Click-Through Rate Prediction and Customer Representation

Subjects: Computer Science, Interdisciplinary Applications | Business

Contributor: Miguel Alves Gomes , Richard Meyes , Philipp Meisen , Tobias Meisen

Click-Through Rate Prediction is a significant subject in e-commerce for both academia and industry. In order to accurately predict the customer's click intent, it is necessary to create a personalized customer representation. Learning such a customer representation is currently state-of-the-art.

click-through rate prediction   e-commerce   customer representation

1. Introduction

Recently, large deep learning models have dominated various domains such as natural language processing (NLP) and computer vision (CV) in academia and industry. Since the introduction of the transformer model [1] in 2017, they have been repeatedly archiving state-of-the-art results. Recent examples like ChatGPT, GPT-3 [2], or Dall-E [3] show what such deep models are capable of. A similar trend can be observed in e-commerce, especially with recommender models like "Wide & Deep" [4] or Bert4Rec [5] and Click-Through Rate (CTR) prediction (CTR-P) models like "Deep & Cross" [6].

In the last years, CTR-P became a core task in online advertisement (also called ads) [7][8]. This is mainly because search engines, and especially recommender systems, are playing a significant role in e-commerce businesses [9][10][11][12]. Furthermore, predicting CTR accurately leads to a better user experience which has been shown to have a great impact on business effectiveness [8][13]. Additionally, CTR is a key performance indicator for online ads and therefore, its prediction influences the ranking and price for online ads and revenue sponsored search [13][14][15]. Although there is a huge amount of data in the e-commerce sector, unlike natural language or images which have recurring patterns, customer behavior is subject to constant change as it is highly dependent on a variety of factors such as season, inflation, and local as well as global developments. In addition, the data are typically use case- and user-specific and are therefore limited in their ability to be shared across organizations.

Companies have limited resources and need to plan them accordingly [16]. Consequently, in the e-commerce sector, companies should ideally only use their resources on reactive customers, e.g., only display recommendations to those customers who are most likely to click on them. Lastly, advertising and recommendations can lead to negative experiences for certain customers, resulting in negative attitudes towards the operating company. This leads to shorter visit duration, fewer visits, fewer referral opportunities, and increased negative word-of-mouth. Therefore, it is crucial to only display advertising and recommendations when success is

probable. Therefore, it is of great importance for a business to understand its customers' intentions and engage them with personalized targeting.

2. Approaching Click-Through Rate Prediction

CTR-P received a lot of attention in industry and academia in the past years. It is approached as a binary classification problem, where the probability of an item click should be predicted regardless of the use case, e.g., retrieved item in a search, clicked ad, or clicked product. In the literature, there is not one CTR-P use case, but multiple kinds of use cases. For example, Chen et al. [17], Ge et al. [18], and Fan et al. [9] propose a CTR-P model to optimize the retrieved items of a search engine. Others predict the CTR for shown ads [6][19] or products in general [10][20]. **Table 1** presents a comprehensive overview of state-of-the-art CTR-P approaches including information of the authorship, publication year, proposed approach, and used dataset. All approaches are based on deep neural networks which means mixtures and ensembles of multi-layer perceptrons, recurrent layers, and attention layers that should capture customers' behavioral information. Furthermore, all models contain an embedding input layer to embed available information, which is usually given by the use case and/or selected by the data engineers. Typical input information is the user id, target item id, additional user information, and additional target item information. The DIN [21], DIEN [10], TIEN [22], and MARN [8] approaches use sequential activity information, which Alves Gomes et al. [26] also rely on their approach. They propose a decoupled approach consisting of an activity embedding that learns historical customer behavior from context in a self-supervised manner, and an LSTM that learns to predict whether the customer will click on a product or recommendation based on the embedded behavior. CTR approaches are evaluated on different datasets, some publications and approaches rely only on closed data [13][18][23][24][25] which are not included in **Table 1**. Others, as shown in **Table 1**, use openly available datasets to evaluate their approach. Of all the reviewed publications, the Amazon review dataset is the most used.

Table 1. Overview of publications proposing CTR-P approaches with information on the datasets.

