ISSN: 2688-8645

Explicit and Implicit Recommendation Algorithms Fusing Multi-head Attention Mechanism

Ru Ma

School of Control and Computer Engineering, North China Electric Power University, Baoding 071000, China

Abstract

The accuracy of the ranking stage largely depends on the feature interaction. Although great progress has been made, existing related algorithms and models can realize the feature interaction between shallow, deep explicit and deep implicit, without mining users. Diversity of interests. To address this issue, a Multi-head attention module is introduced on the xDeepFM benchmark model. For the convenience of distinction, the improved model in this paper is named MHA_xDeepFM. Through interactive modeling of the features between users and items, the deep interaction between the features is found, and the combined features are formed to improve the prediction accuracy of the model. The model proposed in this paper firstly combines the Multi-head attention mechanism to adaptively model the correlation between input features, focusing on the key internal relationship between the data, and then adding the compressed interaction network model to learn the optimal order of combined features, Deep interactive features in complex data are learned through deep neural networks. Finally, based on the click rate, the recall list is filtered and sorted. Finally, the corresponding experiments are carried out. The experimental results show that compared with similar algorithms, the AUC indicators of MHA_xDeepFM are improved by 2.7% and 3.2% on the two datasets, respectively.

Keywords

xDeepFM; Recommendation Algorithm; Attention Mechanism.

1. Introduction

With the advent of the Internet era, the rapid development of network technology, the explosive growth of related data, and users are caught in the sea of information[1]. Due to the increasing amount of information, people must face the problem of information overload. To solve the problem of information overload, researchers propose corresponding solutions, which can be roughly divided into two categories, search engines and recommendation systems. Search engines are aimed at users with clear goals[2]. When users enter keywords in the search box, the search engine will return the corresponding query results to the user, but when the user's goal is not clear, the search engine will lose its corresponding search results. effect. The recommendation system comes into being, it models the user's historical behavior, and then searches for the items that the user may like from the candidate item set according to the corresponding recommendation algorithm, and finally displays the recommendation result to the user[6]. Nowadays, recommendation systems have been successfully used in many fields, such as online shopping fields such as JD.com and Amazon, and movie recommendation fields such as YouTube[7].

Nowadays, deep learning has made breakthroughs in many fields. Since 2016, recommendation algorithms based on deep learning have also received more attention[5]. The application of deep learning in recommendation systems has become an important research in artificial

intelligence. direction. Since deep learning has the ability to extract data features from end-to-end models, it overcomes the dependence of traditional recommendation algorithms on artificial features[6]. Secondly, traditional recommendation algorithms can only mine shallow features and cannot mine deep features. In the real world, deep feature interaction is also essential, such as the classic examples of "beer" and "diapers", deep learning-based recommendation algorithms can take into account both deep and shallow interactions[7]. Although deep learning-related recommendation algorithms have made great progress, there are still some deficiencies in current recommendation algorithms. Recommendation algorithms pay more attention to feature combination construction, realize low-level and high-level interaction of features, and solve traditional artificial feature engineering[8]. It is a time-consuming and labor-intensive problem, but most of these ranking algorithms do not take into account the diversity of user interests[9]. Therefore, how to integrate the attention mechanism into the ranking algorithm to solve the problem of single user interest representation is the research focus of this paper. With the continuous development of the Internet, the requirements for the accuracy of the recommendation algorithm are getting higher and higher. The recommendation system not only meets the requirements of personalization and diversification, but also brings certain profits to the market [16]. To sum up, recommendation algorithms based on deep learning are of great research value both in theory and in practical applications.

2. Related Work

In order to obtain a more accurate recommendation effect, the recalled candidate sets are often sorted according to the size of the click-through rate. The Wide&Deep Learning model was first proposed. The Wide part of the model uses a shallow model to improve the memory ability of the model, and the Deep part of the neural network makes the model have generalization ability, which is reflected in the ability to mine user historical data[17]. The features that do not appear in the model structure are shown in Figure 1:

