References

[1] Rajendra Akerkar and Priti Sajja. Knowledge-based systems. Jones & Bartlett Publishers, 2009.

[2] Charalampos P Andriotis and Konstantinos G Papakonstantinou. "Managing engineering systems with large state and action spaces through deep reinforcement learning". In: Reliability Engineering & System Safety 191 (2019), p. 106483.

[3] NICTA ANU. AutoRec: Autoencoders Meet Collaborative Filtering.

[4] R Baeza-Yates. "Modern Information Retrieval". In: Addison Wesley google schola 2 (1999), pp. 127–136.

[5] O. Bastani. "Sample Complexity of Estimating the Policy Gradient for Nearly Deterministic Dynamical Systems". In: arXiv preprint arXiv:1901.08562 (2019). DOI: 10.48550/arxiv.1901.08562. URL: https://doi.org/10.48550/arxiv.1901.08562.

[6] Nicholas J. Belkin and W. Bruce Croft. "Information Filtering and Information Retrieval: Two Sides of the Same Coin?" In: Commun. ACM 35.12 (Dec. 1992), pp. 29–38. ISSN: 0001-0782. DOI: 10.1145/138859.138861. URL: https://doi.org/10.1145/138859.138861.

[7] Uzair Aslam Bhatti et al. "Recommendation system for immunization coverage and monitoring". In: Human vaccines & immunotherapeutics 14.1 (2018), pp. 165–171.

[8] Uzair Aslam Bhatti et al. “Recommendation system using feature extraction and pattern recognition in clinical care systems”. In: Enterprise information systems 13.3 (2019), pp. 329–351.

[9] John S Breese, David Heckerman, and Carl Kadie. “Empirical analysis of predictive algorithms for collaborative filtering”. In: arXiv preprint arXiv:1301.7363 (2013).

[10] Mourad Brik and Mohamed Touahria. “Contextual Information Retrieval within Recommender System: Case Study" E-learning System"”. In: TEM Journal 9.3 (2020), p. 1150.

[11] Greg Brockman et al. “Openai gym”. In: arXiv preprint arXiv:1606.01540 (2016).

[12] Chris Burges et al. “Learning to rank using gradient descent”. In: Proceedings of the 22nd international conference on Machine learning. 2005, pp. 89–96.

[13] Christopher JC Burges. “From ranknet to lambdarank to lambdamart: An overview”. In: Learning 11.23-581 (2010), p. 81.

[14] Robin Burke. “Hybrid recommender systems: Survey and experiments”. In: User modeling and user-adapted interaction 12 (2002), pp. 331–370.

[15] Lucian Busoniu, Robert Babuska, and Bart De Schutter. “A comprehensive survey of multiagent reinforcement learning”. In: IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 38.2 (2008), pp. 156–172.

[16] Fatih Cakir et al. “Deep metric learning to rank”. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019, pp. 1861–1870.

[17] Zhe Cao et al. “Learning to rank: from pairwise approach to listwise approach”. In: Proceedings of the 24th international conference on Machine learning. 2007, pp. 129–136.

[18] Soumen Chakrabarti et al. “Structured learning for non-smooth ranking losses”. In: Proceedings of the 14th ACM SIGKDD international conference on knowledge discovery and data mining. 2008, pp. 88–96.

[19] Haicheng Chen. “A DQN-based recommender system for item-list recommendation”. In: 2021 IEEE International Conference on Big Data (Big Data). IEEE. 2021, pp. 5699–5702.

[20] Minmin Chen et al. “Top-k off-policy correction for a REINFORCE recommender system”. In: Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining. 2019, pp. 456–464.

[21] Xiaocong Chen et al. “A survey of deep reinforcement learning in recommender systems: A systematic review and future directions”. In: arXiv preprint arXiv:2109.03540 (2021).

[22] Xiaocong Chen et al. “A survey of deep reinforcement learning in recommender systems: A systematic review and future directions”. In: arXiv preprint arXiv:2109.03540 (2021).

