scientific reports

OPEN

Session interest model for CTR prediction based on feature co-action network

Qianqian Wang[1,3]✉, Fang'ai Liu[2], Xiaohui Zhao[2] & Qiaoqiao Tan[2]

The main purpose of click-prediction models is to predict the probability of customers clicking on products and provide support for advertising decisions of businesses. However, most previous models often use deep neural networks to capture implicit interaction and can not fully retain the representational power of the original feature interactions. At the same time, one factor that most models ignore is that sequence is made up of sessions. Therefore, how to model user interest features and preserve the representational properties of feature interactions is the main challenge to improve the accuracy of CTR prediction. According to above issues, this study propose session interest model with feature co-action network (SIFAN). First, we used widely used characteristic co-action network module to tap into the interactions in customer single behavior. Then, the sequential behavior of customers is divided into session layers, and considering that various session interests may follow sequential patterns, gated recursive units are applied to predict customer clicks. Then, by analyzing the GRU with attention update gates, the correlation between conversation interest and target items is determined. According to relevant experimental results, under the same experimental conditions, the SIFAN model has significant performance advantages compared to other models.

Keywords CTR, Prediction model, Behavior sequence, Advertising

Click-through prediction is the hot topics in the field of recommender model[1,2] and online advertising[3].It is mainly used to determine whether customers will click on recommended products or information. The CTR prediction directly affects the revenue of advertising platforms in cost per click (CPC)[4] model. At the same time, it is important for people to obtain the information that they want and enhance their experience. Therefore, how to enhance the property of CTR is very important and needed to be deeply investigated.

With the gradual maturity and widespread application of AI technology, efficient analysis of a large number of input features has become a focus of neural network model research. When predicting CTR, the key factor to consider is customer interest. However, most models ignore the impact of user interest on the accuracy of prediction results. At the same time, previous models treat the customer sequential behaviors as customer interest of the sequences. A sequence is made up by some sessions. In each session, the customer behavior is high homogeneous. In different session, the customer behavior is heterogeneous. Meanwhile, the implicit feature interactions learned in most models can not retain the feature interactions representation property. Therefore, how to model user interest features and preserve the representational properties of feature interactions is the main challenge to improve the accuracy of CTR prediction.

Through the above observation, we propose SIFAN for CTR prediction. This model mainly consists of two parts, session interest extractor and interacting module, feature co-action network (FAN) module. We use many historical sessions to model the continuous behavior of customers and apply a self-attention mechanism to determine relevant session interest. At session interacting layer, this study mainly uses the AUGRU model to predict and reveal the impact of conversation interest on corresponding item. For the FAN module, we model feature interaction among raw feature at the input stage and capture the interest with the relevant target item.

This article focuses on the following aspects for discussions:

1. We propose an innovative feature interaction scheme based on previous research findings. We propose SIFAN to find the association in customer behaviors and target item. We input the customer behavioral se-

[1]School of Data and Computer Science, Shandong Women's University, Jinan 250300, Shandong, P.R. China. [2]Shandong Key Laboratory of Computer Network, Shandong Normal University, Jinan, China. [3]Shandong Provincial Key Laboratory of Network based Intelligent Computing, Jinan 250022, P.R. China. ✉email: wangqq_sdnu@163.com

quences and target item into the model to explore the association in a single customer historical behavior and a target item, as well as the potential session interest behind historical behavioral sequences.

2. Interacting unit, this article adopts multiple previous sessions to model the continuous behavior of customer and employ a self-attention mechanism to obtain session interest. Meanwhile, we apply AUGRU to determine the influence of interest on target item. In feature co-action network module (FAN), each feature is assigned to the MLP to determine the relation between different features. In this mode, FAN can more efficiently describe the interaction between features and significantly reduce the number of parameters, thus having high application value.
3. According to relevant experimental research results, the model established in this article has significant advantages in CTR prediction. In addition, this article also analyzed the impact of key variable on the predictive property of the SIFAN to demonstrate its advantages.

The content of the rest parts of present research as follows. In part 2, we discuss the existing studies of CTR prediction. In part 3, the composition structure and features of the SIFAN model are detailed. The fourth section conducts experimental verification of this model, then analyzes the experimental results. In part 5, we prospect the future research directions and elaborates on the future development trends in this field.

Related work

In such prediction task, several studies have been devoted to modeling feature interactions and capturing customer interest behind customer historical behavior. LR[5] is widely applied in the industrial field for CTR. Its advantages are easy to understand, fast computation speed, and suitability for large-scale data analysis. Kumar et al.[6] applied LR to predict the CTR. The LR converts characteristic into vectors as input. In this way, feature interaction is ignored. Factorization Machine[7] (FM) adds second-order feature interaction and achieve better result. The field-aware factorization machines (FFM)[8] is proposed based on FM with field aware latent vectors. However, these models are weak in obtaining high-order feature interaction and limit representation of feature interaction. Learning feature interaction is very superior in many task.