Author	Year	Approach	Dataset
Alves Gomes et al. [26]	2024	Decoupled Embedding + LSTM	Amazon Review
			closed
Fan et al. [9]	2022	RACP	Avito
			Taobao (closed)
C. Li et al. [20]	2021	Mul-AN	Criteo
			MovieLens-100k
X. Li et al. [8]	2020	MARN	Amazon Review Electro
			Amazon Review Clothing

Author	Year	Approach	Dataset
X. Lie et al. [22]	2020	TIEN	Taobao (closed)
			Amazon Review Beauty
			Amazon Review Clothing
			Amazon Review Grocery
			Amazon Review Phones
			Amazon Review Sports
Zeng et al. [27]	2020	USRF	RetailRocket datasets
			Amazon Review Digital Music
			MovieLense-1M
Zhou et al. [10]	2019	DIEN	Amazon Review Electro
			Amazon Review Books
			Taobao
Zhou et al. [21]	2018	DIN	Amazon Review Electro
			MovieLense-20M
			Alibaba (closed)
Wang et al. [28]	2017	DCN	Criteo

3. Customer Representation

Traditionally, customer behavior is modeled by domain experts to make predictions of their intentions and future behavior. Therefore, data like clickstream data or demographic information are incorporated into the data analysis and feature engineering process [29][30][31][32][33]. As shown by Alves Gomes et al. [34] most customer representations are modeled with manual features extracted by experts or with the RFM analysis [35]. For example, Perisic et al. [36] and Friedrich et al. [37] extracted RFM-based features by extending the RFM analysis from historical data for customer representation. Wu et al. [38] modeled and analyzed customer behavior with an extended RFM approach by adding customer contribution time and repeat purchase attributes and combining it with a k-means clustering. K-means clustering is also used by Hamed Fazlollahtabar [39]. The author chose different customer information gathered from their transactions and applied k-means clustering of different combinations of two features, e.g., gender and product or age and product. Wang et al. [40] analyzed influence factors of second-hand customer-to-customer e-commerce platforms by using questioner and demographic information of customers. Esmeli et al. [33] modeled customers based on twelve features solely based on session information. Berger et al. [41] used features that describe the change in customer behavior based on actual session

information and the information retrieved from previous session history. This manual customer representation process is time-consuming and expensive, especially since it needs to be repeated for each new use case or marketing campaign.

Recent approaches that use embedding layers simplify customer modeling by only inserting information into the learning model without a proper feature engineering process. Most of the aforementioned CTR-P approaches utilize embedding layers to learn customer behavior. Sheil et al. [42] proposed an end-to-end three-layered LSTM to predict future customer behavior by learning patterns of the product the customer interacts with, the interaction time, and additional product-related information. Ni et al. [11] proposed a Deep User Perception Network (DUPN) an end-to-end Long-Short Term Memory (LSTM) with an embedding input that is trained on multiple tasks for a general customer representation. Yang et al. [43] and Wu et al. [44] represented customers based on textual features like product names, categories, and reviews written by the customers. However, in addition to using an embedding layer for input data, embeddings can also be used to represent features. Especially in the e-commerce context, embeddings were used in recommendation scenarios. For this purpose, product embeddings were created and trained [45][46][47][48]. A recent approach using pre-trained embedding features to represent customer behavior was proposed by Alves Gomes et al. [49][50]. The authors pre-trained an embedding to encode customers' behavior and used the representation to predict customers' purchase intention.

References

1. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention is All you Need. In Advances in Neural Information Processing Systems; Guyon, I., Luxburg, U.V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2017; Volume 30.

2. Brown, T.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J.D.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A.; et al. Language Models are Few-Shot Learners. In Advances in Neural Information Processing Systems; Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., Lin, H., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2020; Volume 33, pp. 1877–1901.

3. Ramesh, A.; Pavlov, M.; Goh, G.; Gray, S.; Voss, C.; Radford, A.; Chen, M.; Sutskever, I. Zero-Shot Text-to-Image Generation. In Proceedings of the 38th International Conference on Machine Learning, PMLR, Virtual, 18–24 July 2021; Volume 139, pp. 8821–8831.

4. Cheng, H.T.; Koc, L.; Harmsen, J.; Shaked, T.; Chandra, T.; Aradhye, H.; Anderson, G.; Corrado, G.; Chai, W.; Ispir, M.; et al. Wide & Deep Learning for Recommender Systems. In Proceedings of the DLRS 2016 1st Workshop on Deep Learning for Recommender Systems, Boston, MA, USA, 15 September 2016; pp. 7–10.