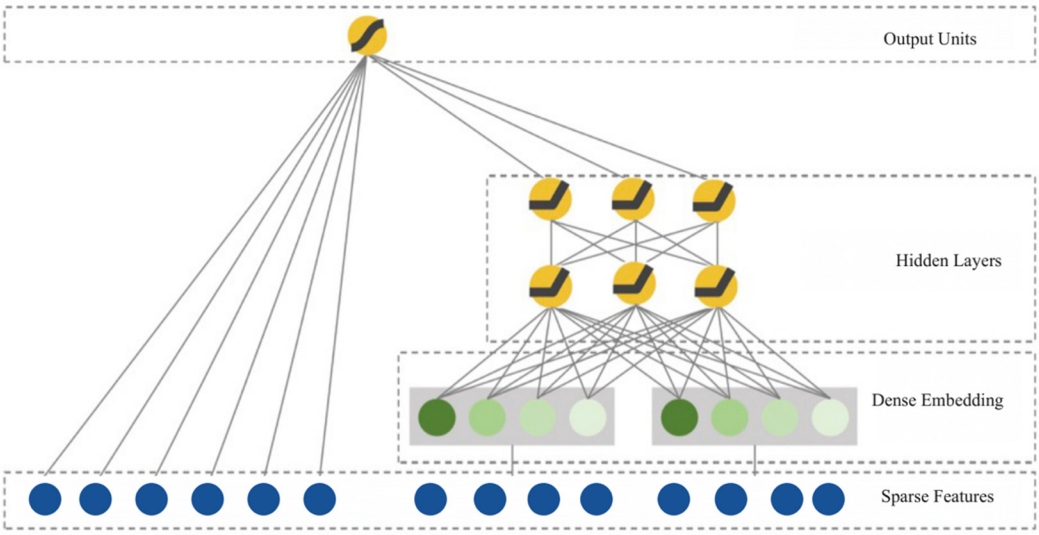

Figure 1. Wide&Deep model structure

Like the deep-wide recommendation model, He et al. [18]used a deep neural network to model the complex interaction between users and items in a nonlinear way, and proposed a neural collaborative filtering algorithm, which uses a multi-layer neural network to learn users. By instantiating the model, a generalization structure of matrix factorization and a multi-layer perceptron structure are obtained, which are then used to model the linear and nonlinear

ISSN: 2688-8645

interactions between users and items, respectively. structure. Combining features is the key to the success of recommendation algorithms. With the great success of deep neural networks in various fields, researchers have successively proposed several recommendation algorithms based on deep learning. xDeepFM was proposed by Lian et al. the purpose of this model is to generate the interaction of features in an explicit manner and at the vector level. The advantage of xDeepFM is that it proposes a CIN structure that can efficiently learn high-order explicit feature interactions, feature interactions occur at the vector level, and the network complexity does not grow exponentially with the degree of interaction at any time. However, the user preferences obtained by xDeepFM are consistent, because the average pooling method is adopted, resulting in the same user interest representation for any candidate item.

3. Model Structure

The MHA_xDeepFM model is sent to the Multi-head attention module after the embedding layer, the purpose is to allow the model to adaptively learn the importance of the input features and help to learn effective combined features. Then input the Compressed Interaction Network (CIN) and DNN modules. The CIN part is responsible for high-order explicit learning of the interactive features of the optimal order; the DNN part is responsible for learning high-order implicit feature interactions, and the linear part is responsible for low-order feature interactions. In addition, the model adds context information to the input part, and embeds the interaction time between the user and the item through the Embedding method, so that the recommendation algorithm can make full use of the time context information. In this way, when making accurate recommendations for users, the sequence characteristics behind the behavior sequence can be taken into account, thereby reflecting the diversity of user interests. The model structure of MHA_xDeepFM is shown in Figure 2:

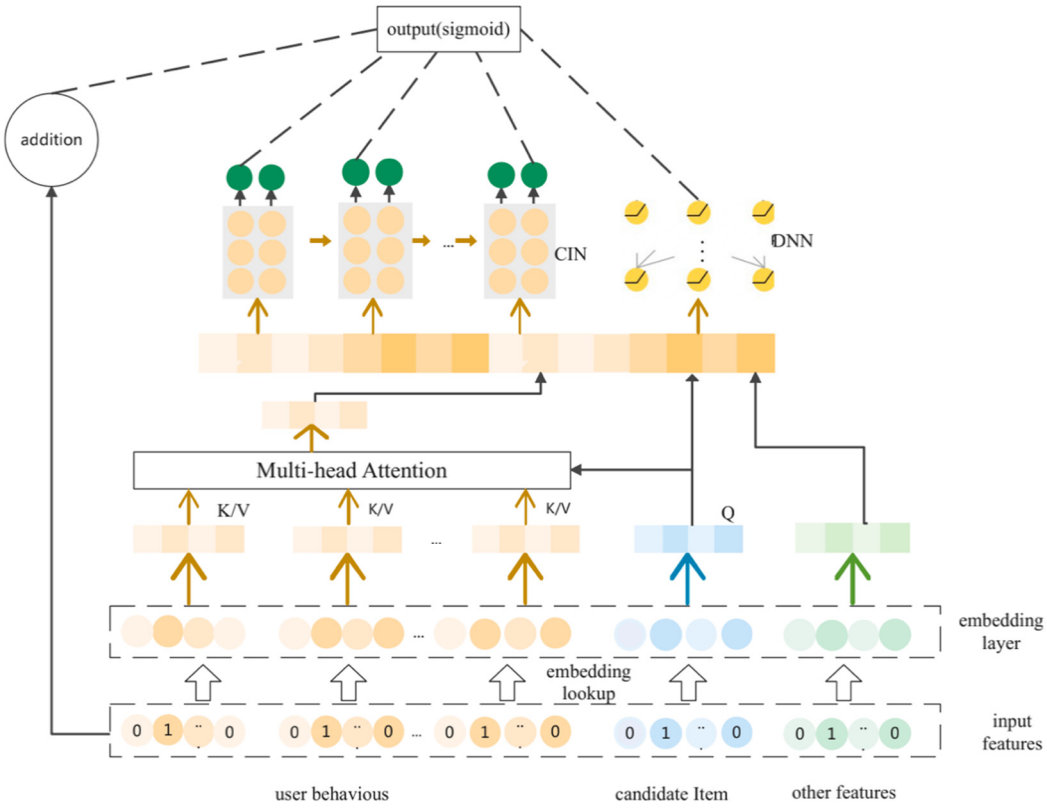

Figure 2. MHA_xDeepFM model structure diagram

ISSN: 2688-8645

The characteristics of the MHA_xDeepFM model are as follows:

(1) By using the Multi-head attention mechanism to learn the relevant weights of the items that the user has interacted with and the candidate items, the model can further effectively combine the original features and learn more meaningful feature combinations, thus improving the accuracy of the model. Spend.

(2) Integrate three methods, namely linear model, CIN and DNN, to mine the dependencies among first-order interactions, higher-order explicit interactions and higher-order implicit interactions of user behavior sequences, to obtain rich behavior sequence information.

(3) In the input part, the model adds contextual information, and by embedding the interaction time between the user and the item, the recommendation algorithm can make full use of the temporal contextual information.

3.1. Multi-head Attention

Multi-head attention mechanism module is shown in Figure 3[19]. First, the three feature vectors Q, K, and V are sub-linearly mapped, and then the corresponding mapping results are spliced. Finally, the final output structure is obtained through linear operation. The output of the module is shown below.

$$head_i = Att(QW_i^Q, KW_i^K, VW_i^V) \tag{1}$$

$$MHead(Q,K,V) = Concat(head_1, head_2, ...head_l)W \tag{2}$$

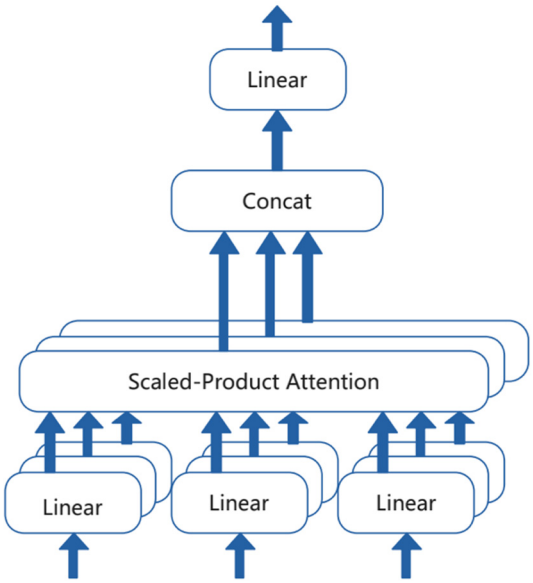

Figure 3. Structure diagram of Multi-head attention mechanism

3.2. Model Output

The MHA_xDeepFM model first linearly combines the outputs of the CIN, DNN, and linear parts, and then processes them by the Sigmoid function, and finally obtains the result of click-through rate prediction. The output of the MHA_xDeepFM model is shown in Equation 3.