Deep neural network (DNN) is an algorithm constructed and has achieved better results in many research areas such as artificial intelligence[9,10], image processing[11,12] and target recognition[13,14]. The advantage of this model is its strong adaptive ability, which can effectively learn complex patterns and characteristic from data. In addition, it can also handle nonlinear problems and has good generalization ability.Hence, it was applied in many CTR prediction models. DeepFM[15] can enhance the accuracy of CTR prediction by capture high-order feature interaction. It is important to learn combined feature. Product-based Neural Networks (PNN)[16] capture characteristic interactions by outer product. The Operation-aware Neural Networks (ONN)[17] constructs a new deep neural network and learns different representations for different operations. Deep and Cross Network (DCN)[18] captures the characteristic combination of specific order, which replaces the Wide &Deep[19] with Cross Network. In practical applications, different features contribute differently to the prediction results. Attention mechanism[20] is powerful in distinguishing the importance of features. Attentional Factorization Machines (AFM)[21] assigns the different importance to different features and improves the CTR precision. A hierarchical attention network was constructed by Long et al.[22]. The model refines the characteristic representation at each layer based on Attention mechanism. In addition, several CTR prediction models are all built on the basis of convolutional networks (CNN)[23]. Liu et al.[24] established a CSCNN model and used it to predict CTR. The results showed that it exhibited high application performance advantages. At the same time, some methods utilize RNN[25] to obtain the sequence dependency among the customer behavior. But, these models also have certain limitations, namely the inability to determine the impact of customer interest on CTR prediction. Meanwhile, implicit characteristic interactions learned in DNN do not retain the full representational power of different feature interaction in a completely loss manner.

Due to not considering the influence of customer interest factors, which has a adverse influence on the reliability of the obtained results. Based on extensive application experience, models that can capture dynamic customer interest have significant advantages in CTR prediction, with improved prediction accuracy. Deep Interest Network (DIN)[26] captured customer interest below historical behavioral sequences. This attention mechanism module can accurately analyze the customer historical behavior and determine the corresponding interest changes during the application process, which can establish a better set of customer behavior characteristics and provide support for subsequent model predictions, thus effectively improving the predictive effect of the model. The model can enhance the accuracy of the final result. Since customer interest are constantly changing, Deep Interest Evolution Network (DIEN)[27] model was constructed by Zhou et al. based on DIN. The DIEN models customer interest evolving process. Search-based Interest Model (SIM)[28] is constructed to find customer interest in the target item, which can accurately analyze the correlation between target projects and behavioral sequences, providing support for subsequent model construction. We figure out that those models enrich the feature representation and improve the precision of the model. The concept of conversation is mentioned in[29,30], but this concept is rarely applied in CTR. The session based recommendation model can adjust the weights of relevant parameters according to changes in customer interest during prediction, thereby improving prediction accuracy. Liu et al.[31] introduced short-term attention model to improve prediction performance, which can capture customer interest and their changes based on the background information of the conversation. In each session, the customer behavior is high homogeneous. In different session, the customer behavior is heterogeneous. We design session interest extractor to obtain session interest.

Most of the above models do not fully consider the explicit and implicit customer interest perspectives when mining interest preferences through customer behavior. Meanwhile, most previous work has not adequately processed the raw data and not mine the customer single behavior. This will greatly alter the reliability and effectiveness of the model, leading to limitations in the results obtained. To treat the above issue, this research design feature co-action network module (FAN), and propose a session interest model with feature con-action network (SIFAN) to enhance CTR prediction accuracy.

Materials and methods

We propose SIFAN for CTR prediction. The overall architecture is exhibited in Fig. 1. The model included two main parts, session interest extractor and interacting module, feature co-action network (FAN) module. The two modules capture information from: characteristic-level interactions in the customer historical behavior and target item, and the session interest behind customer historical behavioral sequences.

Characteristic representation and embedding

This research takes four sets of features as input information. Each characteristic is described by $E \in \mathbb{R}^{M \times d_{\text{model}}}$, where M refer to sparse features size. Customer Profile were described by $X^U \in \mathbb{R}^{N_u \times d_{\text{model}}}$, where$N_u$ refer to sparse characteristic number. Scene Profile and Target Item were described as .$X^{\mathrm{S}} \in \mathbb{R}^{N_s \times d_{\text{model}}}$.,.$X^{\mathrm{I}} \in \mathbb{R}^{N_i \times d_{\text{model}}}$., where .$N_i$.refer to Target Item sparse characteristic number. Customer Behavior can be described by.$X^{\mathrm{B}}{=}[x_1; ?; x_i; ?x_{\mathrm{N}}] \in \mathbb{R}^{N \times d_{\text{model}}}$., $x_i$is the embedding of the *i*-th behavior.

Feature co-action network module

The collaborative unit is the micro MLP for each feature pair, and during the model analysis process, the weight and bias related parameters of the MLP are given through the feature pairs. Combining the following diagram for analysis, it can be seen that the constituent units of the micro MLP in Fig. 2. The characteristic of a customer and target item were described as $U_{user}$and .M_{item}. For $u_0 \in U_{user}$, we use parameter lookup to obtain learnable parameters $P_{user} \in \mathbb{R}^D$. At the same time, item feature $m_0 \in M_{\text{item}}$ to obtain … In the micro-MLP P_{item} is reshaped and split. It can be formulated as follow.

