# Exceptional Model Mining

Duivesteijn, W.

# References

[1] T. Aidt and Z. Tzannatos, Unions and Collective Bargaining, The World Bank, 2002.

[2] P. M. Anglin, R. Gençay, Semiparametric Estimation of a Hedonic Price Function, Journal of Applied Econometrics 11 (6), pp. 633–648, 1996.

[3] K. Bache, M. Lichman, UCI Machine Learning Repository, `http://archive.ics.uci.edu/ml`, Irvine, CA, University of California, School of Information and Computer Science, 2013.

[4] S. D. Bay, M. J. Pazzani, Detecting Group Differences: Mining Contrast Sets, Data Mining and Knowledge Discovery 5 (3), pp. 213–246, 2001.

[5] H. Blockeel, L. De Raedt, J. Ramon, Top-Down Induction of Clustering Trees, Proc. ICML, pp. 55–63, 1998.

[6] M. R. Boutell, J. Luo, X. Shen, C. M. Brown, Learning Multi-Label Scene Classification, Pattern Recognition 37 (9), pp. 1757–1771, 2004.

[7] L. Breiman, J. H. Friedman, R. A. Olshen, C. J. Stone, Classification and Regression Trees, Wadsworth & Brooks/Cole Advanced Books & Software, Monterey, CA, 1984.

[8] W. L. Buntine, Theory Refinement on Bayesian Networks, Proc. UAI, pp. 52–60, 1991.

[9] L. M. de Campos, J. M. Fernández-Luna, J. F. Huete, Bayesian networks and information retrieval: an introduction to the special issue, Information Processing & Management 40 (5), pp. 727–733, 2004.

[10] C.-C. Chang, C.-J. Lin, LIBSVM: a Library for Support Vector Machines, ACM Transactions on Intelligent Systems and Technology 2 (27), pp. 1–27, 2011.

[11] W. Cheng, E. Hüllermeier, Combining Instance-Based Learning and Logistic Regression for Multilabel Classification, Machine Learning 76 (2-3), pp. 211–225, 2009.

[12] D. M. Chickering, A Transformational Characterization of Equivalent Bayesian Network Structures, Proc. UAI, pp. 87–98, 1995.

[13] R. D. Cook, Detection of Influential Observation in Linear Regression, Technometrics 19 (1), pp. 15–18, 1977.

[14] R. D. Cook, S. Weisberg, Characterizations of an Empirical Influence Function for Detecting Influential Cases in Regression, Technometrics 22 (4), pp. 495–508, 1980.

[15] R. D. Cook, S. Weisberg, Residuals and Influence in Regression, Chapman & Hall, London, 1982.

[16] M. Costanigro, R. C. Mittelhammer, J. J. McCluskey, Estimating Class-Specific Parametric Models under Class Uncertainty: Local Polynomial Regression Clustering in an Hedonic Analysis of Wine Markets, Journal of Applied Econometrics 24, pp. 1117–1135, 2009.

[17] R. G. Cowell, A. P. Dawid, S. L. Lauritzen, D. J. Spiegelhalter, Probabilistic Networks and Expert Systems, Springer-Verlag, New York, pp. 31–33, 1999.

[18] G. A. Davis, Bayesian Reconstruction of Traffic Accidents, Law, Probability and Risk 2, pp. 69–89, 2003.

[19] J. Demšar, Statistical Comparisons of Classifiers over Multiple Data Sets, Journal of Machine Learning Research 7, pp. 1–30, 2006.

[20] F. J. Díez, J. Mira, E. Iturralde, S. Zubillaga, DIAVAL, a Bayesian Expert System for Echocardiography, Artificial Intelligence in Medicine 10, pp. 59–73, 1997.

[21] G. Dong, J. Li, Efficient Mining of Emerging Patterns: Discovering Trends and Differences, Proc. KDD, pp. 43–52, 1999.

[22] C. Dougherty, Introduction to Econometrics (4th edition), Oxford University Press, Oxford, 2011.

[23] W. Duivesteijn, A. Feelders, A. Knobbe, Different Slopes for Different Folks – Mining for Exceptional Regression Models with Cook's Distance, Proc. KDD, pp. 868–876, 2012.

[24] W. Duivesteijn, A. Knobbe, Exploiting False Discoveries – Statistical Validation of Patterns and Quality Measures in Subgroup Discovery, Proc. ICDM, pp. 151–160, 2011.

