



Universiteit
Leiden

The Netherlands

Information-theoretic partition-based models for interpretable machine learning

Yang, L.

Citation

Yang, L. (2024, September 20). *Information-theoretic partition-based models for interpretable machine learning*. SIKS Dissertation Series. Retrieved from <https://hdl.handle.net/1887/4092882>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4092882>

Note: To cite this publication please use the final published version (if applicable).

Bibliography

- Abdelhamid, Neda and Fadi Thabtah (2014). “Associative classification approaches: review and comparison”. In: *Journal of Information & Knowledge Management* 13.03.
- Adebayo, Julius, Justin Gilmer, Michael Muelly, Ian Goodfellow, Moritz Hardt, and Been Kim (2018). “Sanity checks for saliency maps”. In: *Advances in neural information processing systems* 31.
- Angelino, Elaine, Nicholas Larus-Stone, Daniel Alabi, Margo Seltzer, and Cynthia Rudin (2017). “Learning certifiably optimal rule lists for categorical data”. In: *arXiv preprint arXiv:1704.01701*.
- Antos, András and Ioannis Kontoyiannis (2001). “Convergence properties of functional estimates for discrete distributions”. In: *Random Structures & Algorithms* 19.3-4, pp. 163–193.
- Bay, Stephen D (2001). “Multivariate discretization for set mining”. In: *Knowledge and Information Systems* 3.4, pp. 491–512.
- Biba, Marenglen, Floriana Esposito, Stefano Ferilli, Nicola Di Mauro, Teresa Maria Altomare Basile, et al. (2007). “Unsupervised Discretization Using Kernel Density Estimation.” In: *International Joint Conference on Artificial Intelligence*. Vol. 7, pp. 696–701.
- Boullé, Marc (2004). “Khiops: A statistical discretization method of continuous attributes”. In: *Machine learning* 55.1, pp. 53–69.
- (2006). “MODL: a Bayes optimal discretization method for continuous attributes”. In: *Machine learning* 65.1, pp. 131–165.
- Breiman, Leo, Jerome Friedman, Charles J Stone, and Richard A Olshen (1984). *Classification and regression trees*. CRC press.

Bibliography

- Cabeli, Vincent, Louis Verny, Nadir Sella, Guido Uguzzoni, Marc Verny, and Hervé Isambert (2020). “Learning clinical networks from medical records based on information estimates in mixed-type data”. In: *PLoS computational biology* 16.5, e1007866.
- Caruana, Rich, Yin Lou, Johannes Gehrke, Paul Koch, Marc Sturm, and Noemie Elhadad (2015). “Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission”. In: *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 1721–1730.
- Chen, Guoqing, Hongyan Liu, Lan Yu, Qiang Wei, and Xing Zhang (2006). “A new approach to classification based on association rule mining”. In: *Decision Support Systems* 42.2, pp. 674–689.
- Clark, Peter and Robin Boswell (1991). “Rule induction with CN2: Some recent improvements”. In: *European Working Session on Learning*. Springer, pp. 151–163.
- Clark, Peter and Tim Niblett (1989). “The CN2 induction algorithm”. In: *Machine learning* 3.4, pp. 261–283.
- Cohen, William W (1995). “Fast effective rule induction”. In: *Machine learning proceedings 1995*. Elsevier, pp. 115–123.
- Colombo, Diego and Marloes H Maathuis (2012). “A modification of the PC algorithm yielding order-independent skeletons”. In: *CoRR*, abs/1211.3295.
- Cover, Thomas M and Joy A Thomas (2012). *Elements of information theory*. John Wiley & Sons.
- Cuevas, Antonio, Ricardo Fraiman, et al. (1997). “A plug-in approach to support estimation”. In: *The Annals of Statistics* 25.6, pp. 2300–2312.
- Darbellay, Georges A and Igor Vajda (1999). “Estimation of the information by an adaptive partitioning of the observation space”. In: *IEEE Transactions on Information Theory* 45.4, pp. 1315–1321.
- Dash, Sanjeeb, Oktay Gunluk, and Dennis Wei (2018). “Boolean Decision Rules via Column Generation”. In: *Advances in Neural Information Processing Systems* 31, pp. 4655–4665.
- De Bie, Tijl (2011a). “An information theoretic framework for data mining”. In: *Proceedings of the 17th ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 564–572.