[23] Tianshu Chu et al. “Multi-agent deep reinforcement learning for large-scale traffic signal control”. In: IEEE Transactions on Intelligent Transportation Systems 21.3 (2019), pp. 1086–1095.

[24] Tianshu Chu et al. “Multi-agent deep reinforcement learning for large-scale traffic signal control”. In: IEEE Transactions on Intelligent Transportation Systems 21.3 (2019), pp. 1086–1095.

[25] Stéphane Clinchant and Eric Gaussier. “A Theoretical Analysis of Pseudo-Relevance Feedback Models”. In: Proceedings of the 2013 Conference on the Theory of Information Retrieval. ICTIR ’13. Copenhagen, Denmark: Association for Computing Machinery, 2013, pp. 6–13. ISBN: 9781450321075. DOI: 10.1145/2499178.2499179. URL: https://doi.org/10.1145/2499178.2499179.

[26] Koby Crammer and Yoram Singer. “Pranking with ranking”. In: Advances in neural information processing systems 14 (2001).

[27] Abhishek Das et al. “Tarmac: Targeted multi-agent communication”. In: International Conference on Machine Learning. PMLR. 2019, pp. 1538–1546.

[28] H. Dong et al. Deep Reinforcement Learning: Fundamentals, Research, and Applications. Springer Nature, 2020.

[29] Ali Dorri, Salil S Kanhere, and Raja Jurdak. “Multi-agent systems: A survey”. In: Ieee Access 6 (2018), pp. 28573–28593.

[30] Gabriel Dulac-Arnold et al. “Challenges of real-world reinforcement learning: definitions, benchmarks and analysis”. In: Machine Learning 110.9 (2021), pp. 2419–2468.

[31] Gabriel Dulac-Arnold et al. “Deep reinforcement learning in large discrete action spaces”. In: arXiv preprint arXiv:1512.07679 (2015).

[32] Jun Feng et al. “Learning to Collaborate: Multi-Scenario Ranking via Multi-Agent Reinforcement Learning”. In: Proceedings of the 2018 World Wide Web Conference. WWW ’18. Lyon, France: International World Wide Web Conferences Steering Committee, 2018, pp. 1939–1948. ISBN: 9781450356398. DOI: 10.1145/3178876.3186165. URL: https://doi.org/10.1145/3178876.3186165.

[33] Jun Feng et al. “Learning to Collaborate: Multi-Scenario Ranking via Multi-Agent Reinforcement Learning”. In: Proceedings of the 2018 World Wide Web Conference. WWW ’18. Lyon, France: International World Wide Web Conferences Steering Committee, 2018, pp. 1939–1948. ISBN: 9781450356398. DOI: 10.1145/3178876.3186165. URL: https://doi.org/10.1145/3178876.3186165.

[34] Yue Feng et al. “From Greedy Selection to Exploratory Decision-Making: Diverse Ranking with Policy-Value Networks”. In: The 41st International ACM SIGIR Conference on Research and Development in Information Retrieval. SIGIR ’18. Ann Arbor, MI, USA: Association for Computing Machinery, 2018, pp. 125–134. ISBN: 9781450356572. DOI: 10.1145/3209978.3209979. URL: https://doi.org/10.1145/3209978.3209979.

[35] S. Firouzi. “Learning to Rank from Noisy and Biased Click Data”. In: Proceedings of the 27th ACM International Conference on Information and Knowledge Management (CIKM 2018). ACM, 2018. DOI: DOIifavailable. URL: URL%20if%20available.

[36] Jakob N Foerster et al. “Learning with opponent-learning awareness”. In: arXiv preprint arXiv:1709.04326 (2017).

[37] Scott Fujimoto, Herke Hoof, and David Meger. “Addressing function approximation error in actor-critic methods”. In: International conference on machine learning. PMLR. 2018, pp. 1587–1596.