[25] W. Duivesteijn, A. Knobbe, A. Feelders, M. van Leeuwen, Subgroup Discovery meets Bayesian Networks – An Exceptional Model Mining Approach, Proc. ICDM, pp. 158–167, 2010.

[26] W. Duivesteijn, E. Loza Mencía, J. Fürnkranz, A. Knobbe, Multi-label LeGo – Enhancing Multi-label Classifiers with Local Patterns, Proc. IDA, pp. 114-125, 2012.

[27] W. Duivesteijn, E. Loza Mencía, J. Fürnkranz, A. Knobbe, Multi-label LeGo – Enhancing Multi-label Classifiers with Local Patterns, Technical Report, Knowledge Engineering Group, Technische Universität Darmstadt, TUD-KE-2012-02, 2012.

[28] A. Elisseeff, J. Weston, A Kernel Method for Multi-Labelled Classification, Advances in Neural Information Processing Systems 14, pp. 681–687, MIT Press, Cambridge, MA, 2002.

[29] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, C.-J. Lin, LIBLINEAR: A Library for Large Linear Classification, Journal of Machine Learning Research 9, pp. 1871–1874, 2008.

[30] U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, From Data Mining to Knowledge Discovery in Databases, AI Magazine 17 (3), pp. 37-54, 1996.

[31] M. Friedman, The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance, Journal of the American Statistical Association 32, pp. 675–701, 1937.

[32] M. Friedman, A Comparison of Alternative Tests of Significance for the Problem of $m$ Rankings, Annals of Mathematical Statistics 11, pp. 86–92, 1940.

[33] N. Friedman, M. Linial, I. Nachman, D. Peér, Using Bayesian Networks to Analyze Expression Data, Journal of Computational Biology 7 (3-4), pp. 601–620, 2000.

[34] N. Friedman, I. Nachman, D. Peér, Learning Bayesian Network Structure from Massive Datasets: The “Sparse Candidate” Algorithm, Proc. UAI, pp. 196–205, 1999.

[35] J. Fürnkranz, P. A. Flach, ROC ‘n’ Rule Learning – Towards a Better Understanding of Covering Algorithms, Machine Learning 58 (1), pp. 39–77, 2005.

[36] J. Fürnkranz, E. Hüllermeier, E. Loza Mencía, K. Brinker, Multilabel Classification via Calibrated Label Ranking, Machine Learning 73 (2), pp. 133–153, 2008.

[37] J. Fürnkranz, A. Knobbe (eds.), Special Issue: Global Modeling Using Local Patterns, Data Mining and Knowledge Discovery journal 20 (1), 2010.

[38] E. Galbrun, P. Miettinen, From Black and White to Full Color: Extending Redescription Mining Outside the Boolean World, Statistical Analysis and Data Mining 5 (4), pp. 284–303, 2012.

[39] A. Gallo, P. Miettinen, H. Mannila, Finding Subgroups having Several Descriptions: Algorithms for Redescription Mining, Proc. SDM, pp. 334–345, 2008.

[40] G. C. Garriga, H. Heikinheimo, J. K. Seppänen, Cross-Mining Binary and Numerical Attributes, Proc. ICDM, pp. 481-486, 2007.

[41] C. F. Gauss, Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium, Friedrich Perthes and I. H. Besser, Hamburg, 1809.

[42] J. F. Gentleman, M. B. Wilk, Detecting Outliers II: Supplementing the Direct Analysis of Residuals, Biometrics 31, pp. 387–410, 1975.

[43] A. Gionis, H. Mannila, T. Mielikäinen, P. Tsarapas, Assessing Data Mining Results via Swap Randomization, Proc. KDD, pp. 167–176, 2006.

[44] S. Godbole, S. Sarawagi, Discriminative Methods for Multi-Labeled Classification, Proc. PAKDD, pp. 22–30, 2004.

[45] H. Grosskreutz, S. Rüping, On Subgroup Discovery in Numerical Domains, Data Mining and Knowledge Discovery 19 (2), pp. 210–226, 2009.

[46] H. Grosskreutz, S. Rüping, S. Wrobel, Tight Optimistic Estimates for Fast Subgroup Discovery, Proc. ECML/PKDD (1), pp. 440–456, 2008.

[47] D. Heckerman, A Tutorial on Learning with Bayesian Networks, Proc. NATO Advanced Study Institute on Learning in Graphical Models, pp. 301–354, Kluwer Academic Publishers, Norwell, MA, 1998.

[48] D. Heckerman, D. Geiger, D. M. Chickering, Learning Bayesian Networks: the Combination of Knowledge and Statistical Data, Machine Learning 20, pp. 197–243, 1995.

