

















- [22] S. Rendle. Factorization machines with libFM. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 3(3):57, 2012.
- [23] R. Rosales, H. Cheng, and E. Manavoglu. Post-click conversion modeling and analysis for non-guaranteed delivery display advertising. In *Proceedings of the fifth ACM international conference on Web search and data mining*, pages 293–302. ACM, 2012.
- [24] G. Song. Criteo display advertising challenge. Available at <https://www.kaggle.com/c/criteo-display-ad-challenge/forums/t/10547/document-and-code-for-the-3rd-place-finish>, 2014.
- [25] C.-H. Tsai, C.-Y. Lin, and C.-J. Lin. Incremental and decremental training for linear classification. In *KDD*, 2014.
- [26] F. Vasile, D. Lefortier, and O. Chapelle. Cost-sensitive learning for utility optimization in online advertising auctions. *arXiv preprint arXiv:1603.03713*, 2016.
- [27] K. Weinberger, A. Dasgupta, J. Langford, A. Smola, and J. Attenberg. Feature hashing for large scale multitask learning. In *Proceedings of the 26th Annual International Conference on Machine Learning*, pages 1113–1120. ACM, 2009.
- [28] W. Zhang, T. Du, and J. Wang. Deep learning over multi-field categorical data. In *European Conference on Information Retrieval*, pages 45–57. Springer, 2016.
- [29] M. Zinkevich, M. Weimer, L. Li, and A. J. Smola. Parallelized stochastic gradient descent. In J. D. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. S. Zemel, and A. Culotta, editors, *Advances in Neural Information Processing Systems 23*, pages 2595–2603. Curran Associates, Inc., 2010.