[38] Jiafeng Guo et al. “A deep relevance matching model for ad-hoc retrieval”. In: Proceedings of the 25th ACM international on conference on information and knowledge management. 2016, pp. 55–64.

[39] Chuan He et al. “A survey on learning to rank”. In: 2008 International Conference on Machine Learning and Cybernetics. Vol. 3. Ieee. 2008, pp. 1734–1739.

[40] Xiangnan He et al. “Neural collaborative filtering”. In: Proceedings of the 26th international conference on world wide web. 2017, pp. 173–182.

[41] Katja Hofmann, Shimon Whiteson, and Maarten Rijke. “Balancing Exploration and Exploitation in Listwise and Pairwise Online Learning to Rank for Information Retrieval”. In: Inf. Retr. 16.1 (Feb. 2013), pp. 63–90. ISSN: 1386-4564. DOI: 10.1007/s10791-012-9197-9. URL: https://doi.org/10.1007/s10791-012-9197-9.

[42] Yujing Hu et al. “Reinforcement learning to rank in e-commerce search engine: Formalization, analysis, and application”. In: Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining. 2018, pp. 368–377.

[43] Kalervo Järvelin and Jaana Kekäläinen. “Cumulated Gain-Based Evaluation of IR Techniques”. In: ACM Trans. Inf. Syst. 20.4 (Oct. 2002), pp. 422–446.

ISSN: 1046-8188. DOI: 10.1145/582415.582418. URL: https://doi.org/10.1145/582415.582418.

[44] Kalervo Järvelin and Jaana Kekäläinen. "IR Evaluation Methods for Retrieving Highly Relevant Documents". In: Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. SIGIR '00. Athens, Greece: Association for Computing Machinery, 2000, pp. 41–48. ISBN: 1581132263. DOI: 10.1145/345508.345545. URL: https://doi.org/10.1145/345508.345545.

[45] Thorsten Joachims. "Optimizing search engines using clickthrough data". In: Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining. 2002, pp. 133–142.

[46] Thorsten Joachims, Adith Swaminathan, and Tobias Schnabel. "Unbiased Learning-to-Rank with Biased Feedback". In: Proceedings of the Tenth ACM International Conference on Web Search and Data Mining. WSDM '17. Cambridge, United Kingdom: Association for Computing Machinery, 2017, pp. 781–789. ISBN: 9781450346757. DOI: 10.1145/3018661.3018699. URL: https://doi.org/10.1145/3018661.3018699.

[47] Sumitkumar Kanoje, Debajyoti Mukhopadhyay, and Sheetal Girase. "User profiling for university recommender system using automatic information retrieval". In: Procedia Computer Science 78 (2016), pp. 5–12.

[48] Vijay Konda and John Tsitsiklis. "Actor-critic algorithms". In: Advances in neural information processing systems 12 (1999).

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

[50] Branislav Kveton et al. "Cascading Bandits: Learning to Rank in the Cascade Model". In: Proceedings of the 32nd International Conference on Machine Learning. Ed. by Francis Bach and David Blei. Vol. 37. Proceedings of

Machine Learning Research. Lille, France: PMLR, July 2015, pp. 767–776. URL: https://proceedings.mlr.press/v37/kveton15.html.

[51] Lihong Li et al. "A contextual-bandit approach to personalized news article recommendation". In: Proceedings of the 19th international conference on World wide web. 2010, pp. 661–670.

[52] Lihong Li et al. "A contextual-bandit approach to personalized news article recommendation". In: Proceedings of the 19th international conference on World wide web. 2010, pp. 661–670.

[53] Shihui Li et al. "Robust multi-agent reinforcement learning via minimax deep deterministic policy gradient". In: Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 33. 01. 2019, pp. 4213–4220.

[54] Weimin Li et al. "Personalization recommendation algorithm based on trust correlation degree and matrix factorization". In: IEEE Access 7 (2019), pp. 45451–45459.

