Northeastern University, Khoury College of Computer Science

CS 6220 Data Mining | Assignment 6

Due: February 29, 2023

Haoping Lin, Yang Yao
ConnrLin, stervt
lin.haop@northeastern.edu,
yao.yan@northeastern.edu

Recommendation System On Transformer

Problem

The problem we want to discuss and solve is when a user adds some items into an online shopping cart, what other items should be recommended to this user next. Using classical recommendation methods, such as Matrix Factorization, does not provide an ideal prediction rate, since they do not take potential sequential patterns of users' behavior into consideration. This may lead to ignoration of hidden features and thus low accuracy.

Background

In early recommendation systems, Collaboration Filter is used to measure users' interest. For example, Matrix Factorization is a common model of user's behavior. However, these models may ignore possible sequence patterns behind and thus leading to a low predicting accuracy. In recent years, with rapid exploration and gradually improved models in deep learning, more and more Recommendation Systems manage to implement neural networks to strengthen ability in analyzing users' behavior, since these deep learning models are able to find potential features based on time sequence. For example, a Recurrent Neural Network encodes previous behaviors into vectors, which will be considered in predicting future behavior. Hence, we are going to implement a recommendation model based on Transformer, which considers not only the sequence from past to now, but also from

future to now. Theoretically, models using Transformer perform better than ones using RNN.

Models implementing Transformers demonstrate the relations between one behavior with behaviors before and after, leading to a more logical analysis in proper cases. For example, a user who buys a phone case is more likely to be someone who just bought a phone, and maybe in the near future, this user will also buy a screen protector. In this case, the sequence of items the user bought should be considered as a key feature to understand the user's behavior. By implementing the attention mechanism behind the Transformer, we believe that the recommendation system should make a wiser choice based on complicated situations, especially situations demonstrating sequence.

Approach

We are going to use deep neural networks to mine latent information. In more detail, we plan to use Transformer as a recommender model. Recommendation tasks can be seen as given a sequence, predict the next behavior, which is very similar to Cloze task. Colze task is similar to gap-filling task, which randomly masks some items in a sequence and asks the model to predict the contents that are being masked. This task is used in training BERT(Bidirectional Encoder Representations from Transformers), so we think we could use BERT for recommendation tasks. For example, given a sequence [b1, b2, b3,...,bn], we can feed [b1,b2, b3,..., bn, [MASK]] to the model, and we can gain the prediction of [MASK] which is the recommended item. In addition, we can change the number and the position of the mask to gain different recommendations. This idea comes from the paper “BERT4Rec: Sequential Recommendation with Bidirectional Encoder Representations from Transformer” and we are going to follow the steps in the paper to build and train a model using our datasets.

We plan to use MovieLens-20m dataset, a dataset containing 20 million ratings and 465,000 tag applications applied to 27,000 movies by 138,000 users. We plan to use this movie dataset to train our movie recommender model. More specifically, we will only use users' browser history as data to train our model, other information like rating and movie information will be deprecated since these features have a loose relation with training our model.