**xDeepFM: Combining Explicit and Implicit Feature Interactions for Recommender Systems**

**TÓM TẮT**

Trong lĩnh vực hệ thống gợi ý (hệ khuyến nghị), việc kết hợp tương tác giữa các đặc trưng mà ta có thể nhìn thấy (rõ ràng) và những tương tác ẩn (không rõ ràng) đang là một vấn đề thách thức. Trong bài báo này, tác giả giới thiệu mô hình xDeepFM, một phương pháp mới kết hợp cả hai loại tương tác này, nhằm cải thiện hiệu suất của hệ thống gợi ý. Mô hình xDeepFM kết hợp hai kiến trúc chính là mạng nơ-ron sâu (DNNs) và mạng nơ-ron tương tác thông tin (FM). DNN được sử dụng để học các tương tác đặc trưng ngầm, trong khi FM được sử dụng để mô hình hóa tương tác đặc trưng rõ ràng. Với xDeepFM, tác giả sử dụng một tầng tương tác thông tin (Cross Network) để kết hợp cả hai kiến trúc trên. Tầng tương tác thông tin này cho phép mô hình xDeepFM học được các tương tác đặc trưng phức tạp và tạo ra các đặc trưng mới từ các tương tác này. Mô hình xDeepFM được đánh giá và so sánh với các mô hình gợi ý khác trên các tập dữ liệu gợi ý phổ biến như MovieLens và Criteo. Kết quả thí nghiệm cho thấy xDeepFM vượt trội với hiệu suất tốt hơn so với các mô hình FM, Wide & Deep và DeepFM truyền thống. Tổng quan này chỉ ra rằng mô hình xDeepFM là một phương pháp tiềm năng trong việc kết hợp cả tương tác đặc trưng rõ ràng và những tương tác ẩn để cải thiện hiệu suất của hệ thống gợi ý trong các ứng dụng thực tế.

Từ khoá: xDeepFM, tương tác đặc trưng, hệ khuyến nghị, mạng nơ-ron sâu (DNN), tầng tương tác thông tin (Cross Network)

**1 GIỚI THIỆU**

Trong thời đại số hóa hiện nay, hệ thống gợi ý đã trở thành một công cụ quan trọng để giúp người dùng khám phá và tiếp cận thông tin, sản phẩm và nội dung phù hợp với sở thích cá nhân. Các hệ thống gợi ý dựa trên các thuật toán phức tạp để phân tích dữ liệu người dùng và đề xuất các mục tiêu phù hợp nhằm cung cấp trải nghiệm cá nhân hóa và tăng cường sự tương tác người dùng.

Một trong những thách thức quan trọng của hệ thống gợi ý là kết hợp cả tương tác đặc trưng rõ ràng và tương tác ẩn từ dữ liệu để tạo ra các đề xuất chất lượng và mang tính cá nhân cao. Tương tác đặc trưng rõ ràng là các tương tác dựa trên thông tin đã biết và được khai thác một cách rõ ràng, trong khi tương tác đặc trưng ngầm là các tương tác ẩn chứa trong dữ liệu mà chúng ta cần tìm hiểu và mô hình hóa.

Trong bài báo này, chúng ta nghiên cứu một mô hình tiên tiến có tên là "xDeepFM: Combining Explicit and Implicit Feature Interactions for Recommender Systems". Mô hình này giải quyết vấn đề kết hợp cả tương tác đặc trưng rõ ràng và đặc trưng ngầm bằng cách kết hợp hai kiến trúc quan trọng: mạng nơ-ron sâu (DNN) và mạng nơ-ron tương tác thông tin (FM).

mạng nơ-ron sâu (DNN) được sử dụng để học các tương tác đặc trưng ngầm, cho phép mô hình xDeepFM tìm ra các mẫu phức tạp và đặc trưng ẩn từ dữ liệu gợi ý. Mạng nơ-ron tương tác thông tin (FM) được sử dụng để mô hình hóa tương tác đặc trưng rõ ràng bằng cách tính toán tương tự ma trận đặc trưng.