[55] Timothy P Lillicrap et al. "Continuous control with deep reinforcement learning". In: arXiv preprint arXiv:1509.02971 (2015).

[56] Wenmin Lin et al. "Location-aware service recommendations with privacy-preservation in the Internet of Things". In: IEEE Transactions on Computational Social Systems 8.1 (2020), pp. 227–235.

[57] Yuanguo Lin et al. "A survey on reinforcement learning for recommender systems". In: arXiv preprint arXiv:2109.10665 (2021).

[58] Feng Liu et al. "Deep reinforcement learning based recommendation with explicit user-item interactions modeling". In: arXiv preprint arXiv:1810.12027 (2018).

[59] Feng Liu et al. "Novel Approaches to Accelerating the Convergence Rate of Markov Decision Process for Search Result Diversification". In: CoRR abs/1802.08401 (2018). arXiv: 1802.08401. URL: http://arxiv.org/abs/1802.08401.

[60] Tie-Yan Liu et al. “Learning to rank for information retrieval”. In: Foundations and Trendső in Information Retrieval 3.3 (2009), pp. 225–331.

[61] Yu Liu et al. “A novel deep hybrid recommender system based on auto-encoder with neural collaborative filtering”. In: Big Data Mining and Analytics 1.3 (2018), pp. 211–221.

[62] Renzhi Lu et al. “Reward Shaping-Based Actor-Critic Deep Reinforcement Learning for Residential Energy Management”. In: IEEE Transactions on Industrial Informatics (2022), pp. 1–12. DOI: 10.1109/TII.2022.3183802.

[63] Jiyun Luo, Sicong Zhang, and Hui Yang. “Win-Win Search: Dual-Agent Stochastic Game in Session Search”. In: Proceedings of the 37th International ACM SIGIR Conference on Research and Development in Information Retrieval. SIGIR ’14. Gold Coast, Queensland, Australia: Association for Computing Machinery, 2014, pp. 587–596. ISBN: 9781450322577. DOI: 10.1145/2600428.2609629. URL: https://doi.org/10.1145/2600428.2609629.

[64] Xueguang Lyu et al. “Contrasting centralized and decentralized critics in multi-agent reinforcement learning”. In: arXiv preprint arXiv:2102.04402 (2021).

[65] Andriy Mnih and Russ R Salakhutdinov. “Probabilistic matrix factorization”. In: Advances in neural information processing systems 20 (2007).

[66] Volodymyr Mnih et al. “Asynchronous methods for deep reinforcement learning”. In: International conference on machine learning. PMLR. 2016, pp. 1928–1937.

[67] Volodymyr Mnih et al. “Human-level control through deep reinforcement learning”. In: nature 518.7540 (2015), pp. 529–533.

[68] Ali Montazeralghaem, Hamed Zamani, and James Allan. “A Reinforcement Learning Framework for Relevance Feedback”. In: Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval. New York, NY, USA: Association for Computing Ma-

chinery, 2020, pp. 59–68. ISBN: 9781450380164. URL: https://doi.org/10.1145/3397271.3401099.

[69] Ali Montazeralghaem, Hamed Zamani, and James Allan. "A reinforcement learning framework for relevance feedback". In: Proceedings of the 43rd international acm sigir conference on research and development in information retrieval. 2020, pp. 59–68.

[70] Keiron O'Shea and Ryan Nash. "An introduction to convolutional neural networks". In: arXiv preprint arXiv:1511.08458 (2015).

[71] Bing Pan et al. "In Google we trust: Users decisions on rank, position, and relevance". In: Journal of computer-mediated communication 12.3 (2007), pp. 801–823.

[72] Feiyang Pan et al. "Policy gradients for contextual recommendations". In: The World Wide Web Conference. 2019, pp. 1421–1431.

