Multimedia Tools and Applications
https://doi.org/10.1007/s11042-023-16083-5

Development of an offline OOH advertising recommendation system using negative sampling and deep interest network

Hyun-Woo Seo[1] · Soo-Hyeok Kim[1] · Sang-Gi Ryu[2] · Seung-Kyu Jo[2] · Su-Phil Cho[2] · Jong-Soo Sohn[2] · Chi-Ehyeon Lim[1]

Received: 25 May 2022 / Revised: 9 May 2023 / Accepted: 18 June 2023

Abstract
The out-of-home (OOH) advertising market has been operated exclusively following the know-how of salespeople. Thus, it is difficult to make scientific decisions and systematically provide various options to advertisers. In this regard, this study develops an OOH advertising recommendation system by analyzing past OOH history data. The OOH advertising allocation problem has the characteristics that the training data are implicit feedback, and only one advertisement can be posted per offline billboard. This study proposes a recommendation system suitable for OOH history data using negative sampling and Deep Interest Network. The proposed recommendation system showed a higher performance than excisting models used for comparison purposes, and the findings of this study present implications for solving similar recommendation problems.

Keywords Out-of-home (OOH) advertising · Recommendation system · Deep interest network · Negative sampling

1 Introduction

According to the 2019 advertising industry survey, the market size of digital out-of-home (OOH) advertising increased by 62.3% over the previous year, expanding the total OOH advertising market [15]. However, advertisers continue receiving proposals for OOH advertising media by the know-how of salespeople. Due to the different sales methods of each salesperson, the OOH advertising market has failed to present scientific and systematic options. Furthermore, because advertisers must execute advertising on the allocated media during the contract term, inappropriate media for the advertiser can cause a great loss. In other words, the closed OOH advertising industry has not provided advertisers with an opportunity to sufficiently consider the optimal advertising options. Therefore, it is necessary to build a data-driven OOH advertising recommendation system for objective and

✉ Jong-Soo Sohn
jongsoo.sohn@cj.net

[1] Department of Industrial Engineering, UNIST, Ulsan, Republic of Korea

[2] CJ AI Center, CJ Corporation, Seoul, Republic of Korea

scientific advertising. As shown at the top of Fig. 1, the current OOH advertising business relies on the personal know-how of salespeople to recommend advertising, making it difficult to approach scientifically based on data. Therefore, as shown at the bottom of Fig. 1, if customers directly access the recommendation service and introduce a recommendation system based on the analysis of the surrounding commercial area where OOH is placed, more detailed OOH recommendations are possible, and the cost of salesperson activities can be reduced.

Moreover, it is difficult for OOH history data to reflect the specific preferences of users, unlike explicit feedback such as ratings for films [11]. Hence, the OOH advertising recommendation system must infer preferences from the past advertising history of advertisers. However, it is difficult for them to advertise on occupied billboards because the OOH advertising allocation problem has the characteristic of "exclusive occupation," implying that only one advertisement can be posted per billboard. In other words, the media not used by an advertiser include the media that have not been provided as options due to "exclusive occupation" and the media not preferred by the advertiser. Thus, it is necessary to determine the media the advertiser will not prefer among the media without advertising records by the advertiser. Furthermore, the OOH advertising media should be proposed considering various metadata such as the industry category of the advertiser and the region category of the media.

This study developed a recommendation system for the OOH advertising industry by analyzing OOH history data. This study used the Deep Interest Network (DIN) [22] among deep learning-based click-through rate (CTR) prediction models. Furthermore, a negative sampling method suitable for OOH advertising recommendation was developed by reflecting the exclusive occupation characteristic of OOH advertising and the collaborative filtering idea that "advertisers of the same industry prefer similar media." Then, the model was trained using the developed negative sampling method. The DIN predicts the preferences of OOH advertising media considering the correlations among the advertiser's metadata, OOH advertising history, and candidate OOH advertising media. Hence, the proposed model can recommend the optimal OOH advertising media by reflecting the OOH history and considering the characteristics of the advertiser and media. Moreover, it can make reliable recommendations by quantifying the correlation between the recommended OOH advertising media and the OOH history and providing the result to the advertiser. Negative sampling distinguishes the media not preferred by the advertiser by effectively excluding the media that have not been suggested due to exclusive occupation among the media without advertising records. The DIN trained by negative sampling reflects the characteristics of OOH advertising.