Một đặc điểm nổi bật của mô hình xDeepFM là tầng tương tác thông tin (Cross Network), giúp kết hợp thông minh giữa DNN và FM. Tầng này cho phép mô hình học các tương tác đặc trưng phức tạp và tạo ra các đặc trưng mới bằng cách kết hợp cả tương tác đặc trưng rõ ràng và đặc trưng ngầm.

Để đánh giá hiệu suất của mô hình xDeepFM, chúng ta thực hiện các thí nghiệm trên các tập dữ liệu gợi ý phổ biến như MovieLens và Criteo. Kết quả thí nghiệm cho thấy xDeepFM vượt trội so với các mô hình gợi ý truyền thống như FM, Wide & Deep và DeepFM, chứng minh tiềm năng của việc kết hợp cả tương tác đặc trưng rõ ràng và đặc trưng ngầm trong việc cải thiện hiệu suất hệ thống gợi ý trong các ứng dụng thực tế.

Phần còn lại của báo cáo sẽ trình bày chi tiết về kiến trúc và phương pháp của mô hình xDeepFM, kết quả thí nghiệm và thảo luận về những điểm đáng chú ý và tiềm năng cho công trình nghiên cứu này.

**5 CÔNG TRÌNH LIÊN QUAN**

**5.1 Hệ thống gợi ý cổ điển**

Hệ thống gợi ý đã được nghiên cứu rộng rãi trong lĩnh vực truy xuất thông tin và học máy. Các hệ thống gợi ý cổ điển có thể được phân loại dựa trên hai mô hình sau: Mô hình không phân rã (Non-factorization Models) và Mô hình phân rã ma trận (Factorization Models).

Mô hình không phân rã (Non-factorization Models) ở đây là các mô hình tuyến tính, điển hình như hồi quy logistic. Ưu điểm của mô hình này là áp dụng được với các thuộc tính hỗn hợp, nhiều chiều và dễ quản lý, triển khai trên các nền tảng. Nhược điểm ở đây là mô hình không học được các đặc trưng của thuộc tính cũng như tốn nhiều thời gian tìm các thuộc tính kết hợp để nâng cao hiệu suất.

Mô hình phân rã ma trận (Factorization Models) bao gồm các pương pháp dựa trên lọc cộng tác (collaborative filtering) và phương pháp dựa trên nội dung (content-based).

Lọc cộng tác là một phương pháp quan trọng trong hệ thống gợi ý, nó được chia thành hai dạng chính là lọc cộng tác dựa trên người dùng (user-based) và lọc cộng tác dựa trên mục tiêu (item-based). Phương pháp lọc cộng tác dựa trên người dùng dựa trên việc tìm kiếm sự tương đồng giữa người dùng dựa trên lịch sử hoạt động của họ, trong khi phương pháp lọc cộng tác dựa trên mục tiêu tìm kiếm sự tương đồng giữa các mục tiêu dựa trên cách người dùng đã tương tác với chúng. Cả hai phương pháp đều sử dụng thông tin từ các người dùng khác hoặc các mục tiêu khác để đưa ra các đề xuất gợi ý.

Phương pháp dựa trên nội dung tập trung vào việc phân tích nội dung của các mục tiêu hoặc người dùng để tìm ra các đặc trưng quan trọng. Các phương pháp này xem xét các đặc trưng của mục tiêu hoặc người dùng, chẳng hạn như từ khoá, thể loại, hoặc thuộc tính khác, và sử dụng chúng để xác định sự tương đồng và đề xuất các mục tiêu tương tự hoặc người dùng tương tự.

Cả hai phương pháp truyền thống này đã đạt được kết quả khá ấn tượng trong các ứng dụng gợi ý. Tuy nhiên, chúng vẫn gặp một số hạn chế, bao gồm sự thiếu, mất mát thông tin và khả năng mô hình hóa tương tác đặc trưng phức tạp giữa các đối tượng gợi ý.

