

REFERENCES

- [1] Zeyuan Allen-Zhu and Yuanzhi Li. 2020. Towards Understanding Ensemble, Knowledge Distillation and Self-Distillation in Deep Learning. *arXiv preprint arXiv:2012.09816* (2020).
- [2] Ting Bai, Ji-Rong Wen, Jun Zhang, and Wayne Xin Zhao. 2017. A neural collaborative filtering model with interaction-based neighborhood. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*. 1979–1982.
- [3] Yu Bai, Sally Goldman, and Li Zhang. 2017. Tapas: Two-pass approximate adaptive sampling for softmax. *arXiv preprint arXiv:1707.03073* (2017).
- [4] Immanuel Bayer, Xiangnan He, Bhargav Kanagal, and Steffen Rendle. 2017. A generic coordinate descent framework for learning from implicit feedback. In *WWW*. 1341–1350.
- [5] Yoshua Bengio, Jean-Sébastien Senécal, et al. 2003. Quick Training of Probabilistic Neural Nets by Importance Sampling. In *AISTATS*. 1–9.
- [6] Guy Blanc and Steffen Rendle. 2018. Adaptive sampled softmax with kernel based sampling. In *International Conference on Machine Learning*. PMLR, 590–599.
- [7] Mingyue Cheng, Runlong Yu, Qi Liu, Vincent W Zheng, Hongke Zhao, Hefu Zhang, and Enhong Chen. 2019. Alpha-Beta Sampling for Pairwise Ranking in One-Class Collaborative Filtering. In *2019 IEEE International Conference on Data Mining (ICDM)*. IEEE, 1000–1005.
- [8] Sumit Chopra, Raia Hadsell, and Yann LeCun. 2005. Learning a similarity metric discriminatively, with application to face verification. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, Vol. 1. IEEE, 539–546.
- [9] Paul Covington, Jay Adams, and Emre Sargin. 2016. Deep neural networks for youtube recommendations. In *Proceedings of the 10th ACM conference on recommender systems*. 191–198.
- [10] Tommaso Furlanello, Zachary C Lipton, Michael Tschannen, Laurent Itti, and Anima Anandkumar. 2018. Born again neural networks. *arXiv preprint arXiv:1805.04770* (2018).
- [11] Xin Geng. 2016. Label distribution learning. *IEEE Transactions on Knowledge and Data Engineering* 28, 7 (2016), 1734–1748.
- [12] Marco Gori, Augusto Pucci, V Roma, and I Siena. 2007. Itemrank: A random-walk based scoring algorithm for recommender engines. In *IJCAI*, Vol. 7. 2766–2771.
- [13] Xiangnan He and Tat-Seng Chua. 2017. Neural factorization machines for sparse predictive analytics. In *Proceedings of the 40th International ACM SIGIR*. 355–364.
- [14] Balázs Hidasi and Alexandros Karatzoglou. 2018. Recurrent neural networks with top-k gains for session-based recommendations. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*. 843–852.
- [15] Balázs Hidasi, Alexandros Karatzoglou, Linas Baltrunas, and Domonkos Tikk. 2015. Session-based recommendations with recurrent neural networks. *arXiv preprint arXiv:1511.06939* (2015).
- [16] Geoffrey Hinton, Oriol Vinyals, and Jeff Dean. 2015. Distilling the knowledge in a neural network. *arXiv preprint arXiv:1503.02531* (2015).
- [17] Folasade Olubusola Isinkaye, YO Folajimi, and Bolande Adefowoke Ojokoh. 2015. Recommendation systems: Principles, methods and evaluation. *Egyptian Informatics Journal* 16, 3 (2015), 261–273.
- [18] Sébastien Jean, Kyunghyun Cho, Roland Memisevic, and Yoshua Bengio. 2014. On using very large target vocabulary for neural machine translation. *arXiv preprint arXiv:1412.2007* (2014).
- [19] Jeff Johnson, Matthijs Douze, and Hervé Jégou. 2019. Billion-scale similarity search with GPUs. *IEEE Transactions on Big Data* (2019).
- [20] Rafal Jozefowicz, Oriol Vinyals, and Schuster et al. 2016. Exploring the limits of language modeling. *arXiv preprint arXiv:1602.02410* (2016).
