Curriculum Vitae

Luis Torgo

2020-03-31

Contents

Career Highlights

Luis Torgo is a Canada Research Chair (Tier 1) on Spatiotemporal Ocean Data Analytics and a Professor of Computer Science at the Faculty of Computer Science of the Dalhousie University, Canada, an Associate Professor of the Department of Computer Science of the Faculty of Sciences of the University of Porto, Portugal. He is a member of the Institute for Big Data Analytics at Dalhousie, and he is also a senior researcher of LIAAD / INESC Tec, and a current member of the board of this research lab.

Luis Torgo is also an invited professor of the Stern Business School of the New York University where he has been collaborating since 2014 in the Master of Science in Business Analytics.

He has been doing research in the area of Data Mining and Machine Learning since 1990, and has published over 100 papers in several forums of these areas. Luis Torgo is the author of the widely acclaimed *Data Mining with R* book published by CRC Press in 2010 with a strongly revised second edition that appeared in January of 2017. He has been involved in many research projects under different roles and involving different types of organizations.

His current broad research interests revolve around analyzing data from dynamic environments, with a particular focus on time and space-time dependent data sets, in the search for unexpected events. In terms of application domains his research is frequently linked with ecological/biological as well as financial domains.

Luis Torgo main contributions to the state of the art on data mining and machine learning are related with tree-based regression methods and more recently with utility-based forecasting methods.

He has a strong experience of teaching different subjects at different academic levels but also in non-academic settings. He is frequently invited for giving short courses on using R for data mining around the world.

Luis Torgo is the CEO and one of the founding partners of KNOYDA a company devoted to training and consulting within data science.

1 Qualifications

- *2000*, Ph.D. on Computer Science, Faculty of Sciences, University of Porto
- *1994*, Summer School on Artificial Intelligence for Multi-Agent Systems: methodologies and applications , Troia, Portugal
- *1992*, Pedagogical and Scientific abilities examination at the Faculty of Economics of the University of Porto (equivalent to MSc. by Portuguese regulations)

- *1991*, Summer School on Advanced Topics in Artificial Intelligence, Czech Republic
- *1989*, Degree in Systems and Informatics Engineering from the University of Minho, Portugal

2 Professional Experience

- [*2019* – . . .], Tier 1 Canada Research Chair on Spatiotemporal Data Aanalytics at the Faculty of Computer Science of the Dalhousie University, Canada
- [*2018* – . . .], Professor of the Faculty of Computer Science of the Dalhousie University, Canada
- [*2009* – . . .], Associate Professor of the Department of Computer Science of the Faculty of Sciences of the University of Porto, Portugal
- [*2007* – . . .], Senior Researcher of the Laboratory of Artificial Intelligence and Data Analysis belonging to the INESC Tec Associated Lab, Portugal
- [*2000* – *2009*], Assistant Professor of the Faculty of Economics of the University of Porto, Portugal
- [*1993* – *2000*], Assistant of the Faculty of Economics of the University of Porto, Portugal
- [*1989* – *2007*], Researcher of the Laboratory of Artificial Intelligence and Computer Science (LIACC) of the University of Porto, Portugal

3 Teaching Activities

3.1 Faculty Subjects

Undergraduate Level

- [2019/20 – . . .]
  Foundations of Data Science using R
  Faculty of Computer Science, Dalhousie University , *responsible*
- [2015/16 – 2017/18]
  Computers Laboratory
  Degree on Computer Science and Master Degree in Network and Information Systems Engineering, Faculty of Sciences, University of Porto, *responsible*

- [2015/16]
  Initiation to Scientific Research
  Degree on Computer Science and Master Degree in Network and Information Systems Engineering, Faculty of Sciences, University of Porto, *responsible*
- [2012/13 – 2015/16]
  Data Structures and Algorithms
  Offered to several degree programs of the Faculty of Sciences, University of Porto, *member of the team*
- [2009/10 – 2011/12]
  Functional Programming
  Degree on Computer Science and Master's Degree in Network and Information Systems Engineering, Faculty of Sciences, University of Porto, *member of the team*
- [2011/12 – 2015/16]
  Introduction to Programming
  Offered to several degree programs of the Faculty of Sciences, University of Porto, *member of the team*
- [2009/10 – 2010/2011]
  Introduction to Programming
  Offered to several degree programs of the Faculty of Sciences, University of Porto, *responsible*
- [2008/09 – 2011/12]
  Data Structures
  Degree on Computer Science and Master Degree in Network and Information Systems Engineering, Faculty of Sciences, University of Porto, *member of the team*
- [2006/07]
  Informatics
  Degree in Economics, Faculty of Economics, University of Porto, *responsible*
- [2004/05 – 2008/09]
  Applied Informatics
  Degree in Economics, Faculty of Economics, University of Porto, *member of the team*
- [1990/91- 2004/05]
  Informatics
  Degree in Economics, Faculty of Economics, University of Porto, *member of the team*
- [1989 – 2007]
  Complements of Artificial Intelligence

