Methods and tools for mining multivariate time series

De Gouveia da Costa Cachucho, R.E.

Author: de Gouveia da Costa Cachucho, R.E.
Title: Methods and tools for mining multivariate time series
Issue Date: 2018-12-10

Chapter 7

Conclusions

In this thesis, we addressed data mining methods, tools and applications for multivariate time series. The sequential nature of time series imposes specific algorithmic solutions to address this type of data. Additionally, this thesis narrows the focus to *multivariate* time series. The multivariate aspect of the data reflects the complexity and multi-faceted nature of the system under investigation. Whether one is measuring infrastructures, athletic performance, monitoring human activities or analysing life style, there are numerous aspects that can be measured. This increasing growth of data frequency and sources stimulate this data centric era. But with new opportunities also come several challenges from the perspective of the methods and tools used.

In this chapter, we reflect on our main contributions to meet some of the knowledge discovery challenges surrounding multivariate time series. We can separate our contributions into two sides: unsupervised and supervised learning. Next to these two paradigms, we focus on a group of three aspects of data science: methods, tools and applications, which for lack of a better term, we refer to as the data science *triad*. Aside from these contributions, we take the opportunity to reflect on two research directions in this era of data science. Firstly, machine learning as an optimization process in two directions: better data representation and better model learning algorithms. Secondly, data science as a paradigm shift of scientific process: scarcity of data versus big data.

Unsupervised Multivariate Time Series In Chapter 2, we introduced a data mining method to find biclusters in multivariate time series, which addresses research questions **Q1** and **Q2**. This task addresses the situation of a system that has some recurring patterns over time in a subset of variables. In contrast with traditional biclustering algorithms, our method is able to find significant periods of time (larger sequences) where multiple variables deviate in a coherent manner. By coherent in this context, we mean that for each selected variable all segments have similar probability distributions.

We argue that framing this pattern recognition challenge as a biclustering task offers considerable benefits. Firstly, pattern recognition tasks in time series have been focusing primarily on the univariate scenario and this neglects relationships between variables. Take as an example the monetary exchange market. We know that the exchange rate of different currencies are connected, but which currencies are strongly related to the Brazilian real and which are related to the Singapore dollar? Furthermore, are these relations always present, or are they triggered by specific events? How long do these relationships last? The same analogy holds for almost any system measured over time. We claim that biclustering multivariate time series can playing an important role in finding such patterns.

Supervised Multivariate Time Series In order to tackle research questions **Q3** and **Q4**, in Chapter 4 we introduced Accordion, a greedy search algorithm that produces good aggregate features, both for regression and classification. With such features, one has a better data representation and this leads to better models. In the supervised learning setting, Accordion is a wrapper algorithm that integrates the feature construction and selection into the learning process. Our method differs from the common practice of considering feature construction as an isolated pre-processing step. We demonstrated the positive effects of searching for good aggregate features automatically by optimizing the selection of three components: time series variable, aggregate function and window size.

Our method automates a feature construction and selection processes, combining multivariate time series with mixed sampling rates. Normally such optimization processes come at the expense of producing features which are not interpretable. We decided to focus on the automatic construction and selection of interpretable aggregate features. By interpretable, we mean the combinations of input variables, aggregate functions and different window sizes. Our optimization process, for each combination of variable and aggregation function, expands and contracts the window size, capturing phenomena that work on different time scales. We believe this method to be highly promising for further development and implementation in efficient computation architectures, and better exploration of functional properties of