-
- (2011b). “Maximum entropy models and subjective interestingness: an application to tiles in binary databases”. In: *Data Mining and Knowledge Discovery* 23, pp. 407–446.
- Demšar, Janez, Tomaž Curk, Aleš Erjavec, Črt Gorup, Tomaž Hočevár, Mitar Milutinović, Martin Možina, Matija Polajnar, Marko Toplak, Anže Starič, Miha Štajdohar, Lan Umek, Lan Žagar, Jure Žbontar, Marinka Žitnik, and Blaž Zupan (2013). “Orange: Data Mining Toolbox in Python”. In: *Journal of Machine Learning Research* 14, pp. 2349–2353. URL: <http://jmlr.org/papers/v14/demsar13a.html>.
- Doshi-Velez, Finale and Been Kim (2017). “Towards a rigorous science of interpretable machine learning”. In: *arXiv preprint arXiv:1702.08608*.
- Dua, Dheeru and Casey Graff (2017). *UCI Machine Learning Repository*. URL: <http://archive.ics.uci.edu/ml>.
- Duong, Tarn and Martin Hazelton (2003). “Plug-in bandwidth matrices for bivariate kernel density estimation”. In: *Journal of Nonparametric Statistics* 15.1, pp. 17–30.
- Duong, Tarn et al. (2007). “ks: Kernel density estimation and kernel discriminant analysis for multivariate data in R”. In: *Journal of Statistical Software* 21.7, pp. 1–16.
- Fayyad, Usama and Keki Irani (1993). “Multi-interval discretization of continuous-valued attributes for classification learning”. In: *Proceedings of the 13th International Joint Conference on Artificial Intelligence (IJCAI-93)*, pp. 1022–1027.
- Ferrandiz, Sylvain and Marc Boullé (2005). “Multivariate discretization by recursive supervised bipartition of graph”. In: *International Workshop on Machine Learning and Data Mining in Pattern Recognition*. Springer, pp. 253–264.
- Filip, Jiri and Tomas Kliegr (2019). “PyIDS-Python Implementation of Interpretable Decision Sets Algorithm by Lakkaraju et al, 2016.” In: *RuleML+RR (Supplement)*.
- Frank, Eibe and Ian H Witten (1998). “Generating accurate rule sets without global optimization”. In: *Working Paper Series ISSN 1170-487X*.
- Frenzel, Stefan and Bernd Pompe (2007). “Partial mutual information for coupling analysis of multivariate time series”. In: *Physical Review Letters* 99.20, p. 204101.

Bibliography

- Friedman, Jerome, Trevor Hastie, and Robert Tibshirani (2001). *The elements of statistical learning*. Vol. 1. 10. Springer series in statistics New York.
- Friedman, Jerome H and Nicholas I Fisher (1999). “Bump hunting in high-dimensional data”. In: *Statistics and computing* 9.2, pp. 123–143.
- Fürnkranz, Johannes and Peter A Flach (2005). “Roc ‘n’ rule learning—towards a better understanding of covering algorithms”. In: *Machine learning* 58.1, pp. 39–77.
- Fürnkranz, Johannes, Dragan Gamberger, and Nada Lavrač (2012). *Foundations of rule learning*. Springer Science & Business Media.
- Galbrun, Esther (2020). “The minimum description length principle for pattern mining: A survey”. In: *arXiv preprint arXiv:2007.14009*.
- (2022). “The minimum description length principle for pattern mining: A survey”. In: *Data mining and knowledge discovery* 36.5, pp. 1679–1727.
- Gao, Weihao, Sreeram Kannan, Sewoong Oh, and Pramod Viswanath (2017). “Estimating mutual information for discrete-continuous mixtures”. In: *Advances in neural information processing systems*, pp. 5986–5997.
- Gao, Weihao, Sewoong Oh, and Pramod Viswanath (2016). “Breaking the Bandwidth Barrier: Geometrical Adaptive Entropy Estimation”. In: Curran Associates, Inc., pp. 2460–2468.
- Gasparini, Mauro (1996). “Bayesian density estimation via Dirichlet density processes”. In: *Journal of Nonparametric Statistics* 6.4, pp. 355–366.
- Gay, Dominique and Marc Boullé (2012). “A bayesian approach for classification rule mining in quantitative databases”. In: *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, pp. 243–259.
- Grünwald, Peter (2004). “A tutorial introduction to the minimum description length principle”. In: *arXiv preprint math/0406077*.
- (2007). *The minimum description length principle*. MIT press.
- Grünwald, Peter and Teemu Roos (2019). “Minimum Description Length Revisited”. In: *arXiv preprint arXiv:1908.08484*.
- Gupta, Ankit, Kishan G Mehrotra, and Chilukuri Mohan (2010). “A clustering-based discretization for supervised learning”. In: *Statistics & probability letters* 80.9-10, pp. 816–824.