Degree in Computer Science, Faculty of Sciences, University of Porto, *responsible*

Graduate Level

- [2014/15 – . . .]
  Data Mining in R
  Master of Science in Business Analytics, Stern Business School, University of New York, *responsible*
- [2014/15 – . . .]
  Predictive Analytics
  Post-graduation course Business Intelligence and Analytics, PBS - University of Porto Business School, *responsible*
- [2014/15 – 2017/18]
  Fraud Detection
  Master's on Information Security, Faculty of Sciences, University of Porto, *responsible*
- [2011/12 – 2017/18]
  Data Mining I
  Master's on Computer Science, Faculty of Sciences, University of Porto, *responsible*
- [2010/11 – 2012/13]
  Prediction Models
  Post-graduation course Information Managment and Marketing Intelligence, PBS - University of Porto Business School, *responsible*
- [2009/10 – 2017/18]
  Knowledge Discovery from Databases
  PhD program on Informatics (MAP-I), Universities of Aveiro, Braga and Porto, *member of the team*
  (*slides of my part of the classes*)
- [2008/09 – 2009/10]
  Analytical Methods of Fraud Detection
  Post-graduation course on Fraud Managment, PBS - University of Porto Business School, *responsible*
- [2007/08 – 2008/09]
  Advanced Topics in Data Mining and Logic Programming
  PhD program on Informatics (MAP-I), Universities of Aveiro, Braga and Porto, *member of the team*
- [2006/07]
  Databases

PhD program on Computacional Biology, Institute Gulbenkian of Science, *responsible*

- [2004/05 – 2008/09]
  Time Series Analysis
  Masters on New Media and E-science, Josef Stefan International Postgraduate School, Ljubljana, Slovenia, *responsible*
- [2004/05 – 2006/07]
  Web Mining
  Masters on Artificial Intelligence and Intelligent Systems, Faculties of Economics and Engeneering, University of Porto, *responsible*
- [2004/05 – 2006/07]
  Databases and Programming
  Masters on Artificial Intelligence and Intelligent Systems, Faculties of Economics and Engeneering, University of Porto, *responsible*
- [2004/05 – 2006/07]
  Data Mining I
  Masters on Artificial Intelligence and Intelligent Systems, Faculties of Economics and Engeneering, University of Porto, *member of the team*
- [2002/03 – 2004/05]
  Web Mining
  Masters on Artificial Intelligence, Faculties of Sciences, Economics and Engeneering, University of Porto, *responsible*
- [2000/01 – 2008/09]
  Data Mining I
  Masters on Data Analysis and Decision Support Systems, Faculty of Economics, University of Porto, *member of the team*
- [2000/01 – 2008/09]
  Databases and Programming
  Masters on Data Analysis and Decision Support Systems, Faculty of Economics, University of Porto, *responsible*
- [2000/01 – 2008/09]
  Artificial Intelligence Methodologies
  Masters on Artificial Intelligence, Faculties of Sciences, Economics and Engeneering, University of Porto, *responsible*

3.2 Tutorials

- [2018]
  Learning with Imbalanced Domains, a tutorial
  2nd International Workshop on Learning with Imbalanced Domains: Theory and Applications Co-located with ECML/PKDD 2018

- [2006]
  Using R for Data Mining and Scientific Discovery
  9th International Conference on Discovery Science (DS-2006), Barcelona, Spain