$$y = \sigma(w_{linear} y_{linear} + w_{cin} y_{cin} + w_{dnn} y_{dnn} + b) \tag{3}$$

ISSN: 2688-8645

4. Experiment and Analysis

In order to verify the effectiveness of the MHA_xDeepFM model, corresponding experiments were carried out on two public datasets. After a series of tuning experiments, the experimental results on MovieLens-1M and Electronics are shown in Table 1.

Table 1. Experimental results under MovileLens-1M and Electronics datasets

	MovieLens AUC	Logloss	Electronics AUC	Logloss
wide&deep	0.7673	0.4792	0.8073	0.2986
DCN	0.7819	0.4564	0.8219	0.2924
deepFM	0.7924	0.4436	0.8324	0.2857
AFM	0.7747	0.4597	0.8196	0.2903
DIN	0.8023	0.4224	0.8591	0.2628
xDeepFM	0.7901	0.4382	0.8524	0.2687
MHA_xDeepFM(ours)	0.8105	0.4128	0.8639	0.2529

From Table 1, the following conclusions can be drawn:

(1) Overall, the model proposed in this paper outperforms other models on MovieLens and Electronics, suggesting that the combination of a Multi-head attention mechanism with explicit and implicit higher-order feature interactions is necessary. In the two datasets, RelaImpr improves by 2.7% and 3.2%, respectively, proving the effectiveness of the model.

(2) The results of the DIN and xDeepFM models outperform other models that do not contain an attention mechanism, proving the effectiveness of the Multi-head attention mechanism in the recommendation domain. However, although the AFM model introduces an attention mechanism, the overall effect is not ideal. The reason is that AFM does not take advantage of the deep neural network and ignores the deep interaction mode.

(3) xDeepFM, MHA_xDeepFM outperform other models that are not able to learn both explicit and implicit higher-order feature interactions, which indicates that it is necessary to combine explicit and implicit higher-order feature interactions.

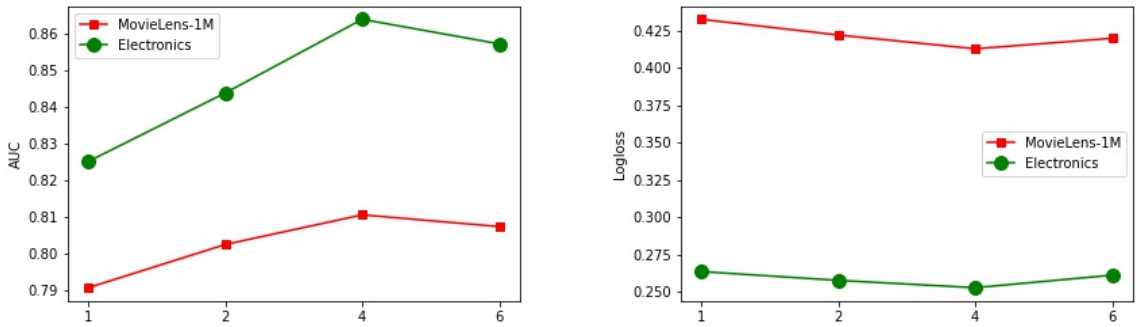

Figure 4. AUC experimental results with different head spaces

In order to study the influence of the number of head spaces of the Multi-head attention mechanism on the model, a set of experiments is designed, in which the number of head spaces is 1, 2, 4, and 6. The experimental results are shown in Figure 4-8. It can be seen from the experimental results that the increase of the number of head spaces of the Multi-head attention mechanism initially improves the performance of the model, but as the number of head spaces

continues to increase, the performance of the model will decrease, which may be the cause of overfitting.