The proposed offline advertising recommendation system based on negative sampling and DIN was verified through the actual OOH history data. A comparative experiment with

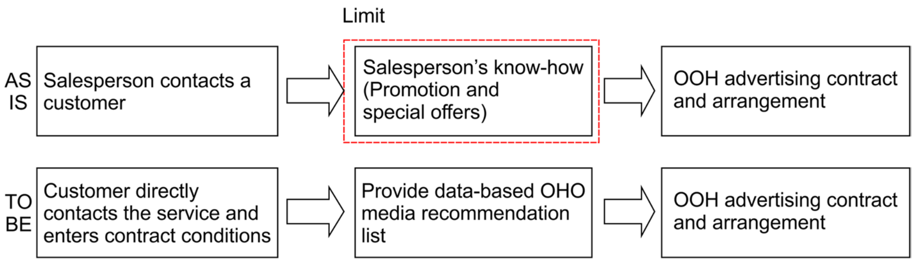

Fig. 1 OOH advertising industry process

benchmark models showed that the proposed system has a higher performance than existing models. The proposed system demonstrated the possibility of improving the explanatory power and reliability of the recommendation system by presenting the correlation information on past OOH advertising records through the local activation network of DIN. Moreover, it proved to be robust to the cold start problem by showing high performance when inferring new advertisers without advertising records. The negative sampling method developed for the OOH advertising recommendation problem improved the performance of the DIN and benchmark model compared to the existing negative sampling method, thus experimentally verifying the appropriateness of the model for the OOH advertising recommendation problem. This study is the first to develop a recommendation system for allocating offline OOH advertising; thus, this study will provide the academic circle with suggestions for solving similar offline recommendation problems.

This paper is organized as follows. Section 2 analyzes related works on offline recommendation and negative sampling for recommendation model training. Section 3 describes the recommendation system for OOH advertising media. Section 4 presents the results of comparative experiment based on real data. Finally, Section 5 presents the discussion and conclusions.

2 Related works

2.1 Offline product recommendation system

The offline product recommendation system considers the physical characteristics of the offline environment with the purchase history of users. Kim and Jeong (2017) [12] suggested an offline method of recommending brands to customers through the brand purchase history data of customers in an offline shopping mall. The indoor location, movement flow, and stay duration were inferred from purchase history, and brands were recommended to customers by applying collaborative filtering. Kim and Jeong (2016) [13] recommends products appropriate for customers by using the customer's dynamic context data in an offline shopping mall. Choi and Lee (2006) [2] analyzed the movement patterns of users to recommend personalized offline products. Several studies on offline product recommendation systems have recommended items based on the purchase history of individual users. However, studies on the recommendation of OOH advertising media are insufficient. Various metadata regarding corporate users should be considered because OOH advertising is recommended to corporate users. Moreover, various physical characteristics of the offline environment should be considered for OOH advertising media, such as local information and the floating population.

2.2 Negative sampling for training the recommendation system based on purchase data

The recommendation system based on offline purchase data is challenging to train due to limiting no negative feedback in purchase data [16]. Negative sampling methods have been developed to identify non-preferred items in implicit feedback such as purchase data. For example, the general negative sampling method extracts items that are not preferred by the user according to an even distribution from items with no purchase history [9, 14, 18]. He et al. (2016) [10] extracted non-preferred items according to a probability distribution based on the total purchase frequency. Zhang et al. (2013) [20] trained a collaborative filtering model by re-using non-preferred items estimated from the recommendation list of collaborative filtering. Furthermore, Ding et al. (2019) [3] and Wang et al. (2020) [17] trained a negative sampler model with a recommendation system by using the user-item exposure data and item characteristic information graph. However, it is practically difficult to train a recommendation system and a sampler model together because fewer offline OOH history data exist due to a long advertisement cycle. Furthermore, the non-preferred media must be extracted by reflecting the industry relationship between the advertiser and the exclusive occupation, a characteristic of the OOH advertising market.

2.3 Hybrid recommendar systems

In area of recommender system, traditional methods are used by collaborative filtering or user based filtering. These recommender systems are sufficiently proven in further researches. However traditional methods hard to be applied in personalized situation. Therefore in recent, various hybrid recommender systems are proposed.

Forouzandeh et al. proposed tourism recommender system by Artificial Bee Colony Algorithm and Fuzzy Topsis model [4, 5]. And Berahmand et al. showed random work based method to apply user preference in complex network [1]. Almost of these researches are useful to apply user preferences in recommendation system. Our research as shown in this paper is a type of hybrid recommender system and proposed method overcomes cold start problem.

