

References

- [1] *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann Series in Data Management Systems. Morgan Kaufmann, 3 edition, 2011.
- [2] Bernard Marr. How much data do we create every day? the mind-blowing stats everyone should read. URL: <https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/#5e3bdedf60ba>, 2018. Online; Accessed 16 March 2019.
- [3] Apache hadoop. URL: <https://hadoop.apache.org/>, 2018. Online; Accessed 16 March 2019.
- [4] Apache spark™- unified analytics engine for big data. URL: <https://spark.apache.org/>, 2018. Online; Accessed 16 March 2019.
- [5] A. Gainaru and E. Slusanschi. Framework for mapping data mining applications on gpus. In *2011 10th International Symposium on Parallel and Distributed Computing*, pages 71–78, July 2011.
- [6] Wenbin Fang, Ka Keung Lau, Mian Lu, Xiangye Xiao, Chi Kit Lam, Philip Yang Yang, Bingsheng He, Qiong Luo, Pedro V. S, and Ke Yang. Parallel data mining on graphics processors. Technical report, 2008.
- [7] Shuai Che, Michael Boyer, Jiayuan Meng, David Tarjan, Jeremy W. Sheaffer, and Kevin Skadron. A performance study of general-purpose applications on graphics processors using cuda. *Journal of Parallel and Distributed Computing*, 68(10):1370 – 1380, 2008. General-Purpose Processing using Graphics Processing Units.

- [8] C. Silvestri and S. Orlando. gpubci: Exploiting gpus in frequent itemset mining. In *2012 20th Euromicro International Conference on Parallel, Distributed and Network-based Processing*, pages 416–425, Feb 2012.
- [9] Kang-Wook Chon, Sang-Hyun Hwang, and Min-Soo Kim. Gminer: A fast gpu-based frequent itemset mining method for large-scale data. *Information Sciences*, 439-440:19 – 38, 2018.
- [10] Wenbin Fang, Mian Lu, Xiangye Xiao, Bingsheng He, and Qiong Luo. Frequent itemset mining on graphics processors. In *Proceedings of the Fifth International Workshop on Data Management on New Hardware, DaMoN '09*, pages 34–42, New York, NY, USA, 2009. ACM.
- [11] Ferenc Bodon. A trie-based apriori implementation for mining frequent item sequences. In *Proceedings of the 1st International Workshop on Open Source Data Mining: Frequent Pattern Mining Implementations, OSDM '05*, pages 56–65, New York, NY, USA, 2005. ACM.
- [12] Christian Böhm, Robert Noll, Claudia Plant, Bianca Wackersreuther, and Andrew Zherdin. *Data Mining Using Graphics Processing Units*, pages 63–90. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.
- [13] Liheng Jian, Cheng Wang, Ying Liu, Shenshen Liang, Weidong Yi, and Yong Shi. Parallel data mining techniques on graphics processing unit with compute unified device architecture (cuda). *J. Supercomput.*, 64(3):942–967, June 2013.
- [14] M. Harris. Optimizing parallel reduction in cuda. URL: <https://developer.download.nvidia.com/assets/cuda/files/reduction.pdf>, 4 2016. Online; Accessed 16-Feb-2019.
- [15] G. Andrade, F. Viegas, G. S. Ramos, J. Almeida, L. Rocha, M. Gonçalves, and R. Ferreira. Gpu-nb: A fast cuda-based implementation of naïve bayes. In *2013 25th International Symposium on Computer Architecture and High Performance Computing*, pages 168–175, Oct 2013.

- [16] John C. Shafer, Rakesh Agrawal, and Manish Mehta. Sprint: A scalable parallel classifier for data mining. In *Proceedings of the 22th International Conference on Very Large Data Bases, VLDB '96*, pages 544–555, San Francisco, CA, USA, 1996. Morgan Kaufmann Publishers Inc.
- [17] Manish Mehta, Rakesh Agrawal, and Jorma Rissanen. Sliq: A fast scalable classifier for data mining. In Peter Apers, Mokrane Bouzeghoub, and Georges Gardarin, editors, *Advances in Database Technology — EDBT '96*, pages 18–32, Berlin, Heidelberg, 1996. Springer Berlin Heidelberg.
- [18] Yisheng Liao, Alex Rubinsteyn, Russell Power, and Jinyang Li. Learning random forests on the gpu. 12 2013.
- [19] Patrick O’Neil and Dallan Quass. Improved query performance with variant indexes. In *Proceedings of the 1997 ACM SIGMOD International Conference on Management of Data, SIGMOD '97*, pages 38–49, New York, NY, USA, 1997. ACM.
- [20] Ian H. Witten, Eibe Frank, and Mark A. Hall. *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 3rd edition, 2011.
- [21] Cécile Favre and Fadila Bentayeb. Bitmap index-based decision trees. In *Proceedings of the 15th International Conference on Foundations of Intelligent Systems, ISMIS'05*, pages 65–73, Berlin, Heidelberg, 2005. Springer-Verlag.
- [22] Weka-c++. URL: <https://sourceforge.net/p/wekacpp/>, 2007. Online; Accessed 16 March 2019.
- [23] Dheeru Dua and Casey Graff. UCI machine learning repository, 2017.
- [24] Us census data (1990) data set. URL: [https://archive.ics.uci.edu/ml/datasets/US+Census+Data+\(1990\)](https://archive.ics.uci.edu/ml/datasets/US+Census+Data+(1990)), 1990. Online; Accessed 16 March 2019.
- [25] Poker hand data set. URL: <https://archive.ics.uci.edu/ml/datasets/Poker+Hand>, 1990. Online; Accessed 16 March 2019.

- [26] Kdd cup (1999). kdd cup 99 intrusion detection datasets. URL: <http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html>, 1999. Online; Accessed 16 March 2019.