Trong bài báo này, chúng ta tiếp tục nghiên cứu và đề xuất một mô hình tiên tiến, xDeepFM, để vượt qua những hạn chế này và nâng cao hiệu suất của hệ thống gợi ý.

**5.2 Hệ thống gợi ý sử dụng học sâu**

Học sâu (Deep Learning) là một lĩnh vực quan trọng trong lĩnh vực Trí tuệ nhân tạo và học máy. Hệ thống gợi ý sử dụng học sâu đã thu hút sự quan tâm của nhiều nhà nghiên cứu và mang lại những kết quả đáng chú ý trong việc cải thiện hiệu suất của các hệ thống gợi ý.

Một trong những mô hình phổ biến trong hệ thống gợi ý sử dụng học sâu là mạng nơ-ron sâu (Deep Neural Network - DNN). Mô hình DNN sử dụng một kiến trúc nhiều tầng để học các đặc trưng phức tạp và tạo ra các dự đoán gợi ý dựa trên đó. Mô hình này đã được áp dụng thành công trong nhiều lĩnh vực, bao gồm gợi ý sản phẩm, âm nhạc, phim ảnh và nội dung trực tuyến.

Ngoài ra, một mô hình khác phổ biến trong hệ thống gợi ý sử dụng học sâu là Mạng nơ-ron tích chập (Convolutional Neural Network - CNN). Mô hình CNN chủ yếu được sử dụng trong việc phân loại hình ảnh, nhưng cũng có thể được áp dụng để xác định sự tương đồng và đề xuất các mục tiêu gợi ý dựa trên nội dung.

Các nghiên cứu trước đây đã chứng minh rằng việc sử dụng học sâu trong hệ khuyến nghị có thể cải thiện hiệu suất so với các phương pháp truyền thống. Tuy nhiên, nhược điểm của một số mô hình học sâu là thiếu khả năng mô hình hóa tương tác đặc trưng phức tạp giữa các đối tượng gợi ý.

Trong bài báo này, chúng ta tiếp tục nghiên cứu và đề xuất mô hình xDeepFM, một sự kết hợp thông minh giữa DNN và FM, nhằm cải thiện khả năng mô hình hóa tương tác đặc trưng rõ ràng và ngầm.

**6 KẾT LUẬN**

Trong bài báo này, chúng tôi đã trình bày mô hình xDeepFM, một phương pháp kết hợp tương tác đặc trưng rõ ràng và ngầm trong hệ thống gợi ý. XDeepFM sử dụng một kiến trúc mạng học sâu kết hợp với mô hình tương tác đặc trưng (Factorization Machine - FM) để khai thác các mối quan hệ phức tạp giữa các đặc trưng trong quá trình gợi ý.

Kết quả thí nghiệm trên các tập dữ liệu gợi ý phổ biến như MovieLens và Criteo đã chứng minh hiệu suất vượt trội của xDeepFM so với các mô hình gợi ý truyền thống như FM, Wide & Deep và DeepFM. Kết quả này đã chứng minh tiềm năng của việc kết hợp cả tương tác đặc trưng rõ ràng và ngầm trong việc cải thiện hiệu suất hệ thống gợi ý trong các ứng dụng thực tế.

Bên cạnh đó, chúng tôi đã đề xuất một số cải tiến cho xDeepFM, như tối ưu hóa hàm mất mát và tăng cường khả năng mô hình hóa tương tác đặc trưng. Các cải tiến này có thể đóng góp vào việc nâng cao hiệu suất và khả năng áp dụng của mô hình trong tương lai.