3 Methodology of OOH advertising recommendation using negative sampling and DIN

3.1 Deep interest network

We present a recommendation system for the OOH advertising industry by leveraging OOH history data. The proposed system utilizes the Deep Interest Network (DIN) trained by negative sampling, which offers significant advantages for identifying non-preferred items in implicit feedback data. Negative sampling has been widely used in various recommendation systems, due to its effectiveness in identifying non-preferred items [8]. By using negative sampling, we can exclude the media that have not been suggested due to exclusive occupation among the media without advertising records, which is a crucial feature of the OOH advertising market.

In addition, the DIN is a state-of-the-art deep learning-based click-through rate (CTR) prediction model, which has been proven to outperform other CTR prediction models [21]. The DIN is designed to capture complex correlations among the user's historical behaviors, candidate items, and the context of the recommendation. By using the DIN trained by negative sampling, we can effectively capture the preferences of OOH advertising media considering the correlations among the advertiser's metadata, OOH advertising history, and candidate OOH advertising media. This enables us to make accurate recommendations that reflect the characteristics of the OOH advertising market.

The OOH history data include customer advertisers, the OOH advertising media used by the advertisers, and various continuous and categorical metadata expressing the advertiser and media. Collaborative filtering uses only the advertisers and the history of

OOH advertising media used by the advertisers- Thus, it is challenging for collaborative filtering to reflect the relationships between similar advertisers and the OOH advertising media due to various metadata. Furthermore, the preferences for candidate media should be inferred by considering the correlations between the media in the OOH history of the advertiser and the candidate OOH advertising media.

Therefore, this study used DIN, a deep learning-based CTR prediction model, for offline OOH advertising recommendations. CTR is the probability of clicking an item in an online recommendation and is generally interpreted as a probability of preferring an item. The DIN calculates candidate OOH advertising media preferences after receiving the advertiser and the advertiser's metadata, OOH history, and the candidate OOH advertising media as input. The DIN calculates the correlation between the OOH history and the candidate OOH advertising media and combines them into an adaptively reflected expression vector. The DIN can calculate the OOH advertising preferences of advertisers by combining the advertiser information, OOH history, and candidate OOH advertising media into an expression vector and reflecting complex relationships between various data through a neural network. The structure of the DIN consists of the following three subnetworks (Fig. 2).

The structure of the Deep Interest Network shown in Fig. 2 consists of three main modules. The inference process starts from the bottom of Fig. 2 and completes at the top. Among these modules, 1) the Embedding Network takes input values such as user preference, user behavior, candidate Ad, context, and passes them through the Concatenation process to 2) the Local Activation Network. In 2), the Activation Unit calculates the weight of each product. 3) In the Multi-Layer Perceptron, Deep Learning is performed using the values computed in 2). Following sections describe detailed explanation of each step.

1) **Embedding network**

The DIN receives advertisers, OOH advertising media, and categorical metadata by one-hot encoding. Because the input value is a high-order binary vector, the embedding network converts each input parameter into a low-order dense vector. For example, the ith input parameter t_i with K_i categories has an embedding network of $\left[w_1{}^i, w_2{}^i, \ldots, w_{K_i}{}^i\right] \in R^{D \times K_i}$.

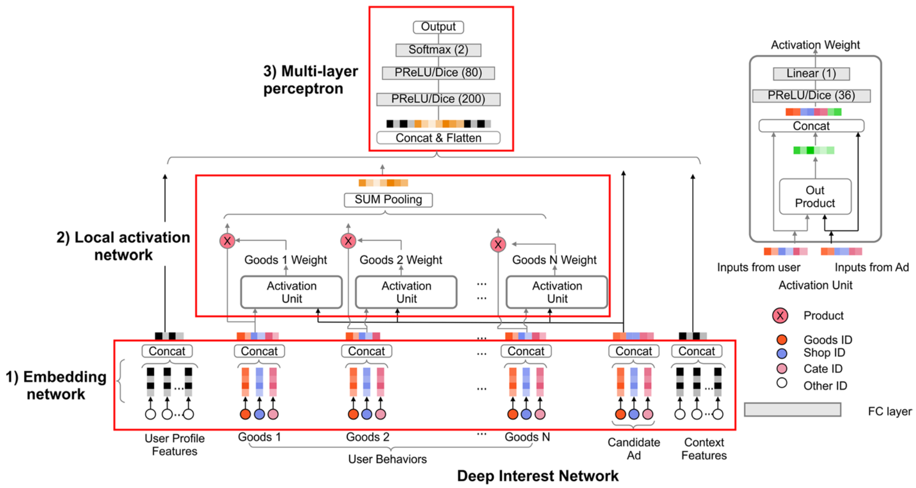

Fig. 2 Deep Interest Network [22]

Each categorical input parameter is converted to a D-order dense vector through embedding network.