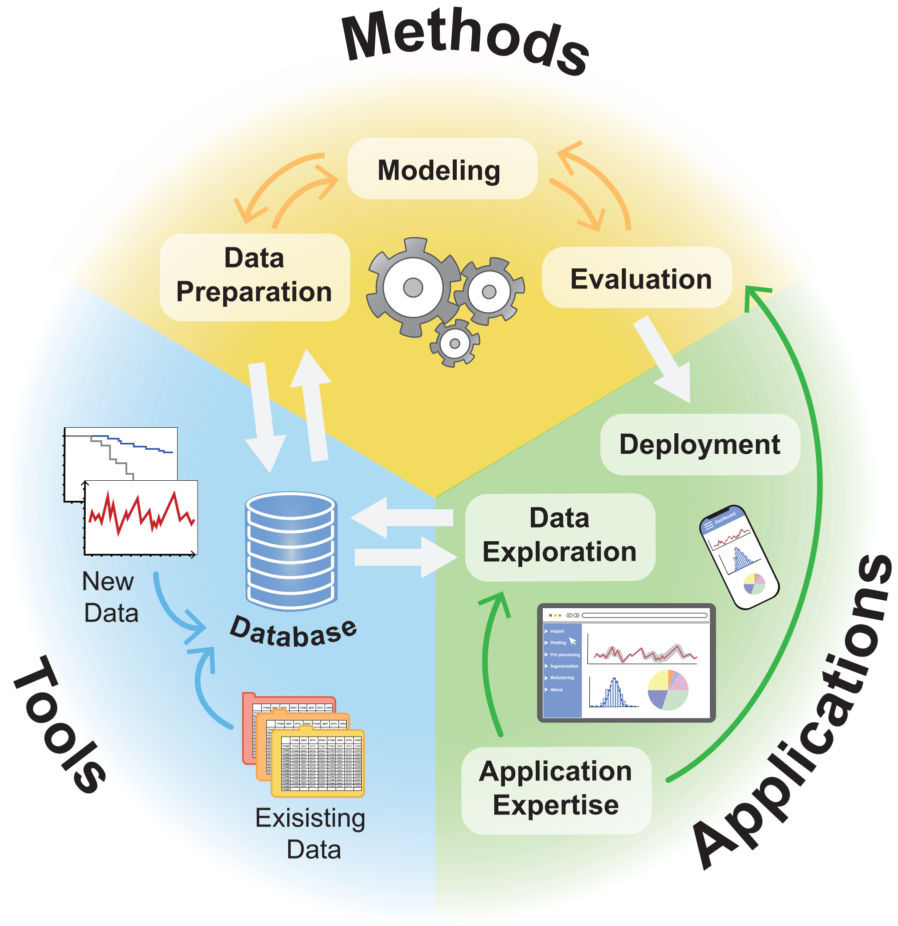

Figure 7.1: The triad of data science projects.

different aggregations. This will allow us to get faster and better results in larger search spaces.

Data Science Triad In addition to the machine learning methods introduced in Chapters 2 and 4, we observe that such methods stand at the intersection between tools and applications. In fact, data science can be seen as the integration of methods, tools and applications. Thus, relevant data science is the implementation of relevant methods, tailored to a specific application, while using the appropriate tools.

We illustrate the data science triad in Figure 7.1, by representing a generic scheme of what a data science project normally entails. Unlike other fields such as data mining, data science is finding an equilibrium between designing generalized machine learning methods and focusing on specific applications with efficient tools. Examples of such relationships can also be found in this thesis. For each of the methods (Chapters 2 and 4) there are corresponding easy to access and intuitive tools in Chapters 3 and 5, respectively. Such tools address research question **Q6**. Additionally, Chapter 6 focuses on a specific application of analyzing and improving the performance of elite speed skating athletes (research question **Q5**). This is done by using tailored features and models for each athlete and a relational database designed for elite sports performance monitoring.

Machine Learning: Optimizing Two Complementary Directions The machine learning process is typically characterized as by being a process of exploration and exploitation of the data. Thus, names for the field such as data mining become intuitive to understand. Maybe the best explanation for this tendency has to do with the uncountable data sources and numerous data structures that exist. Here one could make an analogy with the proverb of Muhammad and the mountain, the machine learning algorithms being Muhammad and the mountain being the data (big data). It makes sense to see the challenge of creating an algorithm that is able to deal with different mountains, an algorithm that is adaptable to different datasets, reliable in its findings and fast in the construction of a new model.