- Hall, Mark, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H Witten (2009). “The WEKA data mining software: an update”. In: *ACM SIGKDD explorations newsletter* 11.1, pp. 10–18.
- Han, Songqiao, Xiyang Hu, Hailiang Huang, Minqi Jiang, and Yue Zhao (2022). “Adbench: Anomaly detection benchmark”. In: *Advances in Neural Information Processing Systems* 35, pp. 32142–32159.
- Han, Yanjun, Jiantao Jiao, and Tsachy Weissman (2015). “Adaptive estimation of shannon entropy”. In: *IEEE International Symposium on Information Theory*. IEEE, pp. 1372–1376.
- Hansen, Mark H and Bin Yu (2001). “Model selection and the principle of minimum description length”. In: *Journal of the American Statistical Association* 96.454, pp. 746–774.
- Holzinger, Andreas (2016). “Interactive machine learning for health informatics: when do we need the human-in-the-loop?” In: *Brain Informatics* 3.2, pp. 119–131.
- Hond, Anne AH de, Ilse MJ Kant, Mattia Fornasa, Giovanni Cinà, Paul WG Elbers, Patrick J Thorat, M Sesmu Arbous, and Ewout W Steyerberg (2023). “Predicting readmission or death after discharge from the ICU: external validation and retraining of a machine learning model”. In: *Critical Care Medicine* 51.2, p. 291.
- Hornik, Kurt, Christian Buchta, and Achim Zeileis (2009). “Open-Source Machine Learning: R Meets Weka”. In: *Computational Statistics* 24.2, pp. 225–232. DOI: 10.1007/s00180-008-0119-7.
- Hu, Xiyang, Cynthia Rudin, and Margo Seltzer (2019). “Optimal sparse decision trees”. In: *Advances in Neural Information Processing Systems* 32.
- Hühn, Jens and Eyke Hüllermeier (2009). “FURIA: an algorithm for unordered fuzzy rule induction”. In: *Data Mining and Knowledge Discovery* 19.3, pp. 293–319.
- Jin, Ruoming, Yuri Breitbart, and Chibuike Muoh (2009). “Data discretization unification”. In: *Knowledge and Information Systems* 19.1, pp. 1–29.
- Kahle, David and Hadley Wickham (2013). “ggmap: Spatial Visualization with ggplot2”. In: *The R Journal* 5.1, pp. 144–161. URL: <https://journal.r-project.org/archive/2013-1/kahle-wickham.pdf>.