[49] H. Heikinheimo, M. Fortelius, J. Eronen, H. Manilla, Biogeography of European Land Mammals Shows Environmentally Distinct and Spatially Coherent Clusters, Journal of Biogeography 34 (6), pp. 1053–1064, 2007.

[50] F. Herrera, C. J. Carmona, P. González, M. J. del Jesus, An Overview on Subgroup Discovery: Foundations and Applications, Knowledge and Information Systems 29 (3), pp. 495–525, 2011.

[51] D. C. Hoaglin, R. Welsh, The Hat Matrix in Regression and ANOVA, American Statistician 32, pp. 17–22, 1978.

[52] Y. Hochberg, A. Tamhane, Multiple Comparison Procedures, Wiley, New York, 1987.

[53] R. T. Jensen, N. H. Miller, Giffen Behavior and Subsistence Consumption, American Economic Review 98 (4), pp. 1553–1577, 2008.

[54] A. M. Jorge, P. J. Azevedo, F. Pereira, Distribution Rules with Numeric Attributes of Interest, Proc. PKDD, pp. 247–258, 2006.

[55] W. Klösgen, Explora: A Multipattern and Multistrategy Discovery Assistant, Advances in Knowledge Discovery and Data Mining, pp. 249–271, 1996.

[56] W. Klösgen, Subgroup Discovery, in: W. Klösgen, J.M. Zytkow (eds.), Handbook of Data Mining and Knowledge Discovery, pp. 354–361, Oxford University Press, Oxford, 2002.

[57] A. Knobbe, B. Crémilleux, J. Fürnkranz, M. Scholz, From Local Patterns to Global Models: the LeGo Approach to Data Mining, Proc. ECML/PKDD Workshop: From Local Patterns to Global Models, pp. 1–16, 2008.

[58] A. Knobbe, E. Ho, Pattern Teams, Proc. PKDD, pp. 577–584, 2006.

[59] A. Knobbe, J. Valkonet, Building Classifiers from Pattern Teams, Proc. ECML PKDD Workshop: From Local Patterns to Global Models, pp. 77–93, 2009.

[60] D. E. Knuth, The Art of Computer Programming, Volume 3: Sorting and Searching, second edition, Addison–Wesley, Reading, MA, 1998.

[61] D. Kocev, C. Vens, J. Struyf, S. Džeroski, Tree Ensembles for Predicting Structured Outputs, Pattern Recognition 46 (3), pp. 817–833, 2013.

[62] R. Kohavi, The Power of Decision Tables, Proc. ECML, pp. 174–189, 1995.

[63] R. M. Konijn, W. Kowalczyk, Hunting for Fraudsters in Random Forests, Proc. HAIS, pp. 174–185, 2012.

[64] E. van de Koppel, I. Slavkov, K. Astrahantseff, A. Schramm, J. Schulte, J. Vandesompele, E. de Jong, S. Dzeroski, A. Knobbe, Knowledge Discovery in Neuroblastoma-related Biological Data, Proc. PKDD Workshop: Data Mining in Functional Genomics and Proteomics, pp. 45-56, 2007.

[65] P. Kralj Novak, N. Lavrač, G. I. Webb, Supervised Descriptive Rule Discovery: A Unifying Survey of Contrast Set, Emerging Pattern and Subgroup Mining, Journal of Machine Learning Research 10, pp. 377–403, 2009.

[66] H.-P. Kriegel, P. Kröger, E. Schubert, A. Zimek, Outlier Detection in Arbitrarily Oriented Subspaces, Proc. ICDM, pp. 379–388, 2012.

[67] P. Larrañaga, M. Poza, Y. Yurramendi, R. H. Murga, C. M. H. Kuijpers, Structure Learning of Bayesian Networks by Genetic Algorithms: A Performance Analysis of Control Parameters, IEEE Transactions on Pattern Analysis and Machine Intelligence 18 (9), pp. 912–926, 1996.

[68] N. Lavrač, B. Kavšek, P. Flach, L. Todorovski, Subgroup Discovery with CN2-SD, Journal of Machine Learning Research 5, pp. 153–188, 2004.

[69] M. van Leeuwen, Maximal Exceptions with Minimal Descriptions, Data Mining and Knowledge Discovery 21 (2), pp. 259–276, 2010.

[70] M. van Leeuwen, A. J. Knobbe, Non-redundant Subgroup Discovery in Large and Complex Data, Proc. ECML PKDD (3), pp. 459–474, 2011.