An alternative to the view above is to have an algorithm that is able to describe properly the data. The process of improving a model outcome can be solved with good representations of the decision space. This space is described with the input data or transformations of it. These variable transformation we normally refer to as feature construction, extraction and selection. In the light of Muhammad and the mountain, instead of the mountain itself, maybe the model can be improved by knowing the height of the mountain, the soil properties, a vast photo gallery of the mountain from different perspectives and exposure. If only one could define an algorithm that searches for such descriptions automatically, maybe linear models could solve the majority of modeling challenges.

The experimental results in Chapter 4 indicate that there are significant gains to be had from focusing on the data representation. In the case of this thesis, the focus was on improving decision trees and linear regression models. Probably many more well-known algorithms can benefit from a layer of data transformation to reach the appropriate representation of the input space.

In fact, some years back, it would have been difficult to motivate such an algorithm for a task that is commonly seen as a pre-processing task. Presently, with developments in fields such as convolutional neural networks, motivating automatic data transformation has become more acceptable. As it seems with the advances in convolutional neural networks, these two optimization directions are actually perfectly complementary.

Data Science as a Scientific Paradigm Shift Data science being a discipline at the intersection of multiple other disciplines, it brings together multiple empirical sciences, which depend on observations to draw founded conclusions. In order to avoid drawing incorrect conclusions from observations, traditionally *hypothesis testing* is put forward as an essential, if not the only, scientific paradigm. With the current wealth of data, the field of data science is put at the center of a great scientific discussion. Is hypothesis testing presently still a sufficient scientific paradigm for research?

Hypothesis testing has been at the core of empirical sciences. This importance is due to two main reasons: data was scarce and costly to gather, and starting from a fixed, prior hypothesis is traditionally how to determine causality. As a start, in this period of data abundance, we are often analyzing the whole population instead of only a tiny sample. Concepts such as big data could trigger a deeper discussion about our present capacity for inference. Secondly, hypothesis testing is still considered the golden standard to determine causality. But many new developments in data science, for example *causal inference* [76] are demonstrating that there are in fact ways to avoid the *post hoc ergo propter hoc* fallacies. This makes the analysis of observational data, the predominant setting in data science, a very acceptable approach for scientific discovery.

So how could a research project be designed according to this new scientific paradigm? Take the speed skating example of Chapter 6. The central question there is how to increase the performance of elite athletes and win more medals. Following the data science rationale, we need to decide on what all can be monitored that can conceivably influence the performance of the athlete. One could measure training sessions, daily state of mind, eating logs, sleeping habits, metabolic variables, you name it. What and how to measure is no longer governed by preconceived ideas about what is expected to be the principle driver of athletic performance (the 'hypothesis'), but rather by technological, ethical and pragmatic considerations. If it can be practically, ethically and realistically measured, let's just include it in the analysis. After data collection, machine learning will consider numerous promising hypotheses in an exploratory manner, followed by a host of robust evaluation methods and experts to validate any scientific discoveries.

Bibliography

[1] https://www.universiteitleiden.nl/nieuws/2016/10/leidse-onderzoekers-winnen-publieksprijs-slimste-project-van-nederland.

[2] S. Aghabozorgi, A. S. Shirkhorshidi, and T. Y. Wah. Time-series clustering — a decade review. *Information Systems*, 53:16–38, 2015.

[3] E. Angelini, J. Henry, and M. Marcellino. Interpolation and backdating with a large information set. *Journal of Economic Dynamics and Control*, 30(12):2693–2724, Dec. 2006.

[4] M. Armesto. Forecasting with mixed frequencies. *Federal Reserve Bank of St. Louis*, pages 521–536, 2010.