[73] Liang Pang et al. "Setrank: Learning a permutation-invariant ranking model for information retrieval". In: Proceedings of the 43rd international ACM SIGIR conference on research and development in information retrieval. 2020, pp. 499–508.

[74] Romain Paulus, Caiming Xiong, and Richard Socher. "A deep reinforced model for abstractive summarization". In: arXiv preprint arXiv:1705.04304 (2017).

[75] Huseyin Polat and Wenliang Du. "SVD-Based Collaborative Filtering with Privacy". In: Proceedings of the 2005 ACM Symposium on Applied Computing. SAC '05. Santa Fe, New Mexico: Association for Computing Machinery, 2005, pp. 791–795. ISBN: 1581139640. DOI: 10.1145/1066677.1066860. URL: https://doi.org/10.1145/1066677.1066860.

[76] Tao Qin et al. "LETOR: A Benchmark Collection for Research on Learning to Rank for Information Retrieval". In: Inf. Retr. 13.4 (Aug. 2010), pp. 346–374. ISSN: 1386-4564. DOI: 10.1007/s10791-009-9123-y. URL: https://doi.org/10.1007/s10791-009-9123-y.

[77] Filip Radlinski, Robert Kleinberg, and Thorsten Joachims. “Learning diverse rankings with multi-armed bandits”. In: Proceedings of the 25th international conference on Machine learning. 2008, pp. 784–791.

[78] Antonin Raffin et al. “Stable-baselines3: Reliable reinforcement learning implementations”. In: The Journal of Machine Learning Research 22.1 (2021), pp. 12348–12355.

[79] Ashwini Rahangdale and Shital Raut. “Deep neural network regularization for feature selection in learning-to-rank”. In: vol. 7. IEEE, 2019, pp. 53988–54006.

[80] Marc'Aurelio Ranzato et al. “Sequence level training with recurrent neural networks”. In: arXiv preprint arXiv:1511.06732 (2015).

[81] Ruslan Salakhutdinov and Andriy Mnih. “Bayesian probabilistic matrix factorization using Markov chain Monte Carlo”. In: Proceedings of the 25th international conference on Machine learning. 2008, pp. 880–887.

[82] Ruslan Salakhutdinov, Andriy Mnih, and Geoffrey Hinton. “Restricted Boltzmann machines for collaborative filtering”. In: Proceedings of the 24th international conference on Machine learning. 2007, pp. 791–798.

[83] Badrul Sarwar et al. “Item-based collaborative filtering recommendation algorithms”. In: Proceedings of the 10th international conference on World Wide Web. 2001, pp. 285–295.

[84] John Schulman et al. “Proximal policy optimization algorithms”. In: arXiv preprint arXiv:1707.06347 (2017).

[85] Guy Shani et al. “An MDP-based recommender system.” In: Journal of Machine Learning Research 6.9 (2005).

[86] Piyush K Sharma et al. “Survey of recent multi-agent reinforcement learning algorithms utilizing centralized training”. In: Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications III. Vol. 11746. SPIE. 2021, pp. 665–676.

[87] Yue Shi, Martha Larson, and Alan Hanjalic. "List-Wise Learning to Rank with Matrix Factorization for Collaborative Filtering". In: Proceedings of the Fourth ACM Conference on Recommender Systems. RecSys '10. Barcelona, Spain: Association for Computing Machinery, 2010, pp. 269–272. ISBN: 9781605589060. DOI: 10.1145/1864708.1864764. URL: https://doi.org/10.1145/1864708.1864764.

[88] David Simões, Nuno Lau, and Luís Paulo Reis. "Multi-agent actor centralized-critic with communication". In: Neurocomputing 390 (2020), pp. 40–56.

[89] Aleksandrs Slivkins, Filip Radlinski, and Sreenivas Gollapudi. "Ranked Bandits in Metric Spaces: Learning Diverse Rankings over Large Document Collections". In: J. Mach. Learn. Res. 14.1 (Feb. 2013), pp. 399–436. ISSN: 1532-4435.

