CSCE768 Course Project
SPRING 2024

Title: eCommerce Recommender System

Team Members:

1. Venkata Naga Sri Sai Pranavi Kolipaka
2. Sai Krishna Revanth Vuruma

Abstract: Recommender Systems (or RecSys as they are widely known) can be found in many modern day applications ranging from e-commerce websites like Amazon to music platforms like Spotify. Through this work, we wish to develop a Deep Learning model that can leverage user behavioral data to make similar or relevant product recommendations for the user to consider. We will use a publicly available dataset [1] to train our model and then evaluate it on popular evaluation metrics [4] used to quantify recommender system performance.

Background: A recommender system leverages machine learning algorithms to give users a list of items that are relevant to the item that they are currently looking for. Popular systems also use behavioral data from other users that have shown similar tendencies as the current user to make more personalized recommendations. They use different criteria such as past purchases, demographic information and behavioral data among others.

Model: BERT4Rec [3] is a variant of the popular BERT model that is trained and designed for Recommendation Systems. Building on the strong points for the original BERT model such as the multi-headed self-attention layer, BERT4Rec adopts transformers to sequential recommendation tasks. Thanks to the self-attention mechanism, BERT4Rec can directly capture dependencies in any distances. In contrast to other recommender systems, BERT4Rec uses bidirectional self-attention to model users' behavior sequences to improve quality of recommendations. The model architecture can be found in the figure below.

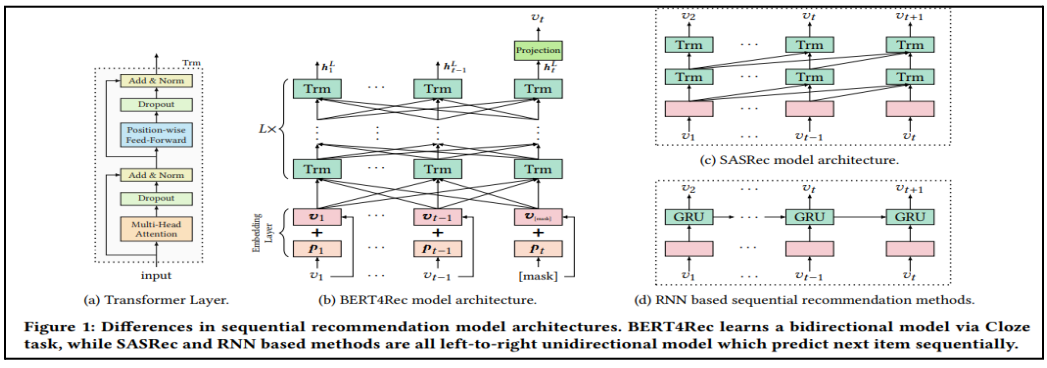

Figure 1: Differences in sequential recommendation model architectures. BERT4Rec learns a bidirectional model via Cloze task, while SASRec and RNN based methods are all left-to-right unidirectional model which predict next item sequentially.

Source: [3]

The BERT4Rec model performed best on benchmark datasets such as Amazon Beauty, Steam and MovieLens with a performance increment of 11.03% NDCG@10 against the strongest baselines [3].

Dataset: The dataset is made publicly available on Kaggle by the REES46 Marketing Platform [2]. It contains behavioral data of users from a large eCommerce store. The data was collected over a period of seven months from Oct 2019 - Apr 2020 spanning around 285 million user events on the platform. The schema consists of 9 columns with the following information:

- Event: timestamp, type
- Product: id, category, category_code, brand, price
- User: user_id, user_session_id

There are four possible event types: view, cart, removed_from_cart and purchase that indicate how the user has interacted with the product in this session.

Evaluation: For evaluating the performance of our recommender system, we will use popular metrics [3] in the space such as:

1. **Precision @K, Recall @K:** Similar to how precision and recall are usually calculated, but adjusted for the top-K number of predictions made by the model.
2. **Pearson Coefficient:** Used to calculate similarity between two sets of data such as users or products.
3. **Normalized Discounted Cumulative Gain:** Also known as NDCG, this metric can be used to calculate a cumulative score of an ordered set of items. Items of higher relevance appearing at lower ranks are penalized thus resulting in a more grounded score.

While evaluating recommender systems it is important that we don't look to calculate scores on individual predictions as that could skew the model. Instead, we consider the set of predictions as a whole.

References:

1. Kechinov, Michael, "eCommerce behavior data from multi category store", Kaggle, accessed on March 09, 2024, https://www.kaggle.com/datasets/mkechinov/ecommerce-behavior-data-from-multi-category-store.
2. REES46, "About REES46", REES46, accessed March 10, 2024, https://rees46.com/.
3. Sun, Fei, et al. "BERT4Rec: Sequential recommendation with bidirectional encoder representations from transformer." Proceedings of the 28th ACM international conference on information and knowledge management. 2019.
4. Bhattacharyya, Mayuk, "Metrics of Recommender Systems: An Overview", Towards Data Science, accessed March 10, 2024, https://towardsdatascience.com/metrics-of-recommender-systems-cde64042127a.

5. Shani, Guy, and Asela Gunawardana. "Evaluating recommendation systems." Recommender systems handbook (2011): 257-297.