[71] D. Leman, A. Feelders, A. Knobbe, Exceptional Model Mining, Proc. ECML/PKDD (2), pp. 1–16, 2008.

[72] F. Lemmerich, M. Becker, M. Atzmüller, Generic Pattern Trees for Exhaustive Exceptional Model Mining, Proc. ECML-PKDD (2), pp. 277–292, 2012.

[73] E. Loza Mencía, Efficient Pairwise Multilabel Classification, PhD thesis, Technische Universität Darmstadt, 2012.

[74] E. Loza Mencía, J. Fürnkranz, Efficient Pairwise Multilabel Classification for Large-Scale Problems in the Legal Domain, Proc. ECML/PKDD (2), pp. 50–65, 2008.

[75] A. M. Lyapunov, Nouvelle Forme du Théorème sur la Limite de Probabilité, Mémoires de l'Académie Impériale des Sciences de St. Petersburg 12, pp. 1–24, 1901.

[76] M. Mampaey, S. Nijssen, A. Feelders, A. J. Knobbe, Efficient Algorithms for Finding Richer Subgroup Descriptions in Numeric and Nominal Data, Proc. ICDM, pp. 499–508, 2012.

[77] A. Marshall, Principles of Economics, MacMillan and Co., London, 1895.

[78] M. Meeng, A. J. Knobbe, Flexible Enrichment with Cortana – Software Demo, Proc. Benelearn, pp. 117–119, 2011.

[79] N. Megiddo, R. Srikant, Discovering Predictive Association Rules, Proc. KDD, pp. 274–278, 1998.

[80] A. J. Mitchell-Jones, W. Bogdanowicz, B. Krystufek, P. J. H. Reijnders, F. Spitzenberger, C. Stubbe, J. B. M. Thissen, V. Vohralík, J. Zima, The Atlas of European Mammals, Poyser Natural History, Academic Press, London, 1999.

[81] D. Moore, G. McCabe, Introduction to the Practice of Statistics, Freeman, New York, 1993.

[82] M. Neil, N. Fenton, M. Tailor, Using Bayesian Networks to Model Expected and Unexpected Operational Losses, Risk Analysis 25 (4), 2005.

[83] P. B. Nemenyi, Distribution-Free Multiple Comparisons, PhD thesis, Princeton University, 1963.

[84] J. Neter, M. Kutner, C. J. Nachtsheim, W. Wasserman, Applied Linear Statistical Models, WCB McGraw-Hill, New York, 1996.

[85] M. Ojala, G. C. Garriga, A. Gionis, H. Mannila, Evaluating Query Result Significance in Databases via Randomizations, Proc. SDM, pp. 906–917, 2010.

[86] R. T. Paine, Food Web Complexity and Species Diversity, The American Naturalist 100 (910), pp. 65–75, 1966.

[87] S.-H. Park, J. Fürnkranz, Multi-Label Classification with Label Constraints, Proc. ECML/PKDD Workshop: Preference Learning, pp. 157–171, 2008.

[88] K. Pearson, L. Filon, Mathematical Contributions to the Theory of Evolution, iv. on the Probable Errors of Frequency Constants and on the Influence of Random Selection on Variation and Correlation, Philosophical Transactions of the Royal Society of London, Series A, Containing Papers of a Mathematical or Physical Character, 191, pp. 229–311, 1898.

[89] B. F. I. Pieters, A. Knobbe, S. Džeroski, Subgroup Discovery in Ranked Data, with an Application to Gene Set Enrichment, Proc. ECML PKDD Workshop: Preference Learning, 2010.

[90] J. R. Quinlan, Learning with Continuous Classes, Proc. AJCAI, pp. 343–348, 1992.

[91] N. Ramakrishnan, D. Kumar, B. Mishra, M. Potts, R. F. Helm, Turning CARTwheels: an Alternating Algorithm for Mining Redescriptions, Proc. KDD, pp. 837–844, 2004.

[92] J. Read, B. Pfahringer, G. Holmes, E. Frank, Classifier Chains for Multi-label Classification, Proc. ECML PKDD, pp. 254–269, 2009.

[93] L. Rezende, Econometrics of Auctions by Least Squares, Journal of Applied Econometrics 23, pp. 925–948, 2008.

[94] J. A. Rice, Mathematical Statistics and Data Analysis, second edition, Duxbury Press, Wadsworth Publishing Company, Belmont, CA, 1995.

[95] C. Riggelsen, Approximation Methods for Efficient Learning of Bayesian Networks, IOS Press, Amsterdam, 2008.

