ICT 3299 Data Mining and Predictive Analysis

FISAC Assignment- 3 Minute Thesis Report

Paper Title- Cross Domain Life Long Sequential Modelling for Online Click-Through Rate Prediction

Adithi NR – 220953009

9th April, 2025

Motivation and Problem Statement

One of the fundamental tasks for increasing user interaction with an ad or item is accurately predicting the Click-Through Rate (CTR). With the exponential growth of ads and products shown to users, CTR prediction has become increasingly complex. This complexity arises primarily due to the larger volume and greater diversity of content, which gives users more varied options and makes their behavior harder to predict.

The challenge becomes even more pronounced in small-scale or niche domains (such as newly introduced content types), where user behaviour data is limited. Therefore, making it difficult to make reliable predictions due to the shorter and less consistent behavioural sequences.

One major hurdle in this setting is the ability to extract and align user interest across multiple domains or platforms. Candidate items in the target domain (e.g., live streams) often do not have enough behavioral data, making it hard to match them effectively with relevant user history from other domains (e.g., pre-uploaded videos). Moreover, traditional methods struggle to effectively filter due to use of simpler attention models and relate cross-domain behavior due to use of LSM, leading to suboptimal CTR prediction.

This leads us to question on how we could effectively align the representation space between candidate items in target domain and the sequence items in source domain and effectively filter data through domains, ultimately increasing the accuracy of CTR predictions.

Related Works

It has been proven that longer sequences improve the predictability of CTR which lead to the development of LSMs. These LSM models aim to extract user interests with respect to the candidate items in order to accurately predict CTR.

LSM has two components GSU(general search unit)- responsible for going through the sequential set of the users to identify items closely related to the candidate and ESU(exact search unit) responsible for picking up items from those chosen from GSU that interests the users the most. While this is an effective way for predicting CTR, it falls short when it comes to areas that do not posses a long range of behavioural sequences.

These methods also are insufficient while using cross-domain LSM which requires linking of the source domain(domain that possess long sequence of the users) and target domain(domain where the CTR predictability has to be made).

Solution

Use of Lifelong Cross Networks(LCN) for the cross-domain LSM featuring a cross representation production module that would enhance the learning of relationships among items between domains and Lifelong attention pyramid(LAP) that would effectively filter out user interests.

Methodology

The LCN model comprises of two major components: The CRP and LAP

Cross Representation Model:
In order to enhance learning of relationships between items across domains, CRP pairs positive and negative examples from users' short-term behaviour sequences. Positive pairs are done by choosing positive items(items that are similar) from each source and target domains short term behavioural sequence and then creating a cross-domain positive pair by selecting one item from source and other from target. Negative Pairs are done by using items from different users behaviour sequence(user B) with different interest from their source and target domains and pairing them with the items of this user(User A). The source items of user B and user A and similarly their target domain items and finally the cross-negative pairs(User B's source and User A's target) are created.

Lifelong Adaption Pyramid:

It is divided into three stages:

1. The Complete-Scope Attention (CSA): Here the model goes through the entire lifelong sequence and gives each item an attention score for each item e^h_k based on its similarity to the candidate itemset e_i^v based on the formula stated and the top-K items are selected for the next stage. Weighted summation is calculated

2. The Median-Scope Attention(MSA): A secondary attention search is performed using contextual details(viewing duration etc). The attention scored based on these LCS(vector of contextual information) is calculated using the formula stated. Now based on the scores the top-K2 items are elected to the final stage. Weighted summation is calculated

3. The focused-scope Attention(FSA): The use of attention similar to multi-head transformer is used. From the sub-sequence obtained from the previous stage the output is calculated as the formula and its weighted summation is calculated.

In the end all the weighted summations are integrated and the final vector of similar items to that of candidate items which the user is interested in is created.

It represents improvement as the three tier LAP model ensures better filtering using advance focusing than the LSM's GSU and ESU model. It also has cross domain reference sequences based on the source and target domain integrating cross domain interactions effectively and increase CTR accuracy across small areas.

Evaluation:

The module is evaluated using the data set collected from the Taobao dataset and the Wechat platforms and compared with its competitors using SOTA methods in the LSM field: SIM Soft,ETA,SDIM,TWIN.

For offline evaluation, AUC, GAUC, and Logloss metrics were employed. The public dataset, sourced from Taobao, uses dimensions of 256 and 128, while the industrial dataset scales these up to 512 and 256. The short-term behavior sequence length is fixed at 50 for both datasets. All feature types share a uniform embedding size of 64. Model parameters are initialized using Xavier Initialization and optimized using the Adam optimizer with a learning rate of 0.001, implemented in TensorFlow. Training is conducted with a batch size of 2048, utilizing a single A100 GPU for the public dataset and a distributed setup of 8 A100 GPUs for the industrial dataset. On real-world industrial data, the model achieved a 2.93% improvement in CTR and a 3.27% increase in user stay time.

Contribution:

Contributions of this Paper include:

Improving the CTR accuracy for the areas that have smaller behavioural sequences using cross domain LSM

Bridging the gap between the source and target domains using LCM

Introduced ways to actively test and train the models

Validated the performance of the LCN model by performing proper tests and stating the improvement

Given the readers insight into the methods of data collection and testing

Wide range of testing

Optimized it for real world use

Limitations and Future Direction:
This model even though solves the problem in domains with little sequences it still depends on a closely related source domain to extract sequences making it still harder for fresh completely new domains

Doesn’t have integrations for text/images making it harder for implementing in areas which are more photo dominant like Instagram

There is still no answer to why certain people choose to engage with certain items, hence any new items that a user may like can still have wrong CTR Predictions.

Therefore for the future, these limitations can be eliminated by using models that do not require long sequences but short sequences that can constantly learn with feedback as well as models with meta data to learn and adapt to newer domains can be utilized.