[90] Anongnart Srivihok and Pisit Sukonmanee. "E-commerce intelligent agent: personalization travel support agent using Q Learning". In: Proceedings of the 7th international conference on Electronic commerce. 2005, pp. 287–292.

[91] Anongnart Srivihok and Pisit Sukonmanee. "E-commerce intelligent agent: personalization travel support agent using Q Learning". In: Proceedings of the 7th international conference on Electronic commerce. 2005, pp. 287–292.

[92] Richard S Sutton, Satinder Singh, and David McAllester. "Comparing policy-gradient algorithms". In: IEEE Transactions on Systems, Man, and Cybernetics (2000).

[93] Richard S Sutton et al. "Policy gradient methods for reinforcement learning with function approximation". In: Advances in neural information processing systems 12 (1999).

[94] Richard S. Sutton and Andrew G. Barto. Reinforcement Learning: An Introduction. Second. The MIT Press, 2018. URL: http://incompleteideas.net/book/the-book-2nd.html.

[95] Nima Taghipour and Ahmad Kardan. “A hybrid web recommender system based on q-learning”. In: Proceedings of the 2008 ACM symposium on Applied computing. 2008, pp. 1164–1168.

[96] Jiaxi Tang and Ke Wang. “Personalized Top-N Sequential Recommendation via Convolutional Sequence Embedding”. In: Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining. WSDM ’18. Marina Del Rey, CA, USA: Association for Computing Machinery, 2018, pp. 565–573. ISBN: 9781450355810. DOI: 10.1145/3159652.3159656. URL: https://doi.org/10.1145/3159652.3159656.

[97] Ye Tao et al. “Item trend learning for sequential recommendation system using gated graph neural network”. In: Neural Computing and Applications (2021), pp. 1–16.

[98] Michael Taylor et al. “Softrank: optimizing non-smooth rank metrics”. In: Proceedings of the 2008 International Conference on Web Search and Data Mining. 2008, pp. 77–86.

[99] Justin K Terry et al. “Revisiting parameter sharing in multi-agent deep reinforcement learning”. In: arXiv preprint arXiv:2005.13625 (2020).

[100] Hamed Valizadegan et al. “Learning to Rank by Optimizing NDCG Measure”. In: NIPS. Vol. 22. 2009, pp. 1883–1891.

[101] Aaron Van den Oord, Sander Dieleman, and Benjamin Schrauwen. “Deep content-based music recommendation”. In: Advances in neural information processing systems 26 (2013).

[102] Hado Van Hasselt et al. “Deep reinforcement learning and the deadly triad”. In: arXiv preprint arXiv:1812.02648 (2018).

[103] Pascal Vincent et al. “Extracting and composing robust features with denoising autoencoders”. In: Proceedings of the 25th international conference on Machine learning. 2008, pp. 1096–1103.

[104] Manolis Vozalis and Konstantinos G Margaritis. “Collaborative filtering enhanced by demographic correlation”. In: AIAI symposium on professional practice in AI, of the 18th world computer congress. 2004.

[105] Manolis Vozalis and Konstantinos G Margaritis. “On the enhancement of collaborative filtering by demographic data”. In: Web Intelligence and Agent Systems: An International Journal 4.2 (2006), pp. 117–138.

[106] Hao Wang, Naiyan Wang, and Dit-Yan Yeung. “Collaborative deep learning for recommender systems”. In: Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining. 2015, pp. 1235–1244.

[107] Huazheng Wang et al. “Variance reduction in gradient exploration for online learning to rank”. In: Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval. 2019, pp. 835–844.

[108] Xiting Wang et al. “A reinforcement learning framework for explainable recommendation”. In: 2018 IEEE international conference on data mining (ICDM). IEEE. 2018, pp. 587–596.