$$||_{i=0}^{L-1}(w_i \parallel b_i) = P_{\text{item}} \tag{1}$$

$$\sum_{i=0}^{L-1}(|w_i| + |b_i|) = |P_{\text{item}}| \tag{2}$$

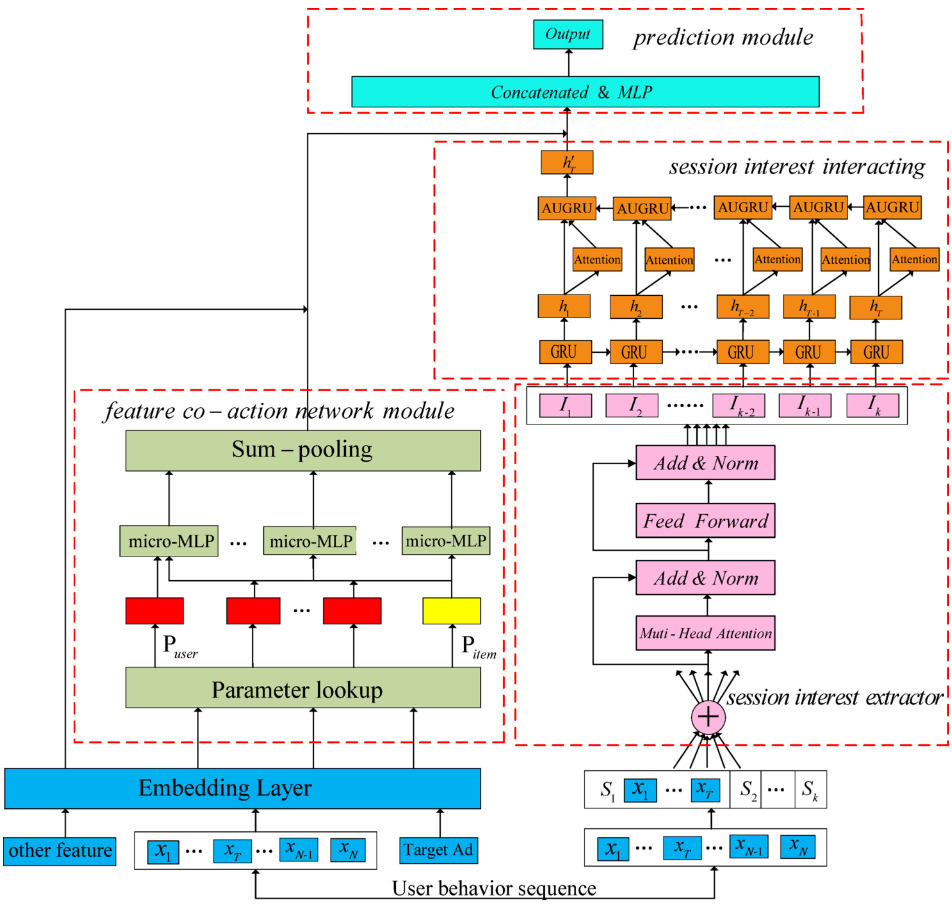

Fig. 1. The construction of SIFAN.

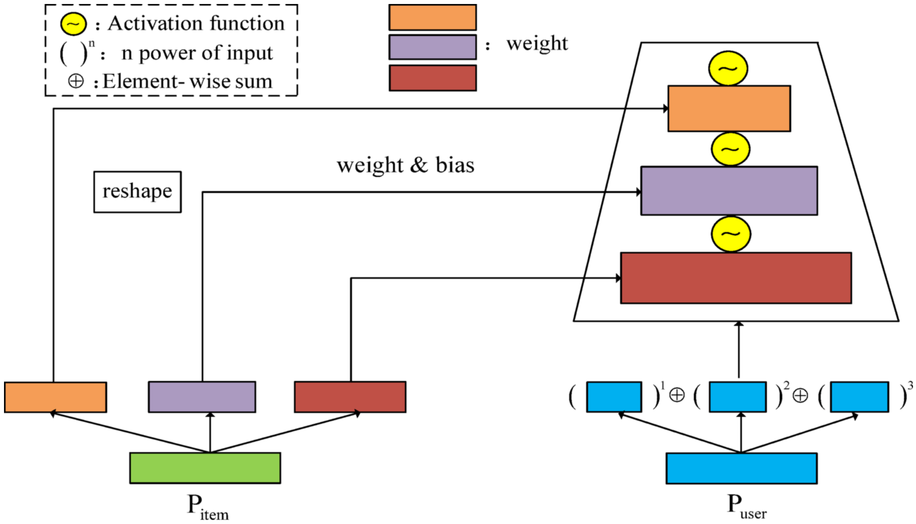

Fig. 2. The structure of micro-MLP in FAN.

Where and are the weight and bias, $|\cdot|$ is the variables size. Then, P_{user} is fed into the micro-MLP through cascading outputs from each layer:

$$h_0 = P_{user} \tag{3}$$

$$h_i = \sigma(w_i \otimes h_i + b_{i-1}), i = 0, 2, \cdots, L-1 \tag{4}$$

$$F(u_0, m_0) = H(P_{\text{item}}, P_{user}) = ||_{i=0}^{L-1} h_i \tag{5}$$

where.$\otimes$ is the matrix multiplication, H is the interaction by passing and P_{item} into the micro-MLP,σdenotes activation function. For sequential feature like customer behavior history $P_{\text{seq}} = \{P_{b(t)}\}_{t=1}^{T}$, we apply a sum-pooling to treat data:

$$H(P_{\text{item}}, P_{\text{seq}}) = H(P_{\text{item}}, \sum_{t=1}^{T} P_{b(t)}) \tag{6}$$

Using this approach we can explore the association between a single customer behaviour and a target item.

Session interest extractor module

Step 1, convert Xto sessions *S*, in which the *k*-th session $S_k = [x_1; .; x_i; .x_T] \in \mathbb{R}^{T \times d_{\text{model}}}$, b$_i$is customer *i*-th behaviors. Many behaviors over 30 min are divided into customer sessions according to Grbovic's[32] method.