[5] M. Atzmüller and F. Lemmerich. Fast subgroup discovery for continuous target concepts. In *ISMIS, the International Symposium on Methodologies for Intelligent Systems*, pages 35–44, 2009.

[6] L. Bao and S. Intille. Activity recognition from user-annotated acceleration data. *Pervasive Computing*, pages 1–17, 2004.

[7] S. Barkow, S. Bleuler, A. Prelić, P. Zimmermann, and E. Zitzler. Bicat: A biclustering analysis toolbox. *Bioinformatics*, 22(10):1282–1283, May 2006.

[8] M. R. Berthold, N. Cebron, F. Dill, T. R. Gabriel, T. Kötter, T. Meinl, P. Ohl, K. Thiel, and B. Wiswedel. Knime - the konstanz information miner: Version 2.0 and beyond. *ACM SIGKDD Explorations Newsletter*, 11(1):26–31, Nov. 2009.

[9] A. Bifet. *Adaptive Learning and Mining for Data Streams and Frequent Patterns*. PhD thesis, Universitat Politècnica de Catalunya, 2009.

[10] A. Bifet, G. Holmes, R. Kirkby, and B. Pfahringer. Moa: Massive online analysis. *Journal of Machine Learning Research*, 11(May):1601–1604, 2010.

[11] W. Bouten, E. W. Baaij, J. Shamoun-Baranes, and K. C. Camphuysen. A flexible gps tracking system for studying bird behaviour at multiple scales. *Journal of Ornithology*, 154(2):571–580, 2013.

[12] A. Brush, J. Krumm, J. Scott, and T. Saponas. Recognizing Activities from Mobile Sensor Data: Challenges and Opportunities. In *Proceedings Ubicomp' 11*, 2011.

[13] R. Cachucho, K. Liu, S. Nijssen, and A. Knobbe. Bipeline: A web-based visualization tool for biclustering of multivariate time series. In B. Berendt, B. Bringmann, É. Fromont, G. Garriga, P. Miettinen, N. Tatti, and V. Tresp, editors, *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2016, Riva del Garda, Italy, September 19-23, 2016, Proceedings, Part III*, pages 12–16, Cham, 2016. Springer International Publishing.

[14] R. Cachucho, M. Meeng, U. Vespier, S. Nijssen, and A. Knobbe. Mining multivariate time series with mixed sampling rates. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '14, pages 413–423, New York, NY, USA, 2014. ACM.

[15] R. Cachucho, S. Nijssen, and A. Knobbe. Biclustering multivariate time series. In *Intelligent Data Analysis*, 2017.

[16] T. Calvert, E. Banister, M. Savage, and T. Bach. A systems model of the effects of training on physical performance. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-6(2):94–102, 1976.

[17] W. Chang, J. Cheng, J. Allaire, Y. Xie, and J. McPherson. *shiny: Web Application Framework for R*, 2016. R package version 0.13.1.

[18] W. Chang, J. Cheng, J. Allaire, Y. Xie, and J. McPherson. shiny: Web application framework for r [computer software]. *URL http://CRAN. R-project. org/package= shiny (R package version 1.0. 0)*, 2017.

[19] Y. Cheng and G. M. Church. Biclustering of expression data. In *Proceedings of the Eighth International Conference on Intelligent Systems for Molecular Biology*, pages 93–103. AAAI Press, 2000.

[20] B. Chiu, E. Keogh, and S. Lonardi. Probabilistic discovery of time series motifs. In *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '03, pages 493–498, New York, NY, USA, 2003. ACM.

[21] G. Creasey, P. Jarvis, and L. Berk. Play and social competence. In O. N. Saracho and B. Spodek, editors, *Early Childhood Education: Inquiries and Insights. Multiple Perspectives on Play in Early Childhood Education*. State University of New York Press, New York, 1998.

[22] M. de Winter. *A Missing Value Ignoring Approach for Whole Time Series Clustering of Longevity Corebody Temperature Data*. Master's thesis, Leiden University, Leiden Institute of Advance Computer Science, 2015.

