

REFERENCES

- [1] Constantin F Aliferis, Alexander Statnikov, Ioannis Tsamardinos, Subramani Mani, and Xenofon D Koutsoukos. 2010. Local causal and Markov blanket induction for causal discovery and feature selection for classification part i: Algorithms and empirical evaluation. *Journal of Machine Learning Research* 11, 1 (2010).
- [2] Kailash Budhathoki and Jilles Vreeken. 2018. Origo: causal inference by compression. *Knowledge and Information Systems* 56, 2 (2018), 285–307.
- [3] Ruichu Cai, Zhenjie Zhang, and Zhifeng Hao. 2011. BASSUM: A Bayesian semi-supervised method for classification feature selection. *Pattern Recognition* 44, 4 (2011), 811–820.
- [4] Ruichu Cai, Zhenjie Zhang, and Zhifeng Hao. 2013. Causal gene identification using combinatorial V-structure search. *Neural Networks* 43 (2013), 63–71.
- [5] Ruichu Cai, Zhenjie Zhang, and Zhifeng Hao. 2013. SADA: A general framework to support robust causation discovery. In *International conference on machine learning*. PMLR, 208–216.
- [6] David Maxwell Chickering. 2002. Learning equivalence classes of Bayesian-network structures. *Journal of machine learning research* 2, Feb (2002), 445–498.
- [7] Anthony Costa Constantinou, Mark Freestone, William Marsh, and Jeremy Coid. 2015. Causal inference for violence risk management and decision support in forensic psychiatry. *Decision Support Systems* 80 (2015), 42–55.
- [8] Gary Doran, Krikamol Muandet, Kun Zhang, and Bernhard Schölkopf. 2014. A permutation-based kernel conditional independence test. In *Proceedings of the Thirtieth Conference on Uncertainty in Artificial Intelligence*. 132–141.
- [9] Byron Ellis and Wing Hung Wong. 2008. Learning causal Bayesian network structures from experimental data. *J. Amer. Statist. Assoc.* 103, 482 (2008), 778–789.
- [10] Kenji Fukumizu, Arthur Gretton, Xiaohai Sun, and Bernhard Schölkopf. 2008. Kernel measures of conditional dependence. In *Advances in neural information processing systems*. 489–496.
- [11] Zhi Geng, Chi Wang, and Qiang Zhao. 2005. Decomposition of search for V-structures in DAGs. *Journal of Multivariate Analysis* 96, 2 (2005), 282–294.
- [12] Mohammad Ali Javidian, Marco Valtorta, and Pooyan Jamshidi. 2019. Order-independent structure learning of multivariate regression chain graphs. In *International Conference on Scalable Uncertainty Management*. Springer, 324–338.
- [13] Mohammad Ali Javidian, Marco Valtorta, and Pooyan Jamshidi. 2020. AMP chain graphs: Minimal separators and structure learning algorithms. *Journal of Artificial Intelligence Research* 69 (2020), 419–470.
- [14] Mohammad Ali Javidian, Marco Valtorta, and Pooyan Jamshidi. 2020. Learning LWF chain graphs: A Markov blanket discovery approach. In *Conference on Uncertainty in Artificial Intelligence*. PMLR, 1069–1078.
- [15] Mohammad Ali Javidian, Zhiyu Wang, Linyuan Lu, and Marco Valtorta. 2020. On a hypergraph probabilistic graphical model. *Annals of Mathematics and Artificial Intelligence* 88, 9 (2020), 1003–1033.
- [16] Jinkyu Kim and John Canny. 2017. Interpretable learning for self-driving cars by visualizing causal attention. In *Proceedings of the IEEE international conference on computer vision*. 2942–2950.
- [17] Daphne Koller and Nir Friedman. 2009. *Probabilistic graphical models: principles and techniques*. MIT press.
- [18] Hui Liu, Shuigeng Zhou, Wai Lam, and Jihong Guan. 2017. A new hybrid method for learning Bayesian networks: Separation and reunion. *Knowledge-Based Systems* 121 (2017), 185–197.
- [19] Saaduddin Mahmud, Md. Mosaddek Khan, Moumita Choudhury, Long Tran-Thanh, and Nicholas R. Jennings. 2020. Learning optimal temperature region for solving mixed integer functional DCOPs. *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence* (Jul 2020).
- [20] Judea Pearl. 2000. *Models, reasoning and inference*. Cambridge, UK: Cambridge University Press (2000).
- [21] Judea Pearl. 2009. *Causality*. Cambridge University Press.
- [22] Jonas Peters, Dominik Janzing, and Bernhard Schölkopf. 2011. Causal inference on discrete data using additive noise models. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33, 12 (2011), 2436–2450.
- [23] Franz Josef Radermacher. 1990. Probabilistic reasoning in intelligent systems: Networks of plausible inference (Judea Pearl). *Society for Industrial and Applied Mathematics Review* 32, 4 (1990), 704–707.
- [24] Mashrur Rashik, Md Musfiqur Rahman, Md Mosaddek Khan, Md Mamun-or Rashid, Long Tran-Thanh, and Nicholas R Jennings. 2020. Speeding up distributed pseudo-tree optimization procedures with cross edge consistency to solve DCOPs. *Applied Intelligence* (2020), 1–14.
- [25] Peter Spirtes, Clark N Glymour, Richard Scheines, and David Heckerman. 2000. *Causation, prediction, and search*. MIT press.
- [26] Xianchao Xie and Zhi Geng. 2008. A recursive method for structural learning of directed acyclic graphs. *Journal of Machine Learning Research* 9, Mar (2008), 459–483.
- [27] Xianchao Xie, Zhi Geng, and Qiang Zhao. 2006. Decomposition of structural learning about directed acyclic graphs. *Artificial Intelligence* 170, 4-5 (2006), 422–439.
- [28] Chuanxu Yan and Shuigeng Zhou. 2020. Effective and scalable causal partitioning based on low-order conditional independent tests. *Neurocomputing* (2020).
- [29] Changwon Yoo and Gregory F Cooper. 2002. Discovery of gene-regulation pathways using local causal search. In *Proceedings of the AMIA Symposium*. American Medical Informatics Association, 914.
- [30] Hao Zhang, Shuigeng Zhou, Chuanxu Yan, Jihong Guan, and Xin Wang. 2019. Recursively learning causal structures using regression-based conditional independence test. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. 3108–3115.
- [31] Kun Zhang, Jonas Peters, Dominik Janzing, and Bernhard Schölkopf. 2011. Kernel-based conditional independence test and application in causal discovery. In *Proceedings of the Twenty-Seventh Conference on Uncertainty in Artificial Intelligence* (Barcelona, Spain) (UAI’11). AUAI Press, Arlington, Virginia, USA, 804–813.