This study utilized a self-attention method[33] to obtain the inner association in behaviors and discover the effects of unrelated behaviors. Then we encode the embedding based on this mechanism. Research has found that many factors can affect customer click behavior, such as color, price, product description, etc., and this influence is closely related to customer age and gender factors. Multi-head self attention is used to determine the correlation of objects in various subspace. So $S_k = [S_{k1}; .; S_{kn}; .S_{kN}]$, the$S_{kn} \in \mathbb{R}^{T \times d_n}$refer to the *n*-th head of S_k. $d_n = \frac{1}{n} d_{\text{model}}$. So the output of $head_n$ can be get:

$$head_n = Attention(S_{kn}W^Q, S_{kn}W^K, S_{kn}W^V) \tag{7}$$

$$Attention(Q, K, V) = \text{softmax}(\frac{S_{kn}W^Q W^{K^T} S_{kn}^T}{\sqrt{d^{\text{model}}}}) S_{kn} W^V \tag{8}$$

$$I_k^S = FNN(Concat(head_1, .head_N)W^O) \tag{9}$$

$$I_k = Avg(I_k^S) \tag{10}$$

where $I_k$represent the customer *k*-th session interest.

It is well known that different session interest can follow sequential patterns. We choose GRU[34,35] to find the interaction of customer various historical session interest.

Here, the activation$h_t$of the GRU at time *t* is a linear interpolation between the previous activation$h_{t\text{-}1}$and the candidate activation$\tilde{h}_t$:

$$h_t = (1 - z_t)h_{t-1} + z_t \tilde{h}_t \tag{11}$$

This article mainly adopts the following methods to determine update gate z_t:

$$z_t = \sigma(W_z \mathrm{I}_t + U_z h_{t-1} + b_z) \tag{12}$$

This article adopts the following methods to determine candidate activation :

$$\tilde{h}_t = \tanh(W I_t + u(r_t \odot h_{t-1}) + b_h) \tag{13}$$

where$r_t$is a set of reset gates and$\odot$is an element-wise multiplication. When off ($r_t$close to 0), the reset gate makes it forget the state of the previous calculation.

$$r_t = \sigma(W_r \mathrm{I}_t + U_r h_{t-1} + b_r) \tag{14}$$

.

Here, $r_t$is a set of reset gates, $\sigma$is activation function, W and b are trained parameter.

The hidden state$h_t$can determine the relation in session interest. The customer session interest associated with target ad can impact customer clicking behavior. So we reassign the weight of the customer session interest for ad. Here, the $I_t{}'$mean input state of GRU. The input to the second GRU can be represented : $I_t{}'{=}h_t$. The attention function as:

$$a_k^I = \frac{\exp(I_k W^I X^I)}{\sum_k^K \exp(I_k W^I X^I)} \tag{15}$$

We combine GRU, using AUGRU to clarify the influence of session interest on target ad:

$$\tilde{u}_t' = a_k^I * u_t' \tag{16}$$

.

$$h_t{}' = (1 - \tilde{u}'{}_t) \circ h_{t-1}' + \tilde{u}'{}_t \circ \tilde{h}'{}_t \tag{17}$$

Where $u_t'$is the original update gate of AUGRU, $\tilde{u}_t'$is the attentional update gate that we use in AUGRU. $h_t{}'$, h_{t-1}' and $\tilde{h}'{}_t$are the hidden states of AUGRU. Figure 3 exhibits the structure of AUGRU.

We calculate the customer click probability through softmax function. The expression corresponding to the loss function is :

$$L = -\frac{1}{N} \sum_{(x,y)\in D}^{N} (y \log p(x) + (1-y)\log(1-p(x))) \tag{18}$$

where$p(x)$means customer click probability on ad.

Experiments

Experiments setting

Datasets. We study performed the experiments at Amazon dataset[36] to prove the performance of SIFAN model, which contains Books and Electronics. To enhance the representation and accuracy of the test results, when validating the performance of the model, two common datasets were applied: Avazu and Criteo. Table 1 shows the specific situation of each dataset, which is divided into training set (80%), according to different tasks.

Evaluation metrics. We employ three metrics: AUC, RMSE. The AUC index was widely used in classification problems[37]. A larger value of AUC indicates a higher accuracy. Logloss is often applied to binary classification problems. A smaller value for Logloss indicates better performance. RMSE[38] can be expressed as:

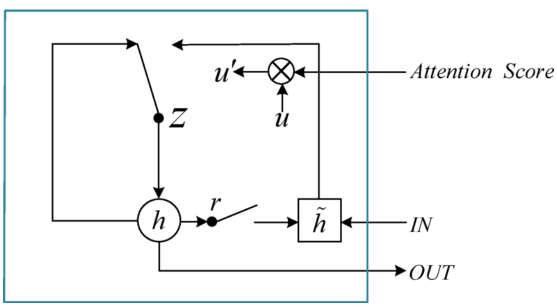

Fig. 3. The architecture of AUGRU.

Dataset	customers	Items	Features	Samples
Books	46,375	43,728	49,852	277,303
Electronics	29,209	34,816	35,784	194,771
Avazu	78,437	69,175	104,242	830,163
Criteo	28,625	30,216	43,127	196,329

Table 1. The datasets.