- [21] Wang-Cheng Kang and Julian McAuley. 2018. Self-attentive sequential recommendation. In *2018 IEEE ICDM*. IEEE, 197–206.
- [22] Weiwei et al Kong. 2020. Rankmax: An Adaptive Projection Alternative to the Softmax Function. *Advances in Neural Information Processing Systems* 33 (2020).
- [23] Walid Krichene and Steffen Rendle. 2020. On sampled metrics for item recommendation. In *ACM SIGKDD*. 1748–1757.
- [24] Jing Li, Pengjie Ren, Zhumin Chen, Zhaochun Ren, Tao Lian, and Jun Ma. 2017. Neural attentive session-based recommendation. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*. 1419–1428.
- [25] Zhi Li, Hongke Zhao, Qi Liu, Zhenya Huang, Tao Mei, and Enhong Chen. 2018. Learning from history and present: Next-item recommendation via discriminatively exploiting user behaviors. In *ACM SIGKDD*. 1734–1743.
- [26] Defu Lian, Qi Liu, and Enhong Chen. 2020. Personalized Ranking with Importance Sampling. In *Proceedings of The Web Conference 2020*. 1093–1103.
- [27] Qi Liu, Enhong Chen, Hui Xiong, Chris HQ Ding, and Jian Chen. 2011. Enhancing collaborative filtering by user interest expansion via personalized ranking. *IEEE TSMC, Part B (Cybernetics)* 42, 1 (2011), 218–233.
- [28] Qi Liu, Yong Ge, Zhongmou Li, Enhong Chen, and Hui Xiong. 2011. Personalized travel package recommendation. In *2011 IEEE ICDM*. IEEE, 407–416.
- [29] Qi Liu, Zhenya Huang, Yu Yin, Enhong Chen, Hui Xiong, Yu Su, and Guoping Hu. 2019. Ekt: Exercise-aware knowledge tracing for student performance prediction. *IEEE Transactions on Knowledge and Data Engineering* 33, 1 (2019), 100–115.
- [30] Takeru et al Miyato. 2018. Virtual adversarial training: a regularization method for supervised and semi-supervised learning. *IEEE transactions on pattern analysis and machine intelligence* 41, 8 (2018), 1979–1993.
- [31] Frederic Morin and Yoshua Bengio. 2005. Hierarchical probabilistic neural network language model. In *Aistats*, Vol. 5. Citeseer, 246–252.
- [32] Mohammad Norouzi, David J Fleet, and Russ R Salakhutdinov. 2012. Hamming distance metric learning. In *NIPS*. 1061–1069.
- [33] Rong Pan, Yunhong Zhou, Bin Cao, Nathan N Liu, Rajan Lukose, Martin Scholz, and Qiang Yang. 2008. One-class collaborative filtering. In *2008 Eighth IEEE International Conference on Data Mining*. IEEE, 502–511.
- [34] Weike Pan and Li Chen. 2013. Gbpr: Group preference based bayesian ranking for one-class collaborative filtering. In *IJCAI*.
- [35] Gabriel et al Pereyra. 2017. Regularizing neural networks by penalizing confident output distributions. *arXiv preprint arXiv:1701.06548* (2017).
- [36] Steffen Rendle. 2021. Item Recommendation from Implicit Feedback. *arXiv preprint arXiv:2101.08769* (2021).
- [37] Steffen Rendle and Freudenthaler et al. 2009. BPR: Bayesian personalized ranking from implicit feedback. *UAI* (2009).
- [38] F. Ricci, L. Rokach, and Bracha Shapira. 2011. Introduction to Recommender Systems Handbook. In *Recommender Systems Handbook*.
- [39] Fei Sun, Jun Liu, Jian Wu, Changhua Pei, Xiao Lin, Wenwu Ou, and Peng Jiang. 2019. BERT4Rec: Sequential recommendation with bidirectional encoder representations from transformer. In *ACM CIKM*. 1441–1450.
- [40] Yang Sun, Fajie Yuan, Min Yang, Guoao Wei, Zhou Zhao, and Duo Liu. 2020. A Generic Network Compression Framework for Sequential Recommender Systems. In *Proceedings of the 43rd International ACM SIGIR*. 1299–1308.