3.3 Short Courses

- [Jan 2020] **Spatio-Temporal Data Mining**
  1 day course together with Stan Matwin at the ACM SIGIR/SIGKDD Africa Summer School on Machine Learning for Data Mining and Search, AFIRM 2020, Cape Town, South Africa
- [Jul 2019] **Advanced Predictive Analytics using R**
  1 day course at LTPlabs, Porto, Portugal
- [Apr 2018, Mar 2019] **Predictive Analytics using R**
  2 days course at Universitat Politècnica de València, Valencia, Spain
- [Jan 2018]
  Data Mining using R
  4 days course at Jozef Stefan Institute, Ljubljana, Slovenia.
- [Sep 2017]
  Data Science in Practice
  5 hours course, IIMT Executive Programs, at International Institute of Managment in Technology, University of Fribourg, Switzerland
- [May 2017]
  Data Mining with R
  2 days course (in Portuguese) at Universidade Federal Fulminense, Niterói Rio de Janeiro, Brazil.
- [Jan 2017]
  Temporal and Spatio-Temporal Data Mining using R
  4 days course at Jozef Stefan Institute, Ljubljana, Slovenia.
- [Aug 2014 and Jul 2015]
  Data Mining in R - learning with case studies
  6 days course at Istanbul Quantitative Lectures, University of Istanbul, Turkey.
- [Feb 2014 and Aug 2014]
  Data Mining in R - learning with case studies
  3 days course at Institute for Big Data Analytics, Dalhousie University, Canada.
- [2013]
  Text Mining (co-teacher together with Nitin Indurkhya)
  regular 4 weeks web course at statistics.com

- [Mar 2013 and April 2014]
  Data Mining in R - learning with case studies
  3 days course at DIKW Academy
- [Nov 2012 and Feb 2014]
  Data Mining for Fraud Detection using R
  7h course at Actuarial Seguros, Lisbon, Portugal.
- [2011 – 2015]
  Data Mining in R - learning with case studies
  regular bi-annual 4 weeks web course at statistics.com
- [2011 – 2012]
  Short introduction to Data Mining
  3h short course at Instituto Superior Técnico, Lisboa, Portugal
- [2011]
  Data Mining using R
  12h course at Summer School, Instituto de Ciências Matemáticas e Computação, University of São Paulo, São Carlos, Brazil
- [2010]
  Data Mining with R: a short course
  10h course at Dipartimento di Informatica, Univerisità degli Studi di Bari, Bari, Italy
- [2007]
  An introduction to the R environment
  30h course at Caixa Economica Federal, Brasilia, Brazil
- [2007]
  An Introduction to R
  5h course at University of Beira Interior, Covilha, Portugal
- [2005]
  Data Mining with R
  6h course at ACAI Summer School on Knowledge Discovery , 561 views on Dez/2016
- [2004]
  An Autonomous Trading System
  International Summer School on Data Analysis, Lisbon, Portugal
- [2003]
  Time Series Analysis
  International Summer School on Neural Networks (NN'03), Porto, Portugal
- [2002]
  Financial Applications
  International Summer School on Neural Networks (NN'02), Porto, Portugal

- [2002]
  Time Series Analysis
  International Summer School on Neural Networks (NN'02), Porto, Portugal

- [1998]
  Introduction to Regression Methods useful in Data Mining
  International Summer School on Knowledge Discovery and Data Mining: Methods and Applications, Caminha, Portugal

- [1994]
  Applications of Propositional Learning Systems: examples and techniques
  Instituto de Ciências Matemáticas e Computação,University of Sao Paulo, Brazil

- [1994]
  Numerical Classification and Prediction
  International Workshop on Artificial Intelligence Techniques, Czech Republic

3.4 Pedagogical Publications

- [2009]
  A Linguagem R - programação para a análise de dados
  Escolar Editora

- [2006]
  Introdução à Programação em R
  Document contributed to the R project, freely available here

- [2006]
  Introdução à Programação em R
  Document (in Portuguese) supporting the subject Databases and Programming, Master on Data Analysis and Decision Support Systems, Faculty of Economics, University of Porto

- [2006]
  Introdução aos Sistemas de Gestão de Bases de Dados
  Document (in Portuguese) supporting the subject Databases and Programming, Master on Data Analysis and Decision Support Systems, Faculty of Economics, University of Porto

- [2003]
  Programação, Análise de Dados e Sistemas de Apoio à Decisão usando o R
  Document (in Portuguese) supporting the subject Informatics and Applied Informatics, degrees in Economics and Managment, Faculty of Economics, University of Porto

- [1997]
  Windows 95, uma breve introdução
  Document (in Portuguese) supporting the subject Informatics, degrees in Economics and Managment, Faculty of Economics, University of Porto