5. Summary

Since the shallow model cannot deeply mine the hidden features of users and items, and cannot simulate complex interaction scenarios between users and items, this paper uses a deep learning model to fine-tune the recall candidate set list. Considering that different input features have different degrees of importance, in order to mine the intrinsic correlation between features and learn meaningful combined features, this paper constructs an xDeepFM algorithm that integrates Multi-head attention mechanism. Conduct interactive modeling to find deep-level interactive relationships between features, and form combined features to improve the prediction accuracy of the model. The model proposed in this paper firstly combines the Multi-head attention mechanism to adaptively model the correlation between input features, focusing on the key internal relationship between the data, and then adding the compressed interaction network model to learn the optimal order of combined features, Deep interactive features in complex data are learned through deep neural networks. Finally, based on the click rate, the recall list is filtered and sorted. Through a series of comparative experiments with popular algorithms such as DIN and xDeepFM, the superiority of the model proposed in this paper in the click-through rate prediction task is verified.

References

[1] Zhang F, Yuan N J, Lian D, et al. Collaborative knowledge base embedding for recommender systems [C]//Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. 2016: 353-362.

[2] Zhou X, He J, Huang G, et al. SVD-based incremental approaches for recommender systems[J]. Journal of Computer and System Sciences, 2015, 81(4): 717-733.

[3] Chaney A J B, Blei D M, Eliassi-Rad T. A probabilistic model for using social networks in personalized item recommendation[C]//Proceedings of the 9th ACM Conference on Recommender Systems. 2015: 43-50.

[4] Li G, Ou W. Pairwise probabilistic matrix factorization for implicit feedback collaborative filtering[J]. Neurocomputing, 2016, 204: 17-25.

[5] Song Y, Elkahky A M, He X. Multi-rate deep learning for temporal recommendation[C]//Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval. 2016: 909-912.

[6] Covington P, Adams J, Sargin E. Deep neural networks for youtube recommendations [C]// Proceedings of the 10th ACM conference on recommender systems. 2016: 191-198.

[7] Huang P, He X, Gao J, et al. Learning deep structured semantic models for web search using clickthrough data [C]//Proceedings of the 22nd ACM international conference on Information & Knowledge Management. 2013: 2333-2338.

[8] Hidasi B, Karatzoglou A, Baltrunas L, et al. Session-based recommendations with recurrent neural networks[J]. arXiv preprint arXiv:1511.06939, 2015.

[9] Wang J, Huang P, Zhao H, et al. Billion-scale commodity embedding for e-commerce recommendation in alibaba[C]//Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2018: 839-848.

[10] Li C, Liu Z, Wu M, et al. Multi-interest network with dynamic routing for recommendation at Tmall [C] //Proceedings of the 28th ACM international conference on information and knowledge management. 2019: 2615-2623.

[11] Wang R, Fu B, Fu G, et al. Deep & cross network for ad click predictions[M]//Proceedings of the ADKDD'17. 2017: 1-7.

[12] Lian J, Zhou X, Zhang F, et al. xdeepfm: Combining explicit and implicit feature interactions for recommender systems [C]//Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining. 2018: 1754-1763.

[13] Xiao J, Ye H, He X, et al. Attentional factorization machines: Learning the weight of feature interactions via attention networks[J]. arXiv preprint arXiv:1708.04617, 2017.

[14] Zhou G, Zhu X, Song C, et al. Deep interest network for click-through rate prediction [C]// Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining. 2018: 1059-1068.

[15] Zhou G, Mou N, Fan Y, et al. Deep interest evolution network for click-through rate prediction [C]// Proceedings of the AAAI conference on artificial intelligence. 2019, 33(01): 5941-5948.

[16] Sun F, Liu J, Wu J, et al. BERT4Rec: Sequential recommendation with bidirectional encoder representations from transformer[C]//Proceedings of the 28th ACM international conference on information and knowledge management. 2019: 1441-1450.

[17] Wu Y, DuBois C, Zheng A X, et al. Collaborative denoising auto-encoders for top-n recommender systems [C]//Proceedings of the ninth ACM international conference on web search and data mining. 2016: 153-162.

[18] He X, Liao L, Zhang H, et al. Neural collaborative filtering[C]//Proceedings of the 26th international conference on world wide web. 2017: 173-182.

[19] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. Advances in neural information processing systems, 2017, 30.