5. Sun, F.; Liu, J.; Wu, J.; Pei, C.; Lin, X.; Ou, W.; Jiang, P. BERT4Rec: Sequential Recommendation with Bidirectional Encoder Representations from Transformer. In Proceedings of the CIKM '19

28th ACM International Conference on Information and Knowledge Management, Beijing, China, 3–7 November 2019; pp. 1441–1450.

6. Huang, G.; Chen, Q.; Deng, C. A New Click-Through Rates Prediction Model Based on Deep&Cross Network. Algorithms 2020, 13, 342.

7. Xia, Z.; Mao, S.; Bai, J.; Geng, X.; Yi, L. A Novel Integrated Network with LightGBM for Click-Through Rate Prediction. Res. Sq. 2021; preprint.

8. Li, X.; Wang, C.; Tan, J.; Zeng, X.; Ou, D.; Zheng, B. Adversarial Multimodal Representation Learning for Click-Through Rate Prediction. In Proceedings of the Web Conference 2020, Taipei, Taiwan, 20–24 April 2020; pp. 827–836.

9. Fan, Z.; Ou, D.; Gu, Y.; Fu, B.; Li, X.; Bao, W.; Dai, X.Y.; Zeng, X.; Zhuang, T.; Liu, Q. Modeling Users' Contextualized Page-wise Feedback for Click-Through Rate Prediction in E-commerce Search. In Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining, Virtual, 21–25 February 2022; pp. 262–270.

10. Zhou, G.; Mou, N.; Fan, Y.; Pi, Q.; Bian, W.; Zhou, C.; Zhu, X.; Gai, K. Deep Interest Evolution Network for Click-Through Rate Prediction. Proc. AAAI Conf. Artif. Intell. 2019, 33, 5941–5948.

11. Ni, Y.; Ou, D.; Liu, S.; Li, X.; Ou, W.; Zeng, A.; Si, L. Perceive Your Users in Depth: Learning Universal User Representations from Multiple E-Commerce Tasks. In Proceedings of the KDD '18 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, London, UK, 19–23 August 2018; pp. 596–605.

12. Carmel, D.; Haramaty, E.; Lazerson, A.; Lewin-Eytan, L. Multi-Objective Ranking Optimization for Product Search Using Stochastic Label Aggregation. In Proceedings of the WWW '20 Web Conference 2020, Taipei, Taiwan, 20–24 April 2020; pp. 373–383.

13. Li, F.; Chen, Z.; Wang, P.; Ren, Y.; Zhang, D.; Zhu, X. Graph Intention Network for Click-through Rate Prediction in Sponsored Search. In Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, Paris, France, 21–25 July 2019; pp. 961–964.

14. Pan, Z.; Chen, E.; Liu, Q.; Xu, T.; Ma, H.; Lin, H. Sparse Factorization Machines for Click-Through Rate Prediction. In Proceedings of the 2016 IEEE 16th International Conference on Data Mining (ICDM), Barcelona, Spain, 12–15 December 2016; pp. 400–409.

15. Ren, K.; Zhang, W.; Rong, Y.; Zhang, H.; Yu, Y.; Wang, J. User Response Learning for Directly Optimizing Campaign Performance in Display Advertising. In Proceedings of the 25th ACM International on Conference on Information and Knowledge Management, Indianapolis, IN, USA, 24–28 October 2016; pp. 679–688.

16. Kumar, V.; Venkatesan, R.; Reinartz, W. Performance Implications of Adopting a Customer-Focused Sales Campaign. J. Mark. 2008, 72, 50–68.

17. Chen, C.; Chen, H.; Zhao, K.; Zhou, J.; He, L.; Deng, H.; Xu, J.; Zheng, B.; Zhang, Y.; Xing, C. EXTR: Click-Through Rate Prediction with Externalities in E-Commerce Sponsored Search. In Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, 14–18 August 2022; pp. 2732–2740.

18. Ge, T.; Zhao, L.; Zhou, G.; Chen, K.; Liu, S.; Yi, H.; Hu, Z.; Liu, B.; Sun, P.; Liu, H.; et al. Image Matters: Visually Modeling User Behaviors Using Advanced Model Server. In Proceedings of the 27th ACM International Conference on Information and Knowledge Management, Torino, Italy, 22–26 October 2018; pp. 2087–2095.

19. Gulhane, P.R.; Kumar, T.S.P. TensorFlow Based Website Click through Rate (CTR) Prediction Using Heat maps. In Proceedings of the 2018 International Conference on Recent Trends in Advance Computing (ICRTAC), Chennai, India, 10–11 September 2018; pp. 97–102.