4 Research Activities

This section presents my main research activities that may help in understanding the impact of my research. Assessing scientific quality is not an easy task and it is still source of debate among many scholars. One of the key issues on this assessment is typically the publication record and associated bibliometric numbers. However, there are many disciplines where other aspects of research activity can be regarded as significant in terms of the impact on the community, and that is frequently the case of Computer Science, the field were I belong. For instance, developping a software tool that is widely used by the community and allows for further advances may well be more important than many publications and yet it is seldom considered as a criterion as it is not easy to assess without being involved in the concrete field of expertise. In my personal case I have developped many software programs that are widely used by the community, like for instance several R packages that have hundreds of monthly downloads. I have also created and still mantain a free regression data set repository that is used in most papers that test new regression algorithms whithin the machine learning community. This type of contributions, although relevant and important in my personal opinion, are hardly reflected in any type of research bliblometrics (with the exception of the data set repository whose citations are counted by Google Scholar although not all papers that use these data sets, cite the repository).

In spite of the previously mentioned limitations I have collected some data concerning bibliometric information related with my publications with the goal of helping in assessing the impact of this aspect of my research activities. The inclusion of this information requires some clarifications. Bibliometric data involves two main decisions: (i) whether to use it or not; and (ii) if yes, from which source. The first decision is hard to escape - it is the norm nowadays to evaluate CVs based on these numbers. Unfortunately this process often leads to unwanted bad decisions. There is a growing awareness of the research community that these numbers may be seriously misleading, particularly in some disciplines (like Computer Science) where the publication *culture* is rather different from other disciplines, for instance in the case of international conferences. While in most disciplines works published in conferences are not subject to peer reviewing, that is not the case in CS where international conferences are always peer reviewed and some are far more competitive than most journals. This leads to a clear bias on some of the frequently used bibliometric indices that frequently disregard conferences. Regarding this issue of the bias and risk of misjudgment of bibliometrics it may be interesting to point out several very interesting recent

papers on this topic:

- Diana Hicks, Paul Wouters, Ludo Waltman, Sarah de Rijcke & Ismael Rafols.*Bibliometrics: The Leiden Manifesto for research metrics.* Nature 520, 429–431 (23 April 2015) doi:10.1038/520429a
- Paula Stephan, Reinhilde Veugelers & Jian Wang. *Reviewers are blinkered by bibliometrics.* Nature 544, 411–412 (27 April 2017) doi:10.1038/544411a
- Rinze Benedictus, Frank Miedema & Mark W. J. Ferguson. *Fewer numbers, better science.* Nature 538, 453–455 (27 October 2016) doi:10.1038/538453a

These and many other works have been trying to raise the awareness of the community for the danger of these bibliometric-based decisions. Still, as this is still the norm I will present some of these numbers with the list of publications given below.

The second decision concerns the source of the bibliometric numbers. I have considered 3 sources: Thomson ISI, Elsevier Scopus and Google Scholar. All of them have potential drawbacks. Still, the decision was to select Google Scholar (GS). This index overcomes some limitations the other indices have, particularly for Computer Science, as it is clearly much more inclusive, although with an increased risk of inflating numbers. A few examples of my personal CV provide illustrations of the problems. My top cited publication is my book *Data Mining with R* published by a major publishing house (CRC Press from Taylor and Francis). For some reason that I cannot explain but most probably related with financial issues between the companies involved, this book is not indexed by Scopus. So, a peer reviewed book by a major publisher is non-existent for this widely used bibliometric index; (ii) one of my recent journal publications is on ACM Computing Surveys, one of the top journals of Computer Science according to the 2017 data from Thomson ISI impact factor. At some point in time (12/Dec/2017), according to Google Scholar (GS) this publication from 2016 had 42 citations. According to Scopus this citation number was 20! If we check the concrete citations which is possible on GS we can observe that 6 of these were self-citations, so we can reduce this to 36, but this is still too far from 20. These and other similar effects lead to divergent values like the fact that my h-index according to Scopus is significantly lower than that reported in GS, a very common phenomenon for computer science researchers as mentioned in [^1]. Still, GS is not without problems either. For instance, the same book I've mentioned above has two quite different editions currently, the second being more than 150 pages longer, with the other pages having been significantly revised. In spite of this GS insists in not considering these two books which obviously has an impact on the bibliometrics (I've inserted it manually but citations are not being counted by GS).

[^1] - Judit Bar-Ilan. *Which h-index? — A comparison of WoS, Scopus and Google Scholar.* Scientometrics, Volume 74, Issue 2, pp. 257–271.