- [41] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. 2016. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE CVPR*. 2818–2826.
- [42] Jiayi Tang and Ke Wang. 2018. Personalized top-n sequential recommendation via convolutional sequence embedding. In *ACM WSDM*. 565–573.
- [43] Ugo Tanielian and Flaviano Vasile. 2019. Relaxed softmax for PU learning. In *Proceedings of the 13th ACM Conference on Recommender Systems*. 119–127.
- [44] Md Mehrab Tanjim, Congzhe Su, Ethan Benjamin, Diane Hu, Liangjie Hong, and Julian McAuley. 2020. Attentive Sequential Models of Latent Intent for Next Item Recommendation. In *WWW*. 2528–2534.
- [45] Hongwei et al Wang. 2018. RippletNet: Propagating user preferences on the knowledge graph for recommender systems. In *ACM CIKM*. 417–426.
- [46] Jiachun Wang and Fajie et al Yuan. 2020. StackRec: Efficient Training of Very Deep Sequential Recommender Models by Layer Stacking. *arXiv preprint arXiv:2012.07598* (2020).
- [47] Le Wu, Xiangnan He, Xiang Wang, Kun Zhang, and Meng Wang. 2021. A Survey on Neural Recommendation: From Collaborative Filtering to Content and Context Enriched Recommendation. *arXiv preprint arXiv:2104.13030* (2021).
- [48] Hongzhi Yin, Bin Cui, Ling Chen, Zhiting Hu, and Zi Huang. 2014. A temporal context-aware model for user behavior modeling in social media systems. In *Proceedings of the 2014 ACM SIGMOD international conference on Management of data*. 1543–1554.
- [49] Runlong Yu, Qi Liu, Yuyang Ye, Mingyue Cheng, Enhong Chen, and Jianhui Ma. 2020. Collaborative List-and-Pairwise Filtering from Implicit Feedback. *IEEE Transactions on Knowledge and Data Engineering* (2020).
- [50] Runlong Yu, Yunzhou Zhang, Yuyang Ye, Le Wu, Chao Wang, Qi Liu, and Enhong Chen. 2018. Multiple pairwise ranking with implicit feedback. In *Proceedings of the 27th ACM CIKM*. 1727–1730.
- [51] Fajie Yuan, Xin Xin, Xiangnan He, Guibing Guo, Weinan Zhang, Chua Tat-Seng, and Joemon M Jose. 2018. fbgd: Learning embeddings from positive unlabeled data with bgd. (2018).
- [52] Fajie et al Yuan. 2016. Lambdafm: learning optimal ranking with factorization machines using lambda surrogates. In *ACM CIKM*. 227–236.
- [53] Fajie et al Yuan. 2019. A simple convolutional generative network for next item recommendation. In *ACM WSDM*. 582–590.
- [54] Fajie et al Yuan. 2020. Parameter-Efficient Transfer from Sequential Behaviors for User Modeling and Recommendation. In *ACM SIGIR*. 1469–1478.
- [55] Li Yuan, Francis EH Tay, and et al Li. 2019. Revisit knowledge distillation: a teacher-free framework. *arXiv preprint arXiv:1909.11723* (2019).
- [56] Sukmin Yun, Jongjin Park, Kimin Lee, and Jinwoo Shin. 2020. Regularizing class-wise predictions via self-knowledge distillation. In *CVPR*. 13876–13885.
- [57] Ying Zhang, Tao Xiang, Timothy M Hospedales, and Huchuan Lu. 2018. Deep mutual learning. In *IEEE CVPR*. 4320–4328.
- [58] Wayne Xin Zhao, Junhua Chen, Pengfei Wang, Qi Gu, and Ji-Rong Wen. 2020. Revisiting Alternative Experimental Settings for Evaluating Top-N Item Recommendation Algorithms. In *Proceedings of the 29th ACM CIKM*. 2329–2332.
- [59] Xiangyu Zhao, Liang Zhang, Zhuoye Ding, Long Xia, Jiliang Tang, and Dawei Yin. 2018. Recommendations with negative feedback via pairwise deep reinforcement learning. In *Proceedings of the 24th ACM SIGKDD*. 1040–1048.