2) **Local activation network**

The local activation network is applied to the OOH history of an advertiser to calculate the correlation between each history and the candidate OOH advertising media. Furthermore, each history is combined into one expression vector by weighted sum according to correlation. In other words, the advertisement history vector v_U for the candidate medium A is calculated as follows:

$$(A) = f\left(v_A, e_1, e_2, \dots, e_H\right) = \sum\nolimits_{j=1}^{H} a\left(e_j, v_A\right)e_j = \sum\nolimits_{j=1}^{H} w_j e_j \tag{1}$$

Where $\{e_1,e_2,\dots,e_H\}$ is the list of OOH history embedding vectors of the advertiser, and v_A is the embedding vector of the candidate OOH advertising media. The local activation network $a(.)$ is a feed forward network that outputs the activation weight. The OOH history is converted to another vector expression depending on the candidate medium. $a(.)$ reflects the explicit information regarding correlation modeling by calculating the outer product from two input vectors and inputting it to the next layer.

3) **Multi-layer perceptron**

The advertiser vector converted to an embedding network, the advertising vector calculated by the local activation network, and the candidate OOH advertising media vector are combined into one vector. Then, the vector is inputted to the multi-layer perceptron. The multi-layer perceptron calculates the preference for the candidate OOH advertising of the advertiser by the combined dense vector. Here, the negative log–likelihood L is used as the objective function. The function is defined as follows:

$$\frac{-1}{N}\sum\nolimits_{(x,y)\in S}(ylogp(x) + (1-y)log(1-p(x))) \tag{2}$$

where S is the training set of the size N, x is the input value of the model, $y \in \{0, 1\}$, and $p(x)$ is the predicted value of the preference for the candidate OOH advertising medium. L is a loss function generally used in a deep learning-based binary classification problem.

3.2 Negative sampling that excludes preferred and occupied media of the same industry

The OOH advertising allocation problem has the characteristic of exclusive occupation, implying that only one advertisement can be posted per billboard. When an advertising contract is signed, the billboards occupied during the target advertising period will not be provided to the customer advertiser. Furthermore, this study assumes that "advertisers of the same industry prefer similar media" under the logic that the advertisers of the same industry target customers of similar ages and occupations. This is similar to the assumption of collaborative filtering that "users with similar characteristics have similar preferences". By reflecting the characteristic of exclusive occupation and the assumption of collaborative filtering, this study proposes a negative sampling that excludes media preferred and occupied in the same industry as explained below.

The negative sampling excluding the preferred and occupied media of the same industry excludes from the candidate negative samples the media that have been used for OOH advertising in the same industry using the industry metadata of advertisers. Using the advertisement contract term included in the OOH advertising purchase history data, the media already occupied during the target advertising period in each OOH advertising purchase history of the advertiser is excluded from the candidate negative samples. This process reflects the situation that occupied media cannot be considered because they cannot be provided as an option due to exclusive occupation when the advertiser performs the corresponding OOH advertising. In other words, the proposed negative sampling method reflects the sales characteristics of the OOH advertising market and identifies non-preferred media by effectively excluding media the advertiser may prefer but has not been provided as an option. The DIN trains the preference and non-preference patterns of the advertiser by receiving training data generated by the proposed negative sampling as input (Table 1).

4 Evaluation

This section shows comprehensive descriptions of the experiment conducted to evaluate the proposed recommendation system. This experiment was conducted based on the real OOH history data, and the proposed recommendation system showed a higher performance than the existing model.

4.1 OOH history data

The OOH history data are from a contract from January 2016 to March 2021, including industries of advertisers, advertiser IDs, IDs of the contracted OOH advertising media, areas of the OOH advertising media, and contract term. In detail, these data include 542 advertisers, 61 media, 94 industries of advertisers, 12 media areas, and a total of 2869

Table 1 Preferred and occupied media of the same industry are excluded

Algorithm 1) Preferred and occupied media of the same industry are excluded.
Negative sampling
Data: out-of-home advertising history
Input: num_negative t
Output: out-of-home advertising history with negative samples
For each user i **do**
cand_neg_media = non-advertised media of user i
cand_neg_media = cand_neg_media - advertised media of users in the same industry of user i
For each media history j **of user** i **do**
non_occupied_media = neg_media – occupied media at advertising period of a media history j
neg_sample = select t samples uniformly from non_occupied_media
append neg_sample to out-of-home advertising history
return out-of-home advertising history

OOH history data. Among them, 133 advertisers have only one advertisement history (Table 2).