Bibliography

- Kameya, Yoshitaka (2011). “Time Series Discretization via MDL-based Histogram Density Estimation”. In: *IEEE 23rd international conference on tools with artificial intelligence*. IEEE, pp. 732–739.
- Kang, Ye, Shanshan Wang, Xiaoyan Liu, Hokyin Lai, Huaiqing Wang, and Baiqi Miao (2006). “An ICA-based multivariate discretization algorithm”. In: *International Conference on Knowledge Science, Engineering and Management*. Springer, pp. 556–562.
- Kerber, Randy (1992). “Chimerge: Discretization of numeric attributes”. In: *Proceedings of the tenth national conference on Artificial intelligence*, pp. 123–128.
- Klemelä, Jussi (2009). “Multivariate histograms with data-dependent partitions”. In: *Statistica sinica*, pp. 159–176.
- Kontkanen, Petri and Petri Myllymäki (2007a). “A linear-time algorithm for computing the multinomial stochastic complexity”. In: *Information Processing Letters* 103.6, pp. 227–233.
- (2007b). “MDL Histogram Density Estimation”. In: *Proceedings of the Eleventh International Conference on Artificial Intelligence and Statistics*. Ed. by Marina Meila and Xiaotong Shen. Vol. 2. Proceedings of Machine Learning Research. PMLR.
- Kontkanen, Petri, Petri Myllymäki, Tomi Silander, and Henry Tirri (1997). “A Bayesian approach to discretization”. In: *Proceedings of the European symposium on intelligent techniques*.
- Kotsiantis, Sotiris and Dimitris Kanellopoulos (2006). “Discretization techniques: A recent survey”. In: *GESTS International Transactions on Computer Science and Engineering* 32.1, pp. 47–58.
- Kotsiantis, Sotiris B (2013). “Decision trees: a recent overview”. In: *Artificial Intelligence Review* 39.4, pp. 261–283.
- Kozachenko, L. F. and N. N. Leonenko (1987). “Sample Estimate of the Entropy of a Random Vector”. In: *Probl. Peredachi Inf.* 23 (2), pp. 9–16.
- Kramer, Andrew A, Thomas L Higgins, and Jack E Zimmerman (2013). “The association between ICU readmission rate and patient outcomes”. In: *Critical care medicine* 41.1, pp. 24–33.
- Kraskov, Alexander, Harald Stögbauer, and Peter Grassberger (2004). “Estimating mutual information”. In: *Physical review E* 69.6, p. 066138.

- Kurgan, Lukasz A and Krzysztof J Cios (2004). “CAIM discretization algorithm”. In: *IEEE transactions on Knowledge and Data Engineering* 16.2, pp. 145–153.
- Kwedlo, Wojciech and Marek Kretowski (1999). “An evolutionary algorithm using multivariate discretization for decision rule induction”. In: *European Conference on Principles of Data Mining and Knowledge Discovery*. Springer, pp. 392–397.
- Lakkaraju, Himabindu, Stephen H Bach, and Jure Leskovec (2016). “Interpretable decision sets: A joint framework for description and prediction”. In: *Proceedings of the 22nd ACM SIGKDD*, pp. 1675–1684.
- Lee, Jaesung and Dae-Won Kim (2013). “Feature selection for multi-label classification using multivariate mutual information”. In: *Pattern Recognition Letters* 34.3, pp. 349–357.
- Li, Bo, Peng Qi, Bo Liu, Shuai Di, Jingen Liu, Jiquan Pei, Jinfeng Yi, and Bowen Zhou (2023). “Trustworthy AI: From principles to practices”. In: *ACM Computing Surveys* 55.9, pp. 1–46.
- Li, Wenmin, Jiawei Han, and Jian Pei (2001). “CMAR: Accurate and efficient classification based on multiple class-association rules”. In: *Proceedings 2001 IEEE international conference on data mining*. IEEE, pp. 369–376.
- Liu, Bing, Wynne Hsu, Yiming Ma, et al. (1998). “Integrating classification and association rule mining.” In: *KDD*. Vol. 98, pp. 80–86.
- Liu, Linxi and Wing Hung Wong (2014). “Multivariate Density Estimation Based on Adaptive Partitioning: Convergence Rate, Variable Selection and Spatial Adaptation”. In: *arXiv preprint arXiv:1401.2597*.
- Lu, Luo, Hui Jiang, and Wing H Wong (2013). “Multivariate density estimation by Bayesian sequential partitioning”. In: *Journal of the American Statistical Association* 108.504, pp. 1402–1410.
- Lud, Marcus-Christopher and Gerhard Widmer (2000). “Relative unsupervised discretization for association rule mining”. In: *European conference on principles of data mining and knowledge discovery*. Springer, pp. 148–158.
- Mandros, Panagiotis, David Kaltenpoth, Mario Boley, and Jilles Vreeken (2020). “Discovering Functional Dependencies from Mixed-Type Data”. In: pp. 1404–1414.