[23] W. Duivesteijn and A. Knobbe. Exploiting false discoveries – statistical validation of patterns and quality measures in subgroup discovery. In *ICDM, the International Conference on Data Mining*, pages 151–160, 2011.

[24] P. Esling and C. Agon. Time-series data mining. *ACM Comput. Surv.*, 45(1):12:1–12:34, Dec. 2012.

[25] D. Figo, P. C. Diniz, D. R. Ferreira, and J. a. M. P. Cardoso. Preprocessing techniques for context recognition from accelerometer data. *Personal and Ubiquitous Computing*, 14(7):645–662, Mar. 2010.

[26] J. Friedman, T. Hastie, and R. Tibshirani. Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 2010.

[27] J. Friedman, T. Hastie, and R. Tibshirani. Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1):1–22, 2010.

[28] E. Ghysels, P. Santa-Clara, and R. Valkanov. Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics*, 2006.

[29] J. P. Gonçalves, S. C. Madeira, and A. L. Oliveira. Biggests: integrated environment for biclustering analysis of time series gene expression data. *BMC Research Notes*, 2(1):124, 2009.

[30] H. Grosskreutz and S. Rüping. On subgroup discovery in numerical domains. *Data Mining and Knowledge Discovery*, 19(2):210–226, 2009.

[31] D. Gujarati. *Basic econometrics*. McGraw Hill, fourth edition, 2003.

[32] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten. The weka data mining software: An update. *ACM SIGKDD Explorations Newsletter*, 11(1):10–18, Nov. 2009.

[33] T. Hastie, R. Tibshirany, and J. Friedman. *The Elements of Statistical Learning Data Mining, Inference, and Prediction*. Springer, second edition, 2009.

[34] J. Heinrich, R. Seifert, M. Burch, and D. Weiskopf. Bicluster viewer: a visualization tool for analyzing gene expression data. In *International Symposium on Visual Computing*, pages 641–652. Springer, 2011.

[35] J. Himberg, K. Korpiaho, H. Mannila, J. Tikanmaki, and H. T. Toivonen. Time series segmentation for context recognition in mobile devices. In *Data Mining, 2001. ICDM 2001, Proceedings IEEE International Conference on*, pages 203–210. IEEE, 2001.

[36] J. A. Hobson. Rem sleep and dreaming: towards a theory of protoconsciousness. *Nature Reviews Neuroscience*, 10(11):803–813, 2009.

[37] T. Huynh and B. Schiele. Analyzing features for activity recognition. In *Proceedings of the 2005 joint conference on Smart objects and ambient intelligence: innovative context-aware services: usages and technologies*, pages 159–163. ACM, 2005.

[38] IBM Corporation. *IBM SPSS Statistics*.

[39] G. John, R. Kohavi, and K. Pfleger. Irrelevant features and the subset selection problem. In *Machine Learning*, pages 121–129, 1994.

[40] S. Jongstra, L. W. Wijsman, R. Cachucho, M. P. Hoevenaar-Blom, S. P. Mooijaart, and E. Richard. Cognitive testing in people at increased risk of dementia using a smartphone app: The ivitality proof-of-principle study. *JMIR mHealth and uHealth*, 5(5):e68, 2017.

[41] A. Jorge, P. Azevedo, and F. Pereira. Distribution rules with numeric attributes of interest. In *PKDD, the European Conference on Principles and Practice of Knowledge Discovery in Databases*, pages 247–258, 2006.

[42] S. Kaiser and F. Leisch. A toolbox for bicluster analysis in r. In *UseR*, 2008.

[43] S. Kaiser and F. Leisch. A generalized motif bicluster algorithm. In *UseR*, 2009.

[44] S. Kaiser, R. Santamaria, T. Khamiakova, M. Sill, R. Theron, L. Quintales, F. Leisch, and E. D. Troyer. *biclust: BiCluster Algorithms*, 2015. R package version 1.2.0.