The training and test data were divided by the contract order. If the advertising history of an advertiser is $(b_1, b_2, ..., b_n)$ the advertisement history of b_k, $k = 1, 2, ..., n - 1$ is defined as training data, and the last advertising history b_n of the advertiser is defined as test data. Furthermore, non-preferred items at the advertising time were added to the training data by performing negative sampling for each advertising history b_k of the advertiser. In this experiment, four negative samples were selected. Because the DIN receives past records as input, it was trained to match the preference of b_k when advertisement history data until the $k - 1$ th were given in the advertising history of each advertiser.

4.2 Comparison model

CF-IF [11]: Collaborative filtering for implicit feedback (CF-IF) is a representative model that applies collaborative filtering to implicit feedback. Collaborative filtering is a traditional recommendation system based on machine learning. The user and item latent vectors are trained by decomposing the user-item interaction matrix based on matrix factorization (MF).

NCF [7]: Neural collaborative filtering (NCF) is a model expanding collaborative filtering, a representative model of the implicit feedback-based recommendation system, into a deep learning-based model. The preference for items is calculated by inputting the user and item latent vectors to MF and MLP to reflect the nonlinear relationship of user and item information. NCF calculates preferences using the deep interactions of the user and item information in MLP, as well as the shallow interactions in MF.

DeepFM [8]: Deep factorization machines (DeepFM) predicts the preferences for items by integrating the shallow interaction between parameters of FM (factorization machines) and the deep interactions of DNN (deep neural network). In contrast to NCF, which receives only user and item vectors as input, DeepFM can calculate the preferences for items by freely receiving metadata for user and item as input.

4.3 Evaluation method

To evaluate proposed method, we use recall, P@3(Precision at 3) [19], and MAP@3(Mean Average Precision at 3) [6] as evaluation metrics to assess the performance of our proposed recommendation system for the OOH advertising industry. Recall measures the proportion of relevant items that are correctly recommended, which is important for ensuring that the recommended OOH advertising media aligns with the advertiser's goals. P@3 measures the precision of the top three recommended items, which is particularly relevant in OOH advertising as

Table 2 OOH history sample data

Advertiser ID	Industry	Start_ date	End_ date	Medium ID	Area
A	Hospital	18.03.01	18.08.31	item_1	Hanam-si
A	Hospital	19.02.01	19.03.31	item_3	Goyang-si
B	Culture	19.03.01	20.02.28	item_2	Gangnam-gu
…	…	…	…	…	…

advertisers often have a limited budget and need to focus their efforts on a small number of high-performing media. Lastly, MAP@3 provides an overall assessment of the ranking quality of the recommended items, taking into account the order of recommendation and the relevance of each recommended item. These metrics have been widely used in the evaluation of recommendation systems. By using these metrics, we can quantitatively measure the effectiveness of our proposed method and compare it with other state-of-the-art recommendation systems.

To determine the suitability of the proposed negative sampling for the OOH advertising recommendation problem, the NCF, DeepFM, and DIN were trained by applying random sampling and the proposed negative sampling. The model trained by random sampling was named'model name-RS', and the model trained by the proposed negative sampling was named 'model name-NS'. Negative sampling was not applied to CF-IF because it assumes all items with no interaction as non-preferred items.

4.4 Comparison of the evaluation indices

Table 3 shows the recall, P@3, and MAP@3 of the comparison models. The proposed recommendation system DIN-NS shows the highest recall at 0.929. CF-IF shows a relatively high recall at 0.860. This is because several advertisers tend to advertise on the billboards they have advertised in the past. However, it is not suitable for the OOH advertising recommendation system because it showed low performance in providing three recommendation lists. Among the deep learning-based models, DeepFM showed the highest recall, followed by NCF and DIN. This suggests that inputting the local category metadata for the industry and item of the advertiser in the model and reflecting the history contribute to improving the inferring accuracy of the OOH advertising preferences. Furthermore, DIN-NS shows the highest P@3 and MAP@3. This result implies that DIN-NS can provide appropriate OOH advertising lists in the OOH advertising industry, where a few OOH advertising lists are given to advertisers.