20. Li, C.; Yi, K.; Fei, M.; Zhou, W.; Wu, X.; Chen, Y. Multiple-structure Attentional Network for Click-through Prediction in Recommendation System. In Proceedings of the 2021 IEEE International Conference on Recent Advances in Systems Science and Engineering (RASSE), Shanghai, China, 12–14 December 2021; pp. 1–6.

21. Zhou, G.; Zhu, X.; Song, C.; Fan, Y.; Zhu, H.; Ma, X.; Yan, Y.; Jin, J.; Li, H.; Gai, K. Deep Interest Network for Click-Through Rate Prediction. In Proceedings of the KDD '18 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, London, UK, 19–23 August 2018; pp. 1059–1068.

22. Li, X.; Wang, C.; Tong, B.; Tan, J.; Zeng, X.; Zhuang, T. Deep Time-Aware Item Evolution Network for Click-Through Rate Prediction. In Proceedings of the 29th ACM International Conference on Information & Knowledge Management, Virtual, 19–23 October 2020; pp. 785–794.

23. Wang, F.; Zhao, L. A Hybrid Model for Commercial Brand Marketing Prediction Based on Multiple Features with Image Processing. Secur. Commun. Netw. 2022, 2022, 5455745.

24. Wong, C.M.; Feng, F.; Zhang, W.; Vong, C.M.; Chen, H.; Zhang, Y.; He, P.; Chen, H.; Zhao, K.; Chen, H. Improving Conversational Recommender System by Pretraining Billion-scale Knowledge Graph. In Proceedings of the 2021 IEEE 37th International Conference on Data Engineering (ICDE), Chania, Greece, 19–22 April 2021; pp. 2607–2612.

25. Yao, S.; Tan, J.; Chen, X.; Yang, K.; Xiao, R.; Deng, H.; Wan, X. Learning a Product Relevance Model from Click-Through Data in E-Commerce. In Proceedings of the Web Conference 2021, Online, 19–23 April 2021; pp. 2890–2899.

26. Miguel Alves Gomes; Richard Meyes; Philipp Meisen; Tobias Meisen; It's Not Always about Wide and Deep Models: Click-Through Rate Prediction with a Customer Behavior-Embedding Representation. *J. Theor. Appl. Electron. Commer. Res.*. **2024**, *19*, 135-151.

27. Zeng, J.; Chen, Y.; Zhu, H.; Tian, F.; Miao, K.; Liu, Y.; Zheng, Q. User Sequential Behavior Classification for Click-Through Rate Prediction. In Proceedings of the Database Systems for Advanced Applications. DASFAA 2020 International Workshops: BDMS, SeCoP, BDQM, GDMA, and AIDE, Jeju, Republic of Korea, 24–27 September 2020; Springer: Berlin/Heidelberg, Germany, 2020; pp. 267–280.

28. Wang, R.; Fu, B.; Fu, G.; Wang, M. Deep & Cross Network for Ad Click Predictions. In Proceedings of the ADKDD'17, Halifax, NS, Canada, 14 August 2017.

29. Sismeiro, C.; Bucklin, R.E. Modeling purchase behavior at an e-commerce web site: A task-completion approach. J. Mark. Res. 2004, 41, 306–323.

30. Romov, P.; Sokolov, E. RecSys Challenge 2015: Ensemble Learning with Categorical Features. In Proceedings of the RecSys '15 Challenge: 2015 International ACM Recommender Systems Challenge, Vienna, Austria, 16–20 September 2015.

31. Li, Q.; Gu, M.; Zhou, K.; Sun, X. Multi-Classes Feature Engineering with Sliding Window for Purchase Prediction in Mobile Commerce. In Proceedings of the 2015 IEEE International Conference on Data Mining Workshop (ICDMW), Atlantic City, NJ, USA, 14–17 November 2015; pp. 1048–1054.

32. Martínez, A.; Schmuck, C.; Pereverzyev, S.; Pirker, C.; Haltmeier, M. A machine learning framework for customer purchase prediction in the non-contractual setting. Eur. J. Oper. Res. 2020, 281, 588–596.

33. Esmeli, R.; Bader-El-Den, M.; Abdullahi, H. Towards early purchase intention prediction in online session based retailing systems. Electron. Mark. 2021, 31, 697–715.