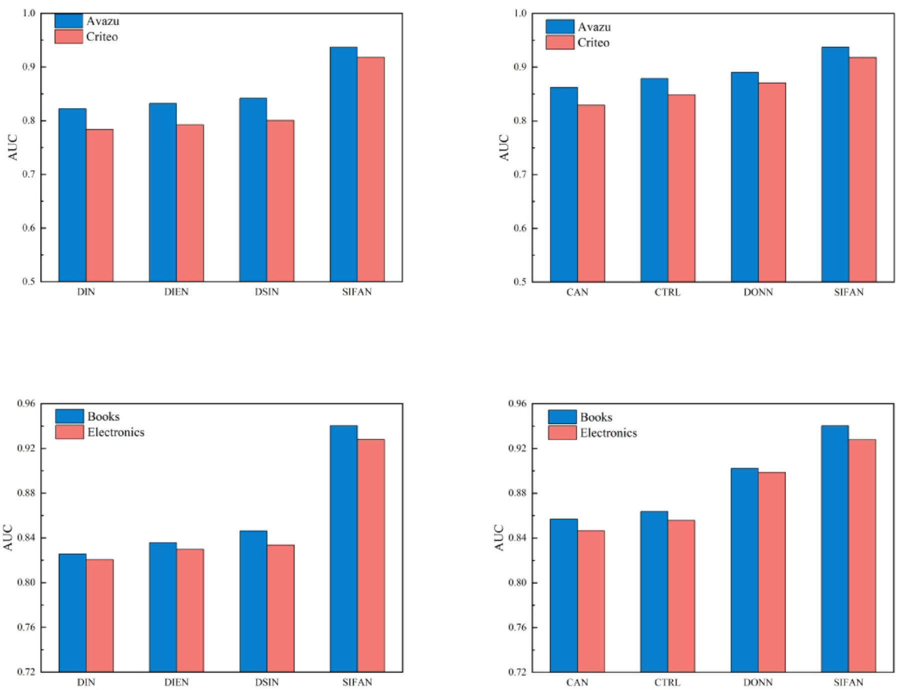

Fig. 4. AUC result in all models.

$$RMSE = \sqrt{\frac{1}{|T|}\sum_{i}\left(y_i - \hat{y}_i\right)^2} \tag{19}$$

where $y_i^t$is the observed scores, T refer to the testing set. When the values of $y_i^t$and $\hat{y}_i^t$are close, the prediction result is more accurate.

Parameter settings. We use dropout rate of 0.5 to the neural networks. Meanwhile, this study use 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1} for verification. Specific analysis shows that, neurons number is in range of 200 to 800.

Comparisons with various models

In the experiment, this article compared SIFAN with some mainstream CTR prediction models during the experimental process, and organized the prediction results. From this, it can obtain a conclusion that the effectiveness of the model is obvious and has extremely high application value, providing support for its practical application and better meeting the requirements of customer click prediction. Analyzing Fig. 4; Tables 2 and 3, it can be seen that the values of each model's prediction results.

1. DIN[26] is a model that use attention mechanisms extract customer interest from customer historical behavior. Its core idea is to analyze and elucidate the characteristics of customer behavior based on attention mechanisms. The model improves the performance of CTR prediction.
2. DIEN[27] is an improved version based on DIN. The model designs an interest extractor layer to get customer interest. Meanwhile, the model can find interest evolving process. DIEN has the better performance than DIN. Therefore, it is favored by customers.
3. DSIN[39] is a model that leverages customer multiple previous sessions in sequences. The model extract customer interest and capture the sequential relation of contextual session interest. It is effectiveness both on advertising and recommender datasets. Analysis shows that this model can better capture user interest, preferences, and behavioral patterns, thereby improving the accuracy of the model. CAN[40] captures the association in customer previous behaviors and target item based on MLP. The model gets higher prediction accuracy.

Model	Logloss			
	Books	Electronics	Avazu	Criteo
DIN	0.1382	0.1406	0.2951	0.3625
DIEN	0.1313	0.1354	0.2836	0.3587
DSIN	0.1142	0.1225	0.2472	0.3473
CAN	0.1094	0.1184	0.2289	0.3366
CTRL	0.0997	0.1152	0.2176	0.3262
DONN	0.0962	0.1079	0.2101	0.3127
SIFAN	0.0873	0.1023	0.2052	0.3004

Table 2. Results for Logloss.

Model	RMSE			
	Books	Electronics	Avazu	Criteo
DIN	0.4527	0.5245	0.4725	0.5825
DIEN	0.4476	0.4913	0.4601	0.5797
DSIN	0.4434	0.4876	0.4579	0.5634
CAN	0.4321	0.4725	0.4432	0.5468
CTRL	0.4252	0.4643	0.4402	0.5346
DONN	0.4063	0.4431	0.4215	0.5105
SIFAN	0.3756	0.4032	0.4001	0.4936

Table 3. Results for RMSE.

4. CTRL[41] model converts the raw tabular data into textual data, and the tabular data and the converted textual data are considered as two different modalities, which are fed into the collaborative CTR model and the pre-trained language model, respectively. The CTRL model achieved better prediction results. However, the model does not take into account the impact of the user interest profile on the prediction results.
5. The DONN[42] model utilises the feature bilinear module to perform outer product calculations between feature pairs, enhancing the MLP to be able to identify more non-linear dependencies. The model achieved high prediction accuracy.
6. In SIFAN, we capture customer session interest. At the same time, we design characteristic co-action network module to mine the customer action based on raw data. The model processes the raw data and mine the customer single behavior. We explore the association in a single customer historical behavior and a target item, as well as the potential session interest behind historical behavioral sequences. We can see that SIFAN model improves accuracy in all datasets.

Sensitivity analysis of model variables

This study analyzed the impact of various parameters on the property of SIFAN model, like epoch, neurons number per layer, dropout rateβ.

Theβgives better results for the SIFAN model as shown in Fig. 5. If the value ofsets properly (from 0.5 to 0.8), the SIFAN model achieves optimal performance on all datasets. In this paper, we set$\beta = 0.5$in later test.