Tổng kết lại, mô hình xDeepFM đại diện cho một tiến bộ quan trọng trong lĩnh vực hệ thống gợi ý. Bằng cách kết hợp cả tương tác đặc trưng rõ ràng và ngầm, xDeepFM có khả năng mô hình hóa những mối quan hệ phức tạp và tạo ra những gợi ý chính xác hơn. Công trình nghiên cứu này cung cấp một cơ sở vững chắc cho sự phát triển và cải tiến các hệ thống gợi ý trong tương lai.

**TÀI LIỆU THAM KHẢO**

[1] Dario Amodei, Sundaram Ananthanarayanan, Rishita Anubhai, Jingliang Bai, Eric Battenberg, Carl Case, Jared Casper, Bryan Catanzaro, Qiang Cheng, Guoliang Chen, et al. 2016. Deep speech 2: End-to-end speech recognition in english and mandarin. In International Conference on Machine Learning. 173–182.

[2] Mathieu Blondel, Akinori Fujino, Naonori Ueda, and Masakazu Ishihata. 2016. Higher-order factorization machines. In Advances in Neural Information Processing Systems. 3351–3359.

[3] Jingyuan Chen, Hanwang Zhang, Xiangnan He, Liqiang Nie, Wei Liu, and TatSeng Chua. 2017. Attentive collaborative filtering: Multimedia recommendation with item-and component-level attention. In Proceedings of the 40th International ACM SIGIR conference on Research and Development in Information Retrieval. ACM, 335–344.

[4] Tianqi Chen, Weinan Zhang, Qiuxia Lu, Kailong Chen, Zhao Zheng, and Yong Yu. 2012. SVDFeature: a toolkit for feature-based collaborative filtering. Journal of Machine Learning Research 13, Dec (2012), 3619–3622.

[5] Heng-Tze Cheng, Levent Koc, Jeremiah Harmsen, Tal Shaked, Tushar Chandra, Hrishi Aradhye, Glen Anderson, Greg Corrado, Wei Chai, Mustafa Ispir, et al. 2016. Wide & deep learning for recommender systems. In Proceedings of the 1st Workshop on Deep Learning for Recommender Systems. ACM, 7–10.

[6] Kyunghyun Cho, Bart Van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. 2014. Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078 (2014).

[7] Xin Dong, Lei Yu, Zhonghuo Wu, Yuxia Sun, Lingfeng Yuan, and Fangxi Zhang. 2017. A Hybrid Collaborative Filtering Model with Deep Structure for Recommender Systems. In AAAI. 1309–1315.

[8] Ali Mamdouh Elkahky, Yang Song, and Xiaodong He. 2015. A multi-view deep learning approach for cross domain user modeling in recommendation systems. In Proceedings of the 24th International Conference on World Wide Web. International World Wide Web Conferences Steering Committee, 278–288.

[9] Huifeng Guo, Ruiming Tang, Yunming Ye, Zhenguo Li, and Xiuqiang He. 2017. Deepfm: A factorization-machine based neural network for CTR prediction. arXiv preprint arXiv:1703.04247 (2017).

[10] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition. 770–778.

[11] Ruining He and Julian McAuley. 2016. VBPR: Visual Bayesian Personalized Ranking from Implicit Feedback. In AAAI. 144–150.

[12] Xiangnan He and Tat-Seng Chua. 2017. Neural factorization machines for sparse predictive analytics. In Proceedings of the 40th International ACM SIGIR conference on Research and Development in Information Retrieval. ACM, 355–364.

[13] Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. 2017. Neural collaborative filtering. In Proceedings of the 26th International Conference on World Wide Web. International World Wide Web Conferences Steering Committee, 173–182.

[14] Xinran He, Junfeng Pan, Ou Jin, Tianbing Xu, Bo Liu, Tao Xu, Yanxin Shi, Antoine Atallah, Ralf Herbrich, Stuart Bowers, et al. 2014. Practical lessons from predicting clicks on ads at facebook. In Proceedings of the Eighth International Workshop on Data Mining for Online Advertising. ACM, 1–9.