Bibliography

- Marx, Alexander and Jilles Vreeken (2019). “Testing Conditional Independence on Discrete Data using Stochastic Complexity”. In: *The 22nd International Conference on Artificial Intelligence and Statistics*, pp. 496–505.
- Marx, Alexander, Lincen Yang, and Matthijs van Leeuwen (2021). “Estimating Conditional Mutual Information for Discrete-Continuous Mixtures using Multi-Dimensional Adaptive Histograms”. In: *Proceedings of the 2021 SIAM International Conference on Data Mining (SDM)*. SIAM, pp. 387–395.
- Mehta, Sameep, Srinivasan Parthasarathy, and Hui Yang (2005). “Toward unsupervised correlation preserving discretization”. In: *IEEE Transactions on Knowledge and Data Engineering* 17.9, pp. 1174–1185.
- Mesner, Octavio César and Cosma Rohilla Shalizi (2020). “Conditional Mutual Information Estimation for Mixed, Discrete and Continuous, Data”. In: *IEEE Transactions on Information Theory*.
- Molnar, Christoph (2020). *Interpretable machine learning*. Lulu.com.
- Mononen, Tommi and Petri Myllymäki (2008). “Computing the Multinomial Stochastic Complexity in Sub-Linear Time”. In: *Proceedings of the 4th European Workshop on Probabilistic Graphical Models*, pp. 209–216.
- Mosqueira-Rey, Eduardo, Elena Hernández-Pereira, David Alonso-Ríos, José Bobes-Bascarán, and Ángel Fernández-Leal (2023). “Human-in-the-loop machine learning: A state of the art”. In: *Artificial Intelligence Review* 56.4, pp. 3005–3054.
- Murdoch, W James, Chandan Singh, Karl Kumbier, Reza Abbasi-Asl, and Bin Yu (2019). “Interpretable machine learning: definitions, methods, and applications”. In: *arXiv preprint arXiv:1901.04592*.
- Nguyen, Hoang-Vu, Emmanuel Müller, Jilles Vreeken, and Klemens Böhm (2014). “Unsupervised interaction-preserving discretization of multivariate data”. In: *Data Mining and Knowledge Discovery* 28.5-6, pp. 1366–1397.
- Paninski, Liam (2003). “Estimation of entropy and mutual information”. In: *Neural Computation* 15.6, pp. 1191–1253.
- Paninski, Liam and Masanao Yajima (2008). “Undersmoothed kernel entropy estimators”. In: *IEEE Transactions on Information Theory* 54.9, pp. 4384–4388.
- Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D.

- Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay (2011). “Scikit-learn: Machine Learning in Python”. In: *Journal of Machine Learning Research* 12, pp. 2825–2830.
- Pfahringer, Bernhard (1995). “Compression-based discretization of continuous attributes”. In: *Machine Learning Proceedings 1995*. Elsevier, pp. 456–463.
- Politis, Dimitris N (1991). “On the entropy of a mixture distribution”. In: *Technical Report 91-67, Purdue University. Department of Statistics*.
- Proença, Hugo M and Matthijs van Leeuwen (2020). “Interpretable multiclass classification by MDL-based rule lists”. In: *Information Sciences* 512, pp. 1372–1393.
- Quinlan, J. Ross (1990). “Learning logical definitions from relations”. In: *Machine learning* 5, pp. 239–266.
- Quinlan, J Ross (2014). *C4. 5: programs for machine learning*. Elsevier.
- Rahimzamani, Arman, Himanshu Asnani, Pramod Viswanath, and Sreeram Kannan (2018). “Estimators for multivariate information measures in general probability spaces”. In: *Advances in Neural Information Processing Systems*, pp. 8664–8675.
- Ram, Parikshit and Alexander G Gray (2011). “Density estimation trees”. In: *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 627–635.
- Ribeiro, Marco Tulio, Sameer Singh, and Carlos Guestrin (2016). “Why should i trust you? Explaining the predictions of any classifier”. In: *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 1135–1144.
- Rissanen, Jorma (1978). “Modeling by shortest data description”. In: *Automatica* 14.5, pp. 465–471.
- (1983). “A universal prior for integers and estimation by minimum description length”. In: *The Annals of statistics* 11.2, pp. 416–431.
- (1996). “Fisher information and stochastic complexity”. In: *IEEE transactions on information theory* 42.1, pp. 40–47.
- Rudin, Cynthia (2019). “Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead”. In: *Nature Machine Intelligence* 1.5, pp. 206–215.