[45] W. Kenney, J. Wilmore, and D. Costill. *Physiology of Sport and Exercise*. Human Kinetics, 2015.

[46] E. Keogh, S. Chu, D. Hart, and M. Pazzani. An online algorithm for segmenting time series. In *Data Mining, 2001. ICDM 2001, Proceedings IEEE International Conference on*, pages 289–296. IEEE, 2001.

[47] E. Keogh, S. Chu, D. Hart, and M. Pazzani. Segmenting time series: A survey and novel approach. *Data mining in time series databases*, 57:1–22, 2004.

[48] E. Keogh and S. Kasetty. On the need for time series data mining benchmarks: a survey and empirical demonstration. *Data Mining and Knowledge Discovery*, pages 349–371, 2002.

[49] E. Keogh and J. Lin. Clustering of time-series subsequences is meaningless: implications for previous and future research. *Knowledge and information systems*, 8(2):154–177, 2005.

[50] W. Klösgen. *Subgroup Discovery, Handbook of Data Mining and Knowledge Discovery*, chapter 16.3. Oxford University Press, 2002.

[51] Y. Kluger, R. Basri, J. T. Chang, and M. Gerstein. Spectral biclustering of microarray cancer data: Co-clustering genes and conditions. *Genome Research*, 13:703–716, 2003.

[52] A. Knobbe, H. Blockeel, and A. Koopman. InfraWatch: Data management of large systems for monitoring infrastructural performance. In *Intelligent Data Analysis*, pages 91–102, 2010.

[53] A. Knobbe, J. Orie, N. Hofman, B. van der Burgh, and R. Cachucho. Sports analytics for professional speed skating. *Data Mining and Knowledge Discovery*, 31(6):1872–1902, Nov 2017.

[54] P. Kralj Novak, N. Lavrač, and G. Webb. Supervised descriptive rule discovery: A unifying survey of contrast set, emerging pattern and subgroup mining. *Journal of Machine Learning Research*, 10:377–403, 2009.

[55] M. Krätzig. The software jmulti. *Applied Time Series Econometrics, Cambridge University Press, Cambridge*, pages 289–299, 2004.

[56] M. Kuhn. Caret package. *Journal of Statistical Software*, 28(5):1–26, 2008.

[57] V. Kuzin, M. Marcellino, and C. Schumacher. MIDAS vs. mixed-

frequency VAR: Nowcasting GDP in the euro area. *International Journal of Forecasting*, 27(2):529–542, Apr. 2011.

[58] J. Kwapisz, G. Weiss, and S. Moore. Activity recognition using cell phone accelerometers. *ACM SIGKDD Explorations Newsletter*, 2011.

[59] F. Lemmerich, M. Atzmueller, and F. Puppe. Fast exhaustive subgroup discovery with numerical target concepts. *Data Mining and Knowledge Discovery*, pages 1–52, 2015.

[60] J. Lin, E. Keogh, S. Lonardi, and B. Chiu. A symbolic representation of time series, with implications for streaming algorithms. *Proceedings of the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery - DMKD '03*, page 2, 2003.

[61] S. C. Madeira and A. L. Oliveira. Biclustering algorithms for biological data analysis: A survey. *IEEE/ACM Trans. Comput. Biol. Bioinformatics*, 1(1):24–45, Jan. 2004.

[62] S. C. Madeira and A. L. Oliveira. A linear time biclustering algorithm for time series gene expression data. In R. Casadio and G. Myers, editors, *Algorithms in Bioinformatics: 5th International Workshop, WABI 2005, Mallorca, Spain, October 3-6, 2005. Proceedings*, pages 39–52, Berlin, Heidelberg, 2005. Springer Berlin Heidelberg.