Table 3 shows that NCF-NS, DeepFM-NS, and DIN-NS have higher performances in general than NCF-RS, DeepFM-RS, and DIN-RS. As seen in the table, training the model by selecting non-preferred billboards of the advertiser, considering the characteristics of the OOH advertising market, contributes more to improving the model performance than randomly extracting non-preferred billboards among the billboards with no OOH history.

4.5 Correlation with OOH history

The local activation network of the DIN selectively combines the history information for the candidate OOH advertising by calculating the correlation between the OOH history and the candidate OOH advertising by feed forward network. The proposed recommendation

Table 3 Evaluation results for the compared models

	Recall	P@3	MAP@3
CF-IF	0.860	0.219	0.240
NCF-RS	0.805	0.315	0.313
DeepFM-RS	0.447	0.218	0.234
DIN-RS	0.775	0.277	0.283
NCF-NS	0.916	0.313	0.315
DeepFM-NS	0.558	0.248	0.257
DIN-NS	0.929	0.338	0.363

system can provide the advertiser recommendations through the activation weight of the local activation network.

Figure 3 shows the activation weight used for calculating the preference of one advertiser for the candidate OOH advertising in test data. For example, media_56 has a preference of 0.995 with the correlations of 0.2151, 0.2341, 0.2667, and 0.2884 with each history of the advertiser. The explanatory power and reliability of OOH advertising recommendation can be improved by presenting a quantitative basis through the model and a qualitative basis such as the area information of billboard regarding correlation with past history through the activation weight of DIN.

4.6 Evaluation in cold start

Cold start refers to the initial state of a new product, service, or system where there are relatively few users or customers who are utilizing or interacting with it. The cold start problem arises because of the lack of initial data, which can lead to reduced performance of recommendation systems and other AI algorithms.

For instance, there may not be enough data on OOH customers who are visiting for the first time, leading the system to struggle in providing appropriate recommendations or even providing irrelevant recommendations to users. Therefore, we evaluated whether our proposed method can effectively address the cold start problem.

A recommendation system that is robust to cold start should make suitable recommendations for new advertisers only with the industry category of the advertiser. This experiment evaluated whether the proposed recommendation system recommends suitable OOH advertising to new advertisers. The model was trained by excluding 133 advertisers with only one advertising history in the OOH history data. The excluded advertisers were assumed as new advertisers with no advertising history.

Table 4 lists the recalls of the proposed recommendation system for a new-user test, targeting the first preferred media of new advertisers with no advertising history and a test set targeting the final preferred media of existing advertisers. The recall of the new-user test set was 0.887, indicating that appropriate OOH advertising can be proposed to a new advertiser only with the industry category of the advertiser. In other words, the proposed recommendation system is robust to cold start. Furthermore, even if the data of advertisers with one advertising history were excluded from training, the recommendation system for the existing advertisers shows a high recall of 0.953.

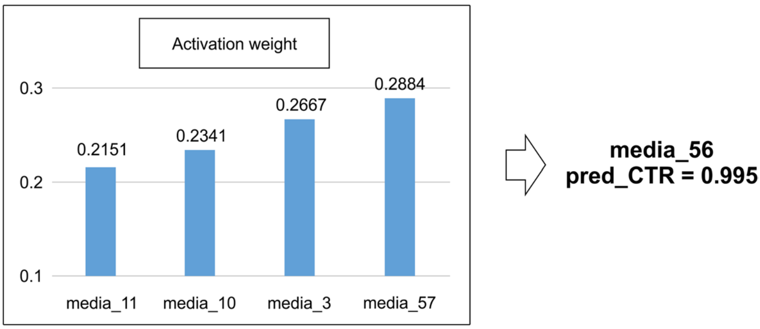

Fig. 3 Prediction of the activation weight and candidate OOH advertising

Table 4 Result of the new user test

	New-user test set	Test set
Recall	0.887	0.953

5 Discussion

In this paper, we proposed a recommendation system for OOH advertising that uses deep learning and negative sampling to improve performance. Our proposed system, DIN-NS, showed higher recall, P@3, and MAP@3 compared to existing models, including CF-IF, NCF, and DeepFM. Furthermore, the proposed system was robust to the cold start problem, which is a significant challenge in recommendation systems.