Holding other parameters constant, we investigate the influence of various numbers of neurons on the prediction results. As can be seen in Fig. 6, because of the overfitting of the model, it caused the SIFAN model to perform stably or even worse when the number of neurons is 600, 700 or 800. So we set it as 500.

The paper experimentally investigated the effect in various epoch values for accuracy rate. When the epoch is too low, no proper parameter can be found, resulting in poorer performance of the model. After in-depth analysis, it can be concluded that the RMSE in the Amazon dataset changed relatively more. This also indicates that this parameter will have impact on the performance of the model, so it should be optimized based on actual test results to obtain more accurate results. This is because training the model on different datasets requires different number of features and the diversity caused some errors. As Fig. 7, the model is superior under condition of epoch is ranged in 10 to 15. Hence, we set epoch to 13 in our experiments.

Conclusion

In present research, this study proposes the novel model SIFAN, which focus on customer multiple session interest and a special characteristic interaction method. First, we input the customer behavioral sequences and target item into the model to explore the association in a single customer previous behavior and target item. In FAN module, each feature will be assigned to a micro-MLP to capture the interaction with other characteristic. Secondly, when modeling customer behavior, their sequential behavior is divided into sessions, and correlation analysis is performed on different historical sessions to determine changes in customer interest. Then, the interaction of interest in each session is determined through GRU, and the customer target advertisements are aggregated through AUGRU to determine the correlation between interest and these advertisements.