[63] S. Makridakis, S. C. Wheelwright, and R. J. Hyndman. *Forecasting methods and applications*. John Wiley & Sons, 1998.

[64] W. McArdle, F. Katch, and V. Katch. *Exercise Physiology: nutrition, energy, and human performance*. Lippincott, Williams & Wilkins, 2014.

[65] M. Meeng, R. Cachucho, and A. Knobbe. Quickie: Quick user interface for convolution kernel-involving experiments. `http://liacs.leidenuniv.nl/~meengm/doc/index.html`. Accessed: 2017-12-07.

[66] M. Meeng and A. Knobbe. Flexible enrichment with Cortana – software demo. In *BeneLearn, the Annual Belgian-Dutch Conference on Machine Learning*, pages 117–119, 2011.

[67] S. Miao, U. Vespier, R. Cachucho, M. Meeng, and A. Knobbe. Predefined pattern detection in large time series. *Information Sciences*, 329:950–964, 2016.

[68] S. Miao, U. Vespier, J. Vanschoren, A. Knobbe, and R. Cachucho. Modeling sensor dependencies between multiple sensor types. In *Proceedings of BeneLearn*, 2013.

[69] E. Miluzzo, N. Lane, and K. Fodor. Sensing meets mobile social networks: the design, implementation and evaluation of the CenceMe application. In *Proceedings of Embedded network sensor systems*, pages 337–350, 2008.

[70] D. Minnen, C. Isbell, I. Essa, and T. Starner. Detecting subdimensional motifs: An efficient algorithm for generalized multivariate pattern discovery. In *Proceedings of the 2007 Seventh IEEE International Conference on Data Mining*, ICDM '07, pages 601–606, Washington, DC, USA, 2007. IEEE Computer Society.

[71] A. Mueen, E. Keogh, Q. Zhu, S. Cash, and B. Westover. Exact discovery of time series motifs. In *Proceedings of the 2009 SIAM International Conference on Data Mining*, pages 473–484, 2009.

[72] T. M. Murali and S. Kasif. Extracting conserved gene expression motifs from gene expression data. In *Pac. Symp. Biocomput*, pages 77–88, 2003.

[73] J. Orie, N. Hofman, J. de Koning, and C. Foster. Thirty-eight years of training distribution in olympic speed skaters. *Int J Sports Physiol Perform*, 9:93–9, 2014.

[74] S. Paraschiakos. *Comparing Sensor Networks for Activity Recognition.* Master's thesis, Leiden University, Leiden Institute of Advance Computer Science, 2017.

[75] J.-g. Park, A. Patel, D. Curtis, S. Teller, and J. Ledlie. Online pose classification and walking speed estimation using handheld devices. *Proceedings of the 2012 ACM Conference on Ubiquitous Computing - UbiComp '12*, page 10, 2012.

[76] J. Pearl. *Causality: Models, Reasoning and Inference.* Cambridge University Press, 2nd edition edition, 2009.

[77] F. Pedregosa, 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. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[78] B. Pieters, A. Knobbe, and S. Džeroski. Subgroup discovery in ranked data, with an application to gene set enrichment. In *Preference Learning 2010 at ECML PKDD, the European Conference on Ma-*

chine Learning and Principles and Practice of Knowledge Discovery in Databases, 2010.

[79] A. Prelić, S. Bleuler, P. Zimmermann, A. Wille, P. Bühlmann, W. Gruissem, L. Hennig, L. Thiele, and E. Zitzler. A systematic comparison and evaluation of biclustering methods for gene expression data. *Bioinformatics*, 22(9):1122–1129, May 2006.

[80] J. Quinlan. Induction of decision trees. *Machine learning*, pages 81–106, 1986.

[81] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2008. ISBN 3-900051-07-0.

[82] R. Santamaría, R. Therón, and L. Quintales. Bicoverlapper: A tool for bicluster visualization. *Bioinformatics*, 24(9):1212, 2008.