The success of our proposed system can be attributed to several factors. First, negative sampling was applied to training data, which improved the model's ability to identify relevant OOH advertising media for each advertiser. The proposed negative sampling considers the characteristics of the OOH advertising market, where advertisers often prefer specific OOH advertising media and avoid others. By selecting non-preferred OOH advertising media for each advertiser, our proposed system can capture their preferences and recommend suitable OOH advertising media. Second, the deep learning-based approach used in our proposed system allowed us to model complex relationships between the advertiser, the OOH advertising media, and the industry. Specifically, we used DIN, which is designed to capture the user's interests and the item's properties. In our case, the user represents the advertiser, and the item represents the OOH advertising media. DIN also considers the industry category of the advertiser and the area of the OOH advertising media to enhance its recommendation accuracy. Third, we evaluated our proposed system using recall, P@3, and MAP@3, which are widely used in the evaluation of recommendation systems. By using these metrics, we were able to quantitatively measure the effectiveness of our proposed method and compare it with other state-of-the-art recommendation systems. The results showed that our proposed system outperformed existing models in terms of recall, P@3, and MAP@3.

One limitation of our study is that we only used one dataset for our evaluation. Future studies should consider using multiple datasets to validate the performance of our proposed system. Furthermore, our study only considered the OOH advertising industry in South Korea. Future studies could expand the scope of the study to other countries or industries to generalize the findings.

Proposed recommendation system shown in this paper for OOH advertising using deep learning and negative sampling showed higher performance than existing models and was robust to the cold start problem. Our study demonstrates the effectiveness of negative sampling and deep learning in recommendation systems and provides insights into improving recommendation accuracy in the OOH advertising industry.

6 Conclusions

Systematic and scientific OOH advertising contracts and arrangements have not been made in the OOH advertising industry despite the importance of the industry. Therefore, this study developed an offline OOH advertising recommendation system based on OOH history data. This study used DIN to reflect the metadata of the advertisers and OOH advertising and calculate preferences through the correlation between the OOH history and the candidate OOH advertising. The model was trained through negative sampling, which reflected the characteristics of the OOH advertising industry. The use of DIN trained by negative sampling in our recommendation system allows us to capture the preferences of OOH advertising media effectively and provide accurate recommendations. The proposed methodology contributes to the advancement of the OOH advertising industry by providing an effective and efficient way to make recommendations to advertisers.

In real OOH history data, the proposed recommendation system showed a higher performance than the previously proposed models. The explanatory power and reliability of the recommendation model were improved through the local activation network of DIN, and the proposed model also showed robustness to a cold start. The proposed recommendation system is expected to present suggestions for similar offline recommendation problems.

In follow-up research, the model will be trained by including additional metadata and the area category information of the OOH advertising. For example, the billboard preferences will be calculated by reflecting the correlation between the billboards and the population information added to this model from various angles. Furthermore, the proposed negative sampling will be developed into a probabilistic sampler model to improve the performance of the recommendation system. Lastly, the proposed recommendation system will be evaluated by applying the findings of this study to other areas such as offline store product recommendations, which should reflect physical metadata of the environment.

Authors' contributions Hyun-Woo Seo: Methodology, Writing – original draft; Soo-Hyeok Kim: Investigation; Sang-Gi Ryu: Project administration, Software, Supervision, Writing – review & editing; Seung-Kyu Jo: Formal Analysis, Writing – review & editing; Su-Phil Cho: Data curation, Writing – review & editing; Jong-Soo Sohn: Resources, Validation, Chi-Ehyeon Lim: Resources, Validation.

Data availability Data supporting the findings of this study is available from [CJ Olive Networks], but the availability of these data is limited, which is not publicly available because it has been used as a license for the current study.

Declarations

Competing interests The authors have no competing interests to declare that are relevant to the content of this article.

Ethics approval and Consent Not applicable.