Bibliography

- Rudin, Cynthia, Chaofan Chen, Zhi Chen, Haiyang Huang, Lesia Semenova, and Chudi Zhong (2022). “Interpretable machine learning: Fundamental principles and 10 grand challenges”. In: *Statistic Surveys* 16, pp. 1–85.
- Rudin, Walter et al. (1964). *Principles of mathematical analysis*. Vol. 3. McGraw-hill New York.
- Schmidberger, Gabi and Eibe Frank (2005). “Unsupervised discretization using tree-based density estimation”. In: *European Conference on Principles of Data Mining and Knowledge Discovery*. Springer, pp. 240–251.
- Scott, David W (2015). *Multivariate density estimation: theory, practice, and visualization*. John Wiley & Sons.
- Scricciolo, Catia (2007). “On rates of convergence for Bayesian density estimation”. In: *Scandinavian Journal of Statistics* 34.3, pp. 626–642.
- Silander, Tomi, Janne Leppä-aho, Elias Jääsaari, and Teemu Roos (2018). “Quotient Normalized Maximum Likelihood Criterion for Learning Bayesian Network Structures”. In: *International conference on artificial intelligence and statistics*. Vol. 84. PMLR, pp. 948–957.
- Silander, Tomi, Teemu Roos, Petri Kontkanen, and Petri Myllymäki (2008). “Factorized normalized maximum likelihood criterion for learning Bayesian network structures”. In: *Proceedings of the 4th European workshop on probabilistic graphical models (PGM-08)*. Citeseer, pp. 257–272.
- Spirtes, Peter, Clark N Glymour, Richard Scheines, and David Heckerman (2000). *Causation, prediction, and search*. MIT press.
- Strobl, Eric V, Kun Zhang, and Shyam Visweswaran (2019). “Approximate kernel-based conditional independence tests for fast non-parametric causal discovery”. In: *Journal of Causal Inference* 7.1.
- Suzuki, Joe (2016). “An estimator of mutual information and its application to independence testing”. In: *Entropy* 18.4, p. 109.
- (2019). “Mutual Information Estimation: Independence Detection and Consistency”. In: *IEEE International Symposium on Information Theory (ISIT)*. IEEE, pp. 2514–2518.
- Teso, Stefano and Kristian Kersting (2019). “Explanatory interactive machine learning”. In: *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, pp. 239–245.

- Valiant, Gregory and Paul Valiant (2011). “Estimating the unseen: an $n/\log(n)$ -sample estimator for entropy and support size, shown optimal via new CLTs”. In: *International Conference on Neural Information Processing*, pp. 685–694.
- Van der Meijden, Siri (2021). “Predicting Intensive Care Unit Readmission: Performance and Explainability of Machine Learning Algorithms”. In: *Master’s thesis, Leiden University*.
- Van Der Pas, Stéphanie and Veronika Rocková (2017). “Bayesian dyadic trees and histograms for regression”. In: *arXiv preprint arXiv:1708.00078*.
- Van Leeuwen, Matthijs and Arno Knobbe (2012). “Diverse subgroup set discovery”. In: *Data Mining and Knowledge Discovery* 25.2, pp. 208–242.
- Veloso, Adriano, Wagner Meira, and Mohammed J Zaki (2006). “Lazy associative classification”. In: *Sixth International Conference on Data Mining (ICDM’06)*. IEEE, pp. 645–654.
- Vinh, Nguyen Xuan, Jeffrey Chan, and James Bailey (2014). “Reconsidering mutual information based feature selection: A statistical significance view”. In: *The AAAI Conference on Artificial Intelligence*.
- Wang, Tong, Cynthia Rudin, Finale Doshi-Velez, Yimin Liu, Erica Klampfl, and Perry MacNeille (2017). “A bayesian framework for learning rule sets for interpretable classification”. In: *The Journal of Machine Learning Research* 18.1.
- Wang, Zijie J, Alex Kale, Harsha Nori, Peter Stella, Mark E Nunnally, Duen Horng Chau, Mihaela Vorvoreanu, Jennifer Wortman Vaughan, and Rich Caruana (2022). “Interpretability, then what? editing machine learning models to reflect human knowledge and values”. In: *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 4132–4142.
- Weiler, Daniel and Julian Eggert (2007). “Multi-dimensional histogram-based image segmentation”. In: Springer, pp. 963–972.
- Woldhek, Annemarie L, Saskia Rijkenberg, Rob J Bosman, and Peter HJ Van Der Voort (2017). “Readmission of ICU patients: A quality indicator?” In: *Journal of critical care* 38, pp. 328–334.
- Wu, Mike, Michael Hughes, Sonali Parbhoo, Maurizio Zazzi, Volker Roth, and Finale Doshi-Velez (2018). “Beyond sparsity: Tree regularization of deep models for interpretability”. In: *Proceedings of the AAAI conference on artificial intelligence*. Vol. 32. 1.