[109] Kangning Wei, Jinghua Huang, and Shaohong Fu. “A survey of e-commerce recommender systems”. In: 2007 international conference on service systems and service management. IEEE. 2007, pp. 1–5.

[110] Zeng Wei et al. “Reinforcement Learning to Rank with Markov Decision Process”. In: Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval. SIGIR ’17. Shinjuku, Tokyo, Japan: Association for Computing Machinery, 2017, pp. 945–948. ISBN: 9781450350228. DOI: 10.1145/3077136.3080685. URL: https://doi.org/10.1145/3077136.3080685.

[111] Jason Weston, Hector Yee, and Ron J Weiss. “Learning to rank recommendations with the k-order statistic loss”. In: Proceedings of the 7th ACM Conference on Recommender Systems. 2013, pp. 245–248.

[112] Frank Wilcoxon. "Individual comparisons by ranking methods". In: Breakthroughs in Statistics: Methodology and Distribution. Springer, 1992, pp. 196–202.

[113] Cathy Wu et al. "Variance reduction for policy gradient with action-dependent factorized baselines". In: arXiv preprint arXiv:1803.07246 (2018).

[114] Yao Wu et al. "Collaborative denoising auto-encoders for top-n recommender systems". In: Proceedings of the ninth ACM international conference on web search and data mining. 2016, pp. 153–162.

[115] Long Xia et al. "Adapting Markov Decision Process for Search Result Diversification". In: Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval. SIGIR '17. Shinjuku, Tokyo, Japan: Association for Computing Machinery, 2017, pp. 535–544. ISBN: 9781450350228. DOI: 10.1145/3077136.3080775. URL: https://doi.org/10.1145/3077136.3080775.

[116] Ruobing Xie et al. "Explore, Filter and Distill: Distilled Reinforcement Learning in Recommendation". In: Proceedings of the 30th ACM international on conference on information and knowledge management. CIKM '21. Virtual Event, Queensland, Australia: Association for Computing Machinery, 2021, pp. 4243–4252. ISBN: 9781450384469. DOI: 10.1145/3459637.3481917. URL: https://doi.org/10.1145/3459637.3481917.

[117] Jun Xu and Hang Li. "AdaRank: A Boosting Algorithm for Information Retrieval". In: Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. SIGIR '07. Amsterdam, The Netherlands: Association for Computing Machinery, 2007, pp. 391–398. ISBN: 9781595935977. DOI: 10.1145/1277741.1277809. URL: https://doi.org/10.1145/1277741.1277809.

[118] Hong-Jian Xue et al. "Deep matrix factorization models for recommender systems." In: IJCAI. Vol. 17. Melbourne, Australia. 2017, pp. 3203–3209.

[119] Yaodong Yang et al. “Mean field multi-agent reinforcement learning”. In: International conference on machine learning. PMLR. 2018, pp. 5571–5580.

[120] Yiming Yang and Siddharth Gopal. “Multilabel classification with meta-level features in a learning-to-rank framework”. In: Machine Learning 88.1-2 (2012), pp. 47–68.

[121] Jing Yao et al. “RLPer: A Reinforcement Learning Model for Personalized Search”. In: Proceedings of The Web Conference 2020. New York, NY, USA: Association for Computing Machinery, 2020, pp. 2298–2308. ISBN: 9781450370233. URL: https://doi.org/10.1145/3366423.3380294.

[122] Yisong Yue and Thorsten Joachims. “Interactively optimizing information retrieval systems as a dueling bandits problem”. In: Proceedings of the 26th Annual International Conference on Machine Learning. 2009, pp. 1201–1208.

[123] Yisong Yue and Thorsten Joachims. “Interactively optimizing information retrieval systems as a dueling bandits problem”. In: Proceedings of the 26th Annual International Conference on Machine Learning. 2009, pp. 1201–1208.