[15] Geoffrey Hinton, Li Deng, Dong Yu, George E Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, Tara N Sainath, et al. 2012. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. IEEE Signal Processing Magazine 29, 6 (2012), 82–97.

[16] Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014).

[17] Yehuda Koren. 2008. Factorization meets the neighborhood: a multifaceted collaborative filtering model. In Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 426–434.

[18] Yehuda Koren, Robert Bell, and Chris Volinsky. 2009. Matrix factorization techniques for recommender systems. Computer 42, 8 (2009).

[19] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. 2012. Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems. 1097–1105.

[20] Joonseok Lee, Seungyeon Kim, Guy Lebanon, and Yoram Singer. 2013. Local low-rank matrix approximation. In International Conference on Machine Learning. 82–90.

[21] Jianxun Lian and Xing Xie. 2016. Cross-Device User Matching Based on Massive Browse Logs: The Runner-Up Solution for the 2016 CIKM Cup. arXiv preprint arXiv:1610.03928 (2016).

[22] Jianxun Lian, Fuzheng Zhang, Min Hou, Hongwei Wang, Xing Xie, and Guangzhong Sun. 2017. Practical Lessons for Job Recommendations in the Cold-Start Scenario. In Proceedings of the Recommender Systems Challenge 2017 (RecSys Challenge ’17). ACM, New York, NY, USA, Article 4, 6 pages. https://doi.org/10.1145/3124791.3124794

[23] Jianxun Lian, Fuzheng Zhang, Xing Xie, and Guangzhong Sun. 2017. CCCFNet: a content-boosted collaborative filtering neural network for cross domain recommender systems. In Proceedings of the 26th International Conference on World Wide Web Companion. International World Wide Web Conferences Steering Committee, 817–818.

[24] Jianxun Lian, Fuzheng Zhang, Xing Xie, and Guangzhong Sun. 2017. Restaurant Survival Analysis with Heterogeneous Information. In Proceedings of the 26th International Conference on World Wide Web Companion. International World Wide Web Conferences Steering Committee, 993–1002.

[25] Xiaoliang Ling, Weiwei Deng, Chen Gu, Hucheng Zhou, Cui Li, and Feng Sun. 2017. Model Ensemble for Click Prediction in Bing Search Ads. In Proceedings of the 26th International Conference on World Wide Web Companion. International World Wide Web Conferences Steering Committee, 689–698.

[26] Guimei Liu, Tam T Nguyen, Gang Zhao, Wei Zha, Jianbo Yang, Jianneng Cao, Min Wu, Peilin Zhao, and Wei Chen. 2016. Repeat buyer prediction for e-commerce. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 155–164.

[27] H Brendan McMahan, Gary Holt, David Sculley, Michael Young, Dietmar Ebner, Julian Grady, Lan Nie, Todd Phillips, Eugene Davydov, Daniel Golovin, et al. 2013. Ad click prediction: a view from the trenches. In Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 1222–1230.

[28] Aditya Krishna Menon and Charles Elkan. 2010. A log-linear model with latent features for dyadic prediction. In Data Mining (ICDM), 2010 IEEE 10th International Conference on. IEEE, 364–373. KDD ’18, August 19–23, 2018, London, United Kingdom J. Lian, X. Zhou, F. Zhang, Z. Chen, X. Xie, and G. Sun

[29] Tomáš Mikolov, Martin Karafiát, Lukáš Burget, Jan Černocky, and Sanjeev Khu- ` danpur. 2010. Recurrent neural network based language model. In Eleventh Annual Conference of the International Speech Communication Association.

[30] Rong Pan, Yunhong Zhou, Bin Cao, Nathan N Liu, Rajan Lukose, Martin Scholz, and Qiang Yang. 2008. One-class collaborative filtering. In Data Mining, 2008. ICDM’08. Eighth IEEE International Conference on. IEEE, 502–511.