Summarizing, I have decided to include bibliometric data from Google Scholar, even-though I'm well aware that this is a source of debate but my decision was

essentially guided by completeness criteria both in terms of computer science in general and in my particular case. Moreover, in order to try to provide more information on the quality of the peer-reviewed conference papers, I have added the CORE rank of the respective conferences. CORE is a widely used source of information on the computer science conferences reputation. It is used wordwide by many funding agencies and universities where the specificities of the computer science field have been recognised. This austrolasian association provides a rank for international conferences using a scale from A* till Unranked. Further details on the criteria and meaning of this classification schema can be obtained here.

4.1 Publications

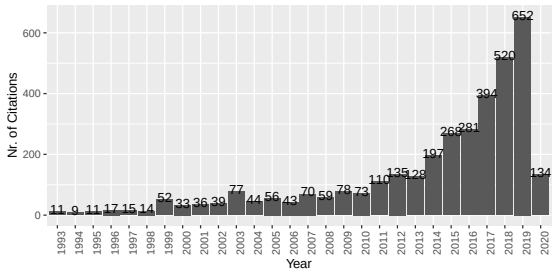
Global bibliometrics

Important Note: All citation numbers were obtained from my Google Scholar profile on 2020-03-31.

Global scores:

Total Nr. of Citations	h-index	i10-index
3596	27	55

Yearly evolution of the number of citations:



Books

[1] L. Torgo. *Data Mining with R: Learning with Case Studies, Second Edition (chinese edition)*. China Machine Press (CMP), 2018. ISBN: 9787111596660. (*extra information*)

[2] L. Torgo. *Data Mining with R: Learning with Case Studies, Second Edition*. Chapman and Hall/CRC, 2017.
(*document*) (*extra information*)

[3] L. Torgo. *Data Mining with R: Learning with Case Studies (Chinese Edition)*. China Machine Press, 2012.
(*document*)

[4] L. Torgo. *Data Mining with R: Learning with Case Studies*. Chapman and Hall/CRC Data Mining and Knowledge Discovery Series. CRC Press, 2010.
(*526 citations*) (*document*) (*extra information*)

[5] L. Torgo. *A Linguagem R, programação para a análise de dados*. Escolar Editora, 2009.
(*document*)

Edition of Books

[1] A. Jorge, L. Torgo, P. Brazdil, et al., ed. *Knowledge Discovery in Databases: PKDD 2005: 9th European Conference on Principles and Practice of Knowledge Discovery in Databases*. LNAI 3721. Springer, 2005.
(*2 citations*)

[2] J. Gama, R. Camacho, P. Brazdil, et al., ed. *Machine Learning: ECML 2005: 16th European Conference on Machine Learning*. LNAI 3720. Springer, 2005.
(*2 citations*)

Chapters in Books

[1] N. Guimarães, L. Torgo, and A. Figueira. "Twitter as a Source for Time- and Domain-Dependent Sentiment Lexicons". In: *Social Network Based Big Data Analysis and Applications*. 2018, pp. 1-19.
(*2 citations*) (*document*)

[2] L. Torgo. "Regression Trees". In: *Encyclopedia of Machine Learning and Data Mining*. Ed. by C. Sammut and G. I. Webb. Springer, 2016, pp. 1080-1083.
(*document*)

[3] L. Torgo. "Model Trees". In: *Encyclopedia of Machine Learning and Data Mining*. Ed. by C. Sammut and G. I. Webb. Springer, 2016, pp. 845-843.
(*document*)

[4] L. Torgo. "Regression Trees". In: *Encyclopedia of Machine Learning*. Ed. by C. Sammut and G. I. Webb. Springer, 2011, pp. 842-845.
(*1 citations*)

[5] L. Torgo. "Model Trees". In: *Encyclopedia of Machine Learning*. Ed. by C. Sammut and G. I. Webb. Springer, 2011, pp. 684-686.

[6] L. Torgo and C. Soares. "Resource-bounded Outlier Detection Using Clustering Methods". In: *Data Mining for Business Applications*. Ed. by C. Soares and R. Ghani. Frontiers in Artificial Intelligence and Applications. IOS Press, 2010,

pp. 84-98.
(*11 citations*)

[7] P. Flach, H. Blockeel, T. Gartner, et al. "Data Mining and Decision Support, Integration and Collaboration". In: *On the road to knowledge: mining 21 years of UK traffic accident reports*. Ed. by D. Mladenic, N. Lavrac, M. Bohanec and S. Moyle. Morgan Kaufmann, 2003, pp. 143-156.
(*13 citations*)