References

1. Berahmand K, Nasiri E, Forouzandeh S, Li Y (2022) A preference random walk algorithm for link prediction through mutual influence nodes in complex networks. J King Saud Univ-Comput Inf Sci 34(8):5375–5387

2. Choi YH, Lee SY (2006) Users' moving patterns analysis for personalized product recommendation in offline shopping malls. J Korean Inst Intell Syst 16(2):185–190. https://doi.org/10.5391/JKIIS.2006.16.2.185
3. Ding, J., Quan, Y., He, X., Li, Y., & Jin, D. (2019) Reinforced Negative Sampling for Recommendation with Exposure Data. Proceeding of IJCAI 2019:2230–2236
4. Forouzandeh S, Berahmand K, Nasiri E, Rostami M (2021) A hotel recommender system for tourists using the Artificial Bee Colony Algorithm and Fuzzy TOPSIS Model: a case study of tripadvisor. Int J Inf Technol Decis Mak 20(01):399–429
5. Forouzandeh S, Rostami M, Berahmand K (2022) A hybrid method for recommendation systems based on tourism with an evolutionary algorithm and topsis model. Fuzzy Inf Eng 14(1):26–50
6. Gemmeke JF, Ellis DPW, Freedman D, Jansen A, Lawrence W, Moore RC, Plakal M, Ritter M (2017) Audio set: An ontology and human-labeled dataset for audio events. 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP): 776–780
7. Guo H, Tang R, Ye Y, Li Z, He X (2017) DeepFM: a factorization-machine based neural network for CTR prediction. arXiv preprint arXiv:1703.04247 . Accessed 23 June 2021
8. He X, Liao L, Zhang H, Nie L, Hu X, Chua T-S (2017). Neural Collaborative Filtering. In Proceedings of the 26th International Conference on World Wide Web (WWW '17). International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE: 173–182
9. He X, Liao L, Zhang H, Nie L, Hu X, Chua TS (2017) Neural collaborative filtering. In: Proceedings of the 26th international conference on World Wide Web: 173–182
10. He X, Zhang H, Kan MY, Chua TS (2016) Fast matrix factorization for online recommendation with implicit feedback. In: Proceedings of the 39th international ACM SIGIR conference on Research and Development in Information Retrieval: 549–558
11. Hu Y, Koren Y, Volinsky C (2008) Collaborative filtering for implicit feedback datasets. In: vol 2008 Eighth IEEE International Conference on Data Mining. IEEE Publications: 263–272
12. Kim NK, Jeong SB (2017) A study on customized brand recommendation system in offline shopping malls using collaboration filtering. Proceeding of 19th Joint Business Adminstration Conference of Korea 2019: 600–619.
13. Kim NK, Jeong SB (2016) A Study on Customized Brand Recommendation based on Customer Behavior for off-line Shopping Malls. J Inf Technol Appl Manag 23(4):55–70
14. Liu Y, Wang S, Khan MS, He J (2018) A novel deep hybrid recommender system based on autoencoder with neural collaborative filtering. Big Data Min Anal 1(3):211–221. https://doi.org/10.26599/BDMA.2018.9020019
15. Ministry of Culture, Sports and Tourism (2020) 2019 advertising industry survey. Searched on https://www.mcst.go.kr/kor/s_policy/dept/deptView.jsp?pSeq=1300&pDataCD=0417000000&pType=08. Accessed 23 June 2021
16. Pan R, Zhou Y, Cao B, Liu NN, Lukose R, Scholz M, Yang Q (2008) One-class collaborative filtering. In: vol 2008 Eighth IEEE International Conference on Data Mining. IEEE Publications: 502–511
17. Wang X, Xu Y, He X, Cao Y, Wang M, Chua TS (2020) Reinforced negative sampling over knowledge graph for recommendation. Proc Web Conf 2020:99–109
18. Xu Z, Chen C, Lukasiewicz T, Miao Y, Meng X (2016) Tag-aware personalized recommendation using a deep-semantic similarity model with negative sampling. In: Proceedings of the 25th ACM international on conference on information and knowledge management. pp 1921–1924
19. Yih W-T (2009) Learning term-weighting functions for similarity measures. In: Proceedings of the 2009 conference on empirical methods in natural language processing. pp 793–802
20. Zhang W, Chen T, Wang J, Yu Y (2013) Optimizing top-n collaborative filtering via dynamic negative item sampling. In: Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval. pp 785–788
21. Zhou G, Zhu X, Song C, Fan Y, Zhu H, Ma X, Yan Y, Jin J, Li H, Gai K (2018) Deep Interest Network for Click-Through Rate Prediction. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '18). Association for Computing Machinery, New York, USA. pp 1059–1068
22. Zhou G, Zhu X, Song C, Fan Y, Zhu H, Ma X,Yan Y, Jin J, Li H, Gai K (2018) Deep interest network for click-through rate prediction. In: Proceedings of the 24th ACM SIGKDD international conference on Knowledge Discovery & Data Mining. pp 1059–1068. https://doi.org/10.1145/3219819.3219823