[124] Yisong Yue et al. “A Support Vector Method for Optimizing Average Precision”. In: Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. SIGIR ’07. Amsterdam, The Netherlands: Association for Computing Machinery, 2007, pp. 271–278. ISBN: 9781595935977. DOI: 10.1145/1277741.1277790. URL: https://doi.org/10.1145/1277741.1277790.

[125] Wei Zeng et al. “Multi Page Search with Reinforcement Learning to Rank”. In: Proceedings of the 2018 ACM SIGIR International Conference on Theory of Information Retrieval. ICTIR ’18. Tianjin, China: Association for Computing Machinery, 2018, pp. 175–178. ISBN: 9781450356565. DOI: 10.1145/3234944.3234977. URL: https://doi.org/10.1145/3234944.3234977.

[126] Chongjie Zhang and Victor Lesser. “Coordinating multi-agent reinforcement learning with limited communication”. In: Proceedings of the 2013 inter-

national conference on Autonomous agents and multi-agent systems. 2013, pp. 1101–1108.

[127] Sicong Zhang, Jiyun Luo, and Hui Yang. "A pomdp model for content-free document re-ranking". In: Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval. 2014, pp. 1139–1142.

[128] Weijia Zhang et al. "Intelligent Electric Vehicle Charging Recommendation Based on Multi-Agent Reinforcement Learning". In: Proceedings of the Web Conference 2021. WWW '21. Ljubljana, Slovenia: Association for Computing Machinery, 2021, pp. 1856–1867. ISBN: 9781450383127. DOI: 10.1145/3442381.3449934. URL: https://doi.org/10.1145/3442381.3449934.

[129] Yang Zhang et al. "Coordination between individual agents in multi-agent reinforcement learning". In: Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 35. 13. 2021, pp. 11387–11394.

[130] Tingting Zhao et al. "Analysis and improvement of policy gradient estimation". In: Advances in Neural Information Processing Systems 24 (2011).

[131] X. Zhao, X. Li, and Z. Zhang. "Joint Structural Learning to Rank with Deep Linear Feature Learning". In: IEEE Transactions on Knowledge and Data Engineering 27.10 (2015), pp. 2756–2769. DOI: 10.1109/tkde.2015.2426707. URL: https://doi.org/10.1109/tkde.2015.2426707.

[132] Xiangyu Zhao et al. "Deep reinforcement learning for page-wise recommendations". In: Proceedings of the 12th ACM Conference on Recommender Systems. 2018, pp. 95–103.

[133] Xiangyu Zhao et al. "Recommendations with negative feedback via pairwise deep reinforcement learning". In: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2018, pp. 1040–1048.

[134] Guanjie Zheng et al. “DRN: A deep reinforcement learning framework for news recommendation”. In: Proceedings of the 2018 world wide web conference. 2018, pp. 167–176.

[135] Lulu Zheng et al. “Episodic multi-agent reinforcement learning with curiosity-driven exploration”. In: Advances in Neural Information Processing Systems 34 (2021), pp. 3757–3769.

[136] R. Zhou et al. “Anchor-changing regularized natural policy gradient for multi-objective reinforcement learning”. In: arXiv preprint arXiv:2206.05357 (2022). URL: https://doi.org/10.48550/arxiv.2206.05357.

[137] Lixin Zou et al. “Neural interactive collaborative filtering”. In: Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval. 2020, pp. 749–758.

[138] Lixin Zou et al. “Pseudo Dyna-Q: A reinforcement learning framework for interactive recommendation”. In: Proceedings of the 13th International Conference on Web Search and Data Mining. 2020, pp. 816–824.

[139] Shihao Zou et al. “Marlrank: Multi-agent reinforced learning to rank”. In: Proceedings of the 28th ACM International Conference on Information and Knowledge Management. 2019, pp. 2073–2076.

[140] Shihao Zou et al. “Marlrank: Multi-agent reinforced learning to rank”. In: Proceedings of the 28th ACM International Conference on Information and Knowledge Management. 2019, pp. 2073–2076.