[96] E. Schubert, J. Wolfe, A. Tarnopolsky, Spectral Centroid and Timbre in Complex, Multiple Instrumental Textures, Proc. 8th International Conference on Music Perception & Cognition, pp. 654–657, 2004.

[97] K. Sechidis, G. Tsoumakas, I. P. Vlahavas, On the Stratification of Multi-label Data, Proc. ECML PKDD (3), pp. 145–158, 2011.

[98] C. Silverstein, S. Brin, R. Motwani, Beyond Market Baskets: Generalizing Association Rules to Dependence Rules, Data Mining and Knowledge Discovery 2 (1), pp. 39–68, 1998.

[99] T. Stengos, E. Zacharias, Intertemporal Pricing and Price Discrimination: A Semiparametric Hedonic Analysis of the Personal Computer Market, Journal of Applied Econometrics 21, pp. 371–386, 2006.

[100] J.-N. Sulzmann, J. Fürnkranz, A Comparison of Techniques for Selecting and Combining Class Association Rules, Proc. ECML/PKDD Workshop: From Local Patterns to Global Models, pp. 154-168, 2008.

[101] P.-N. Tan, V. Kumar, J. Srivastava, Selecting the right interestingness measure for association patterns, Proc. KDD, pp. 32–41, 2002.

[102] A. Tellegen, D. Watson, L. A. Clark, On the Dimensional and Hierarchical Structure of Affect, Psychological Science 10 (4), pp. 297–303, 1999.

[103] K. Trohidis, G. Tsoumakas, G. Kalliris, I. P. Vlahavas, Multi-Label Classification of Music into Emotions, Proc. 9th International Conference on Music Information Retrieval, pp. 325–330, 2008.

[104] G. Tsoumakas, A. Dimou, E. Spyromitros Xioufis, V. Mezaris, I. Kompatsiaris, I. Vlahavas, Correlation Based Pruning of Stacked Binary Relevance Models for Multi-Label Learning, Proc. 1st International Workshop on Learning from Multi-Label Data, pp. 101–116, 2009.

[105] G. Tsoumakas, E. Loza Mencía, I. Katakis, S.-H. Park, J. Fürnkranz, On the Combination of Two Decompositive Multi-Label Classification Methods, Proc. ECML PKDD Workshop: Preference Learning, pp. 114–129, 2009.

[106] G. Tsoumakas, I. Katakis, Multi-Label Classification: An Overview, International Journal of Data Warehousing and Mining 3 (3), pp. 1–13, 2007.

[107] G. Tsoumakas, I. Katakis, I. P. Vlahavas, Mining Multi-label Data, Data Mining and Knowledge Discovery Handbook, Springer, New York, pp. 667–685, 2010.

[108] G. Tsoumakas, J. Vilcek, E. Spyromitros Xioufis, I. P. Vlahavas, Mulan: A Java Library for Multi-Label Learning, Journal of Machine Learning Research 12, pp. 2411–2414, 2011.

[109] L. Umek, B. Zupan, Subgroup Discovery in Data Sets with Multi-Dimensional Responses, Intelligent Data Analysis 15 (4), pp. 533–549, 2011.

[110] A. Veloso, W. Meira Jr., M. A. Gonçalves, M. J. Zaki, Multi-label Lazy Associative Classification, Proc. PKDD, pp. 605–612, 2007.

[111] T. Verma, J. Pearl, Equivalence and Synthesis of Causal Models, Proc. UAI, pp. 255–270, 1990.

[112] G. I. Webb, Discovering Significant Patterns, Machine Learning 68 (1), pp. 1–33, 2007.

[113] I. H. Witten, E. Frank, M. A. Hall, Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann, San Francisco, CA, 2011.

[114] S. Wrobel, An Algorithm for Multi-relational Discovery of Subgroups, Proc. PKDD, pp. 78–87, 1997.

[115] G. Yang, L. Le Cam, Asymptotics in Statistics: Some Basic Concepts, Berlin, Springer-Verlag, 2000.

[116] B. Zhang, Regression Clustering, Proc. ICDM, pp. 451–458, 2003.

[117] M.-L. Zhang, Lift: Multi-Label Learning with Label-Specific Features, Proc. 22nd International Joint Conference on Artificial Intelligence, pp. 1609–1614, 2011.

[118] M.-L. Zhang, K. Zhang, Multi-Label Learning by Exploiting Label Dependency, Proc. KDD 2010, pp. 999–1008, 2010.