recommender systems
    - Extending the reinforcement learning model to suggest the K top actions instead of just one.

  It is essential to remind that this method is the state-of-the-art recommender system that is currently running in the YouTube, therefore, its implementation might be difficult due to our limited access to proper dataset and the complexity of the methods.

- The main idea of the second paper [1], is that the user sessions can be divided into two groups; browsing sessions and clicker sessions. In the first one users are not interested in seeing ads and will not probably click on them. On the other hand, in the clicker sessions, users tend to look for ads and might purchase an item. Clearly, a clicker session is a better choice for showing ads, thus, the goal is to find an effective way to predict wether a session is clicker or browser. The authors claim that an embedding space of user actions followed by a classifier (HCRF recommended, but other methods like logistic regression have been tested as well) can help with the classification of these sessions. The source code of this HCRF model seem to be available in a link on the paper. They have implemented a directed version of skip-gram in which unlike the regular method, each action is only associated to its previous actions (Not its previous and next actions). Training data is composed of users' sessions labeled with their category (browser and clicker), and the sequence of embeddings of a new user session as the input of the HCRF model determines the class of its session.

- The results of the analysis of the AOL dataset are illustrated in the table below;

Dimensions	(3,614,506, 5)
Unique ID	66,000
Unique Query	1,244,495
Unique URL	389403
Entries with no following clicks	1,678,893

Next Week

1. A deeper study on [1] for details.

2. Prepare a convenient dataset for the purpose of remaining tasks.

References

[1] Mihajlo Grbovic et al. Hidden conditional random fields with distributed user embeddings for ad targeting. *SIGIR '21, July 11–15, 2021, Virtual Event, Canada.*, 2021.

[2] Minmin Chen et al. Top-k off-policy correction for a reinforce recommender system. *SIGIR '21, July 11–15, 2021, Virtual Event, Canada.*, 2021.