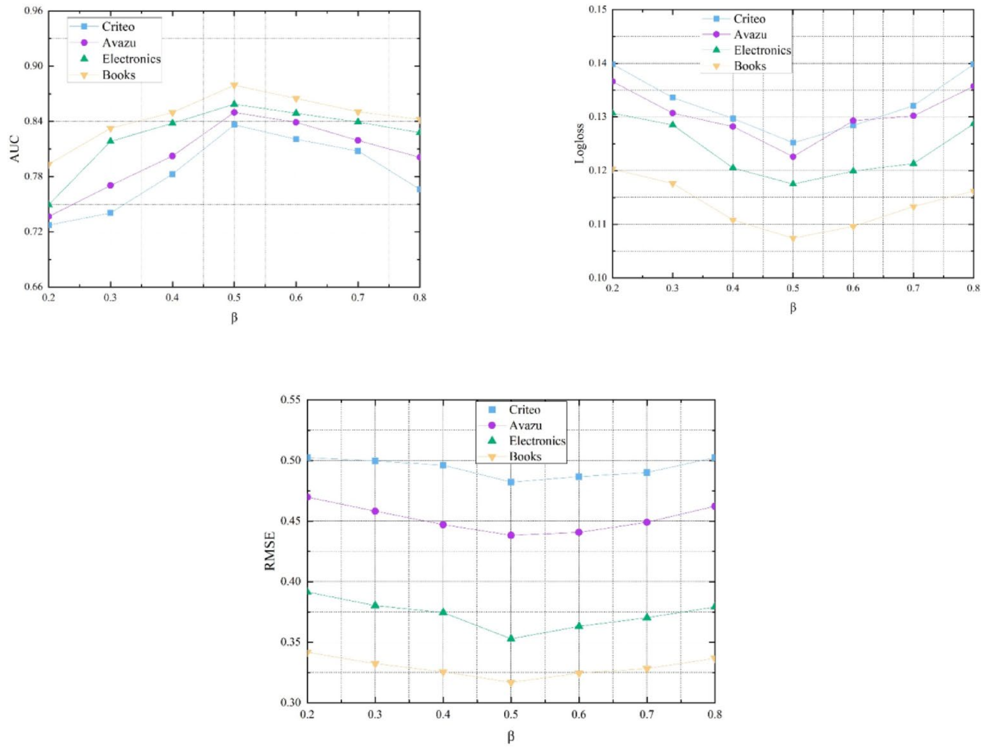

Fig. 5. Comparison of dropout rateβ

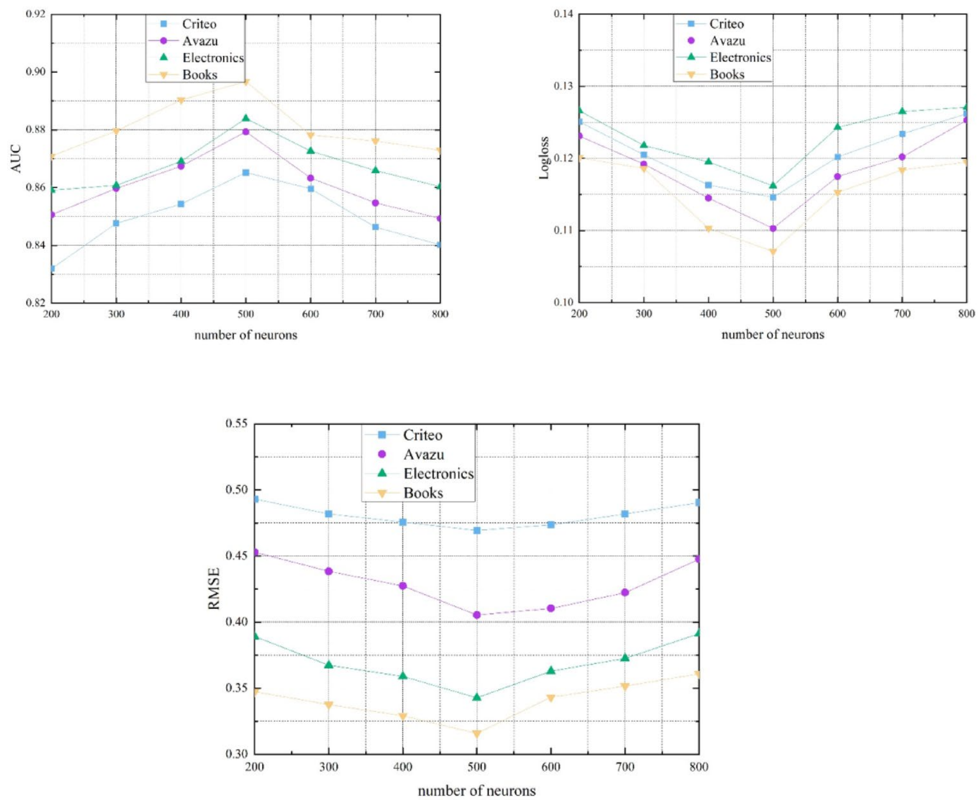

Fig. 6. Comparison of the neurons number.

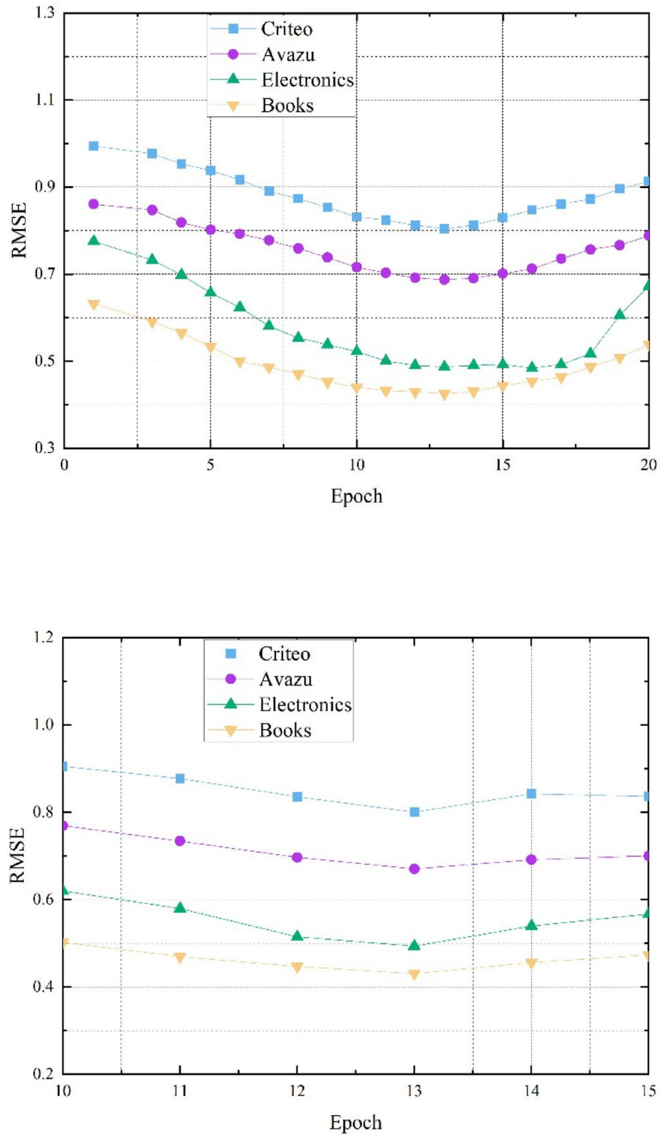

Fig. 7. The influence of the epoch.

According to test results, the SIFAN has better result in CTR prediction. In future research, in order to improve the predictive performance of model in future research, it is necessary to use more customer information and expand the relevant interest features, which can make more accurate judgments on customer behavior and interest, and thus obtain better prediction results. This article has achieved some valuable results in the research on this issue, but there are still some shortcomings. Therefore, improvements are needed in the future to better meet the performance requirements of practical applications and provide support for improving the application value of the obtained results.

Data availability

Correspondence and requests for materials should be addressed to Qianqian Wang. The datasets used during the current study available from the corresponding author on reasonable request.

Received: 19 April 2024; Accepted: 22 April 2025
Published online: 25 April 2025

References

1. Cheng, H. T. et al. Wide & deep learning for recommender systems. In *The Workshop on Deep Learning for Recommender Systems*, 7–10. (ACM, 2016).
2. Melville, P. & Sindhwani, V. Recommender systems. *Encyclopedia Mach. Learn.* **1**, 829–838 (2010).
3. Evans, D. S. The online advertising industry: economics, evolution, and privacy. *J. Economic Perspect.* **23** (3), 37–60 (2009).