[83] D. Shepard. A two-dimensional interpolation function for irregularly-spaced data. In *Proceedings of the 1968 23rd ACM national conference*, pages 517–524. ACM, 1968.

[84] V. Stodden. *Model selection when the number of variables exceeds the number of observations*. Phd thesis, Standford University, 2006.

[85] D. Stranneby and W. Walker. *Digital Signal Processing and Applications*. Elsevier, 2004.

[86] M. Sugiyama, T. Kanamori, T. Suzuki, M. C. d. Plessis, S. Liu, and I. Takeuchi. Density-difference estimation. *Neural Comput.*, 25(10):2734–2775, Oct. 2013.

[87] J. Sun, S. Papadimitriou, and P. S. Yu. Window-based tensor analysis on high-dimensional and multi-aspect streams. In *Proceedings of the Sixth International Conference on Data Mining*, ICDM '06, pages 1076–1080, Washington, DC, USA, 2006. IEEE Computer Society.

[88] Y. Tanaka, K. Iwamoto, and K. Uehara. Discovery of time-series motif from multi-dimensional data based on mdl principle. *Mach. Learn.*, 58(2-3):269–300, Feb. 2005.

[89] R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B*, 58, 1994.

[90] R. Tibshirani. Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society, Series B*, 58:267–288, 1994.

[91] H. Turner, T. Bailey, and W. Krzanowski. Improved biclustering of microarray data demonstrated through systematic performance tests. *Computational Statistics & Data Analysis*, 48(2):235–254, 2005.

[92] D. United Nations. World population ageing: 1950-2050. Technical report, New York, NY, USA, 2002.

[93] A. Vahdatpour, N. Amini, and M. Sarrafzadeh. Toward unsupervised activity discovery using multi-dimensional motif detection in time series. In *Proceedings of the 21st International Jont Conference on Artifical Intelligence*, IJCAI'09, pages 1261–1266, San Francisco, CA, USA, 2009. Morgan Kaufmann Publishers Inc.

[94] O. van de Rest, B. A. Schutte, J. Deelen, S. A. Stassen, E. B. van den Akker, D. van Heemst, P. Dibbets-Schneider, et al. Metabolic effects of a 13-weeks lifestyle intervention in older adults: The growing old together study. *Aging (Albany NY)*, 8(1):111, 2016.

[95] O. van de Rest, B. A. Schutte, J. Deelen, S. A. Stassen, E. B. van den Akker, D. van Heemst, P. Dibbets-Schneider, et al. Metabolic effects of a 13-weeks lifestyle intervention in older adults: The growing old together study. *Aging (Albany NY)*, 8(1):111, 2016.

[96] J. Van Hoof, H. Kort, P. Rutten, and M. Duijnstee. Ageing-in-place with the use of ambient intelligence technology: Perspectives of older users. *International journal of medical informatics*, 80(5):310–331, 2011.

[97] D. Vanderkam, J. Allaire, J. Owen, D. Gromer, P. Shevtsov, and B. Thieurmel. *dygraphs: Interface to 'Dygraphs' Interactive Time Series Charting Library*, 2016. R package version 0.8.

[98] J. Vanschoren, U. Vespier, S. Miao, M. Meeng, R. Cachucho, and A. Knobbe. Large-scale sensor network analysis: applications in structural health monitoring. In *Big Data Management, Technologies, and Applications*, pages 314–347. IGI Global, 2014.

[99] G. Veiga, L. Ketelaar, W. De Leng, R. Cachucho, J. N. Kok, A. Knobbe, C. Neto, and C. Rieffe. Alone at the playground. *European Journal of Developmental Psychology*, 14(1):44–61, 2017.

[100] G. Veiga, W. Leng, R. Cachucho, L. Ketelaar, J. N. Kok, A. Knobbe, C. Neto, and C. Rieffe. Social competence at the playground: Preschoolers during recess. *Infant and Child Development*, 26(1), 2017.