[31] Yanru Qu, Han Cai, Kan Ren, Weinan Zhang, Yong Yu, Ying Wen, and Jun Wang. 2016. Product-based neural networks for user response prediction. In Data Mining (ICDM), 2016 IEEE 16th International Conference on. IEEE, 1149–1154.

[32] Steffen Rendle. 2010. Factorization machines. In Data Mining (ICDM), 2010 IEEE 10th International Conference on. IEEE, 995–1000.

[33] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2009. BPR: Bayesian personalized ranking from implicit feedback. In Proceedings of the twenty-fifth conference on uncertainty in artificial intelligence. AUAI Press, 452–461.

[34] Steffen Rendle and Lars Schmidt-Thieme. 2010. Pairwise interaction tensor factorization for personalized tag recommendation. In Proceedings of the third ACM international conference on Web search and data mining. ACM, 81–90.

[35] Matthew Richardson, Ewa Dominowska, and Robert Ragno. 2007. Predicting clicks: estimating the click-through rate for new ads. In Proceedings of the 16th international conference on World Wide Web. ACM, 521–530.

[36] Suvash Sedhain, Aditya Krishna Menon, Scott Sanner, and Lexing Xie. 2015. Autorec: Autoencoders meet collaborative filtering. In Proceedings of the 24th International Conference on World Wide Web. ACM, 111–112.

[37] Ying Shan, T Ryan Hoens, Jian Jiao, Haijing Wang, Dong Yu, and JC Mao. 2016. Deep crossing: Web-scale modeling without manually crafted combinatorial features. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 255–262.

[38] Nathan Srebro, Jason Rennie, and Tommi S Jaakkola. 2005. Maximum-margin matrix factorization. In Advances in neural information processing systems. 1329– 1336.

[39] Hao Wang, Naiyan Wang, and Dit-Yan Yeung. 2015. Collaborative deep learning for recommender systems. In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 1235–1244.

[40] Ruoxi Wang, Bin Fu, Gang Fu, and Mingliang Wang. 2017. Deep & Cross Network for Ad Click Predictions. arXiv preprint arXiv:1708.05123 (2017).

[41] Xinxi Wang and Ye Wang. 2014. Improving content-based and hybrid music recommendation using deep learning. In Proceedings of the 22nd ACM international conference on Multimedia. ACM, 627–636.

[42] Yao Wu, Christopher DuBois, Alice X Zheng, and Martin Ester. 2016. Collaborative denoising auto-encoders for top-n recommender systems. In Proceedings of the Ninth ACM International Conference on Web Search and Data Mining. ACM, 153–162.

[43] Jun Xiao, Hao Ye, Xiangnan He, Hanwang Zhang, Fei Wu, and Tat-Seng Chua. 2017. Attentional Factorization Machines: Learning the Weight of Feature Interactions via Attention Networks. In Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI 2017, Melbourne, Australia, August 19-25, 2017. 3119–3125. https://doi.org/10.24963/ijcai.2017/435

[44] Fajie Yuan, Guibing Guo, Joemon M Jose, Long Chen, Haitao Yu, and Weinan Zhang. 2016. Lambdafm: learning optimal ranking with factorization machines using lambda surrogates. In Proceedings of the 25th ACM International on Conference on Information and Knowledge Management. ACM, 227–236.

[45] Fuzheng Zhang, Nicholas Jing Yuan, Defu Lian, Xing Xie, and Wei-Ying Ma. 2016. Collaborative knowledge base embedding for recommender systems. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. ACM, 353–362.

[46] Weinan Zhang, Tianming Du, and Jun Wang. 2016. Deep learning over multi-field categorical data. In European conference on information retrieval. Springer, 45–57.

[47] Guorui Zhou, Chengru Song, Xiaoqiang Zhu, Xiao Ma, Yanghui Yan, Xingya Dai, Han Zhu, Junqi Jin, Han Li, and Kun Gai. 2017. Deep interest network for click-through rate prediction. arXiv preprint arXiv:1706.06978 (2017).