4. Zhu, H. et al. Optimized cost per click in taobao display advertising. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2191–2200. (2017).
5. Chapelle, O., Manavoglu, E. & Rosales, R. Simple and scalable response prediction for display advertising. *ACM Trans. Intell. Syst. Technol. (TIST)*. **5**(4), 61 (2015).
6. Kumar, R. et al. Predicting clicks: CTR estimation of advertisements using logistic regression classifier. In *2015 IEEE International Advance Computing Conference (IACC)*, 1134–1138 (IEEE, 2015).
7. Rendle, S. Factorization machines. In *2010 IEEE International Conference on Data Mining*, 995–1000 (IEEE, 2010).
8. Juan, Y. et al. Field-aware factorization machines for CTR prediction. In *Proceedings of the 10th ACM Conference on Recommender Systems*, 43–50. (2016).
9. Zhang, C. & Lu, Y. Study on artificial intelligence: the state of the art and future prospects. *J. Industrial Inform. Integr.* **23**, 100224 (2021).
10. Zhao, S., Blaabjerg, F. & Wang, H. An overview of artificial intelligence applications for power electronics. *IEEE Trans. Power Electron.* **36**(4), 4633–4658 (2020).
11. Berg, S. et al. Ilastik: interactive machine learning for (bio) image analysis. *Nat. Methods.* **16**(12), 1226–1232 (2019).
12. Hagler, D. J. Jr et al. Image processing and analysis methods for the adolescent brain cognitive development study. *Neuroimage* **202**, 116091 (2019).
13. Chowdhary, K. R. & Chowdhary, K. R. Natural language processing. *Fundam. Artif. Intell.* 603–649. (2020).
14. Lauriola, I., Lavelli, A. & Aiolli, F. An introduction to deep learning in natural language processing: models, techniques, and tools. *Neurocomputing* **470**, 443–456 (2022).
15. Guo, H. et al. Deepfm: An end-to-end wide & deep learning framework for CTR prediction. arXiv preprint arXiv:1804.04950 (2018).
16. Qu, Y. et al. Product-based neural networks for customer response prediction. In *2016 IEEE 16th International Conference on Data Mining (ICDM)*, 1149–1154 (IEEE, 2016).
17. Yang, Y. et al. Operation-aware neural networks for customer response prediction. *Neural Netw.* **121**, 161–168 (2020).
18. Wang, R. et al. Deep & cross network for ad click predictions. In *Proceedings of the ADKDD'17*, 1–7 (2017).
19. Cheng, H. T. et al. Wide & deep learning for recommender systems. In *Proceedings of the 1st Workshop on Deep Learning for Recommender Systems*, 7–10. (2016).
20. Niu, Z., Zhong, G. & Yu, H. A review on the attention mechanism of deep learning. *Neurocomputing* **452**, 48–62 (2021).
21. Xiao, J. et al. Attentional factorization machines: learning the weight of feature interactions via attention networks. *ArXiv Preprint* (2017). arXiv:1708.04617.
22. Long, L., Yin, Y. & Huang, F. Hierarchical attention factorization machine for CTR prediction. In *International Conference on Database Systems for Advanced Applications*, 343–358 (Springer International Publishing, 2022).
23. Li, Z. et al. A survey of convolutional neural networks: analysis, applications, and prospects. *IEEE Trans. Neural Netw. Learn. Syst.* (2021).
24. Liu, H. et al. Category-Specific CNN for Visual-aware CTR Prediction at JD. com. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2686–2696 (2020).
25. Cui, Q. et al. MV-RNN: A multi-view recurrent neural network for sequential recommendation. *IEEE Trans. Knowl. Data Eng.* **32**(2), 317–331 (2018).
26. Zhou, G. et al. Deep interest network for click-through rate prediction. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 1059–1068 (2018).
27. Zhou, G. et al. Deep interest evolution network for click-through rate prediction. In *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 33, no. 01, 5941–5948. (2019).
28. Pi, Q. et al. Search-based customer interest modeling with lifelong sequential behavior data for click-through rate prediction. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, 2685–2692 (2020).
29. Li, J. et al. Neural attentive session-based recommendation. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, 1419–1428 (2017).
30. Wu, S. et al. Session-based recommendation with graph neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 33, no. 01, 346–353 (2019).
31. Liu, Q. et al. STAMP: short-term attention/memory priority model for session-based recommendation. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 1831–1839 (2018).
32. Grbovic, M. & Cheng, H. Real-time personalization using embeddings for search ranking at airbnb. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 311–320 (2018).
33. Wu, C. et al. Neural news recommendation with multi-head self-attention. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 6389–6394 (2019).
34. Dey, R. & Salem, F. M. Gate-variants of gated recurrent unit (GRU) neural networks. In *2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS)*, 1597–1600 (IEEE, 2017).
35. Haidong, S. et al. Enhanced deep gated recurrent unit and complex wavelet packet energy moment entropy for early fault prognosis of bearing. *Knowl. Based Syst.* **188**, 105022 (2020).
36. McAuley, J. et al. Image-based recommendations on styles and substitutes. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 43–52 (2015).
37. Bai, Z., Zhang, X. L. & Chen, J. Partial AUC optimization based deep speaker embeddings with class-center learning for text-independent speaker verification. In *ICASSP 2020–2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 6819–6823 (IEEE, 2020).
38. Chen, W. et al. A new design of evolutionary hybrid optimization of SVR model in predicting the blast-induced ground vibration. *Eng. Comput.* **37**(2), 1455–1471 (2021).
39. Feng, Y. et al. Deep session interest network for click-through rate prediction. preprint:1905.06482, (2019).
40. Bian, W. et al. CAN: feature co-action network for click-through rate prediction. In *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining*, 57–65 (2022).
41. Li, X. et al. CTRL: Connect Collaborative and Language Model for CTR Prediction.preprint: 2306.02841 (2023).
42. KimTae-Suk DONN: leveraging heterogeneous outer products for CTR prediction. *Neural Comput. Appl.* (2024).

Acknowledgements
This study has obtain support from the following grants: Natural Science Foundation of Shandong Province (ZR2021QF071); Development Programme for Young Innovation Teams in Shandong Higher Education Institutions (2024KJN022); Cultivation Fund of Shandong Women's University High-level Scientific Research Project (2020GSPSJ02).

Author contributions
W.and L. wrote the main manuscript text and Z.T. prepared all figuress. All authors reviewed the manuscript.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to Q.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.