Bibliography

- Yang, Fan, Kai He, Linxiao Yang, Hongxia Du, Jingbang Yang, Bo Yang, and Liang Sun (2021). “Learning Interpretable Decision Rule Sets: A Submodular Optimization Approach”. In: *Advances in Neural Information Processing Systems* 34.
- Yang, Hongyu, Cynthia Rudin, and Margo Seltzer (2017). “Scalable Bayesian rule lists”. In: *International Conference on Machine Learning*. PMLR, pp. 3921–3930.
- Yang, Kun and Wing Hung Wong (2014). “Density estimation via adaptive partition and discrepancy control”. In: *arXiv preprint arXiv:1404.1425*.
- Yang, Lincen, Mitra Baratchi, and Matthijs van Leeuwen (2023). “Unsupervised discretization by two-dimensional mdl-based histogram”. In: *Machine Learning*, pp. 1–35.
- Yin, Xiaoxin and Jiawei Han (2003). “CPAR: Classification based on predictive association rules”. In: *Proceedings of the 2003 SIAM international conference on data mining*. SIAM, pp. 331–335.
- Zhang, Guangyi and Aristides Gionis (2020). “Diverse Rule Sets”. In: *Proceedings of the 26th ACM SIGKDD*, pp. 1532–1541.
- Zhang, Xiao-Hang, Jun Wu, Ting-Jie Lu, and Yuan Jiang (2007). “A discretization algorithm based on Gini criterion”. In: *2007 International Conference on Machine Learning and Cybernetics*. Vol. 5. IEEE, pp. 2557–2561.

Acknowledgements

I would like to express my deepest gratitude to my promoter and daily supervisor Dr. Matthijs van Leeuwen. It has been an incredible journey for me in the past few years, and thank you for all the support, and for always being the most considerate supervisor ever.

I would like to thank my promoter Prof.dr.Aske Plaat. Thank you for all the support, for always giving super swift responses, and for our inspiring scientific discussions in the past four years.

My heartfelt thanks to my parents. Without your support and love, I would never have made it this far.

I would also like to specially thank Prof.dr.Peter Grunwald, for always being approachable and for always willing to help.

Further, sincere acknowledgements to Dr. Furong Ye, Dr. Fuyu Cai, Yiming Dong, Dr. Xinru Li, Kaihua Liu, Dr. Alexander Marx, and Dr. Meng Jiang, for the nice time we spent together, and the most sincere academic, career, and/or life advices you have given.

I am fortunate to have been working with a lot of smart and warm-hearted colleagues. To my EDA colleagues Daniela, Francesco, Hugo, Iris, Manon, Marieke, Saber, Sander, and Shannon, thank you for creating a “research home” for me at university. I would like to also thank Wenjing Chu, Hui Feng, Rachel de Jong, Shuaiqun Pan, Nan Pu, Minkang Tan, Yumeng Wang, Weikang Weng, Ying Yin, Yuxian Zhao, for our enjoyable lunch, coffee, and other relaxing times we have shared.

Further, special thanks to Dr. Hao Wang, Bing Xu and Zhaoxing Wang, for the best kind of friendship you have given, and for all the help and encouragement over the years. To Zhong Li, Jiayang Shi, and Zhao Yang, for the nice working,

Acknowledgements

exercising, and relaxing time we have had together.

Last, the biggest acknowledgement to my wife Chi Zhang. I feel so lucky to have you always here with me.