34. Alves Gomes, M.; Meisen, T. A review on customer segmentation methods for personalized customer targeting in e-commerce use cases. Inf. Syst. e-Bus. Manag. 2023, 21, 527–570.

35. Hughes, A.M. Strategic Database Marketing: The Masterplan for Starting and Managing a Profitable, Customer-Based Marketing Program; Irwin Professional: Burr Ridge, IL, USA, 1994.

36. Perišić, A.; Pahor, M. RFM-LIR Feature Framework for Churn Prediction in the Mobile Games Market. IEEE Trans. Games 2022, 14, 126–137.

37. Fridrich, M.; Dostál, P. User Churn Model in E-Commerce Retail. Sci. Pap. Univ. Pardubic. Ser. D Fac. Econ. Adm. 2022, 30.

38. Wu, J.; Shi, L.; Yang, L.; Niu, X.; Li, Y.; Cui, X.; Tsai, S.B.; Zhang, Y. User value identification based on improved RFM model and k-means++ algorithm for complex data analysis. Wirel. Commun. Mob. Comput. 2021, 2021, 9982484.

39. Fazlollahtabar, H. Intelligent marketing decision model based on customer behavior using integrated possibility theory and K-means algorithm. J. Intell. Manag. Decis. 2022, 1, 88–96.

40. Wang, L.; Sun, H. Influencing Factors of Second-Hand Platform Trading in C2C E-commerce. J. Intell. Manag. Decis. 2023, 2, 21–29.

41. Berger, P.; Kompan, M. User Modeling for Churn Prediction in E-Commerce. IEEE Intell. Syst. 2019, 34, 44–52.

42. Sheil, H.; Rana, O.; Reilly, R. Predicting purchasing intent: Automatic feature learning using recurrent neural networks. arXiv 2018, arXiv:1807.08207.

43. Yang, B.; Liu, K.; Xu, X.; Xu, R.; Liu, H.; Xu, H. Learning Universal User Representations via Self-Supervised Lifelong Behaviors Modeling. In Proceedings of the ICLR 2022 Conference, Virtual, 25–29 April 2022.

44. Wu, C.; Wu, F.; Qi, T.; Lian, J.; Huang, Y.; Xie, X. Ptum: Pre-training user model from unlabeled user behaviors via self-supervision. arXiv 2020, arXiv:2010.01494.

45. Vasile, F.; Smirnova, E.; Conneau, A. Meta-Prod2Vec: Product Embeddings Using Side-Information for Recommendation. In Proceedings of the RecSys '16: 10th ACM Conference on Recommender Systems, Boston, MA, USA, 15–19 September 2016; pp. 225–232.

46. Tercan, H.; Bitter, C.; Bodnar, T.; Meisen, P.; Meisen, T. Evaluating a Session-based Recommender System using Prod2vec in a Commercial Application. In Proceedings of the 23rd International Conference on Enterprise Information Systems, Virtual, 26–28 April 2021; SciTePress: Setúbal, Portugal, 2021; Volume 1, pp. 610–617.

47. Alves Gomes, M.; Tercan, H.; Bodnar, T.; Meisen, P.; Meisen, T. A Filter is Better Than None: Improving Deep Learning-Based Product Recommendation Models by Using a User Preference Filter. In Proceedings of the 2021 IEEE 23rd Int Conf on High Performance Computing & Communications; 7th Int Conf on Data Science & Systems; 19th Int Conf on Smart City; 7th Int Conf on Dependability in Sensor, Cloud & Big Data Systems & Application (HPCC/DSS/SmartCity/DependSys), Haikou, China, 20–22 December 2021; pp. 1278–1285.

48. Srilakshmi, M.; Chowdhury, G.; Sarkar, S. Two-stage system using item features for next-item recommendation. Intell. Syst. Appl. 2022, 14, 200070.

49. Alves Gomes, M.; Meyes, R.; Meisen, P.; Meisen, T. Will This Online Shopping Session Succeed? Predicting Customer's Purchase Intention Using Embeddings. In Proceedings of the CIKM '22: 31st ACM International Conference on Information & Knowledge Management, Atlanta, GA, USA, 17–21 October 2022; pp. 2873–2882.

50. Alves Gomes, M.; Wönkhaus, M.; Meisen, P.; Meisen, T. TEE: Real-Time Purchase Prediction Using Time Extended Embeddings for Representing Customer Behavior. J. Theor. Appl. Electron. Commer. Res. 2023, 18, 1404–1418.