[8] T. Hellström and L.Torgo. "Post processing trading signals for improved trading performance". In: *Data Mining III*. WIT Press, 2002, pp. 437-447.
(*document*)

[9] P. Brazdil and L. Torgo. "Knowledge Acquisition via Knowledge Integration". In: *Current Trends in Knowledge Acquisition*. Ed. by B. e. a. Wielinga. IOS Press, 1990, pp. 90-104.
(*82 citations*) (*document*)

Journals

[1] M. Etemad, Z. Etemad, A. Soares, et al. "Wise Sliding Window Segmentation: A classification-aided approach for trajectory segmentation". In: *arXiv* arXiv:2003.10248 (2020).
(*document*)

[2] M. Monteiro, M. S. Baptista, J. Séneca, et al. "Understanding the Response of Nitrifying Communities to Disturbance in the McMurdo Dry Valleys, Antarctica". In: *Microorganisms* 8.3 (2020). ISSN: 2076-2607. DOI: 10.3390/microorganisms8030404.
(*document*)

[3] V. Cerqueira, L. Torgo, and C. Soares. "Machine Learning vs Statistical Methods for Time Series Forecasting: Size Matters". In: *arXiv* arXiv:1909.13316 (2019).
(*3 citations*) (*document*)

[4] A. G. G. de Sousa, M. P. Tomasino, P. Duarte, et al. "Diversity and Composition of Pelagic Prokaryotic and Protist Communities in a Thin Arctic Sea-Ice Regime". In: *Microbial ecology* ? (2019), pp. 1-21. DOI: https://doi.org/10.1007/s00248-018-01314-2.
(*2 citations*) (*document*)

[5] N. Moniz and L. Torgo. "A review on web content popularity prediction: Issues and open challenges". In: *Online Social Networks and Media* 12 (2019), pp. 1-20. ISSN: 2468-6964. DOI: https://doi.org/10.1016/j.osnem.2019.05.002.
(*1 citations*) (*document*)

[6] A. Figueira, N. Guimaraes, and L. Torgo. "A Brief Overview on the Strategies to Fight Back the Spread of False Information". In: *Journal of Web Engineering*

18.4 (2019), pp. 319-352. DOI: https://doi.org/10.13052/jwe1540-9589.18463.
(*document*)

[7] V. Cerqueira, L. Torgo, F. Pinto, et al. "Arbitrage of Forecasting Experts". In: *Machine Learning* 108 (2019), pp. 913-944.
(*8 citations*) (*document*)

[8] V. Cerqueira, L. Torgo, and I. Mozetic. "Evaluating time series forecasting models: An empirical study on performance estimation methods". In: *arXiv* arXiv:1905.11744 (2019).
(*2 citations*) (*document*)

[9] P. Branco, L. Torgo, and R. P. Ribeiro. "Pre-processing approaches for imbalanced distributions in regression". In: *Neurocomputing* 343 (2019), pp. 76-99. DOI: 10.1016/j.neucom.2018.11.100.
(*2 citations*) (*document*)

[10] P. Branco, L. Torgo, and R. P. Ribeiro. "REBAGG: REsampled BAGGing for Imbalanced Regression". In: *Proceedings of Machine Learning Research (PMLR)* 94 (2018), pp. 1-15.
(*4 citations*) (*extra information*)

[11] L. Torgo, S. Matwin, G. Weiss, et al. "Cost-Sensitive Learning: Preface". In: *Proceedings of Machine Learning Research (PMLR)* 88 (2018), pp. 1-3.
(*document*)

[12] P. Branco, L. Torgo, and R. P. Ribeiro. "Resampling with neighbourhood bias on imbalanced domains". In: *Expert Systems* 35.4 (2018). DOI: 10.1111/exsy.12311.
(*3 citations*) (*document*)

[13] I. Mozetic, L. Torgo, V. Cerqueira, et al. "How to evaluate sentiment classifiers for Twitter time-ordered data?" In: *PLOS ONE* 13.3 (2018), p. e0194317.
(*12 citations*) (*document*) (*extra information*)

[14] N. Moniz and L. Torgo. "Multi-Source Social Feedback of Online News Feeds". In: *arXiv* arXiv:1801.07055 (2018).
(*12 citations*) (*document*)

[15] H. Ribeiro, T. de Sousa, J. Santos, et al. "Potential of dissimilatory nitrate reduction pathways in polycyclic aromatic hydrocarbon degradation". In: *Chemosphere* 199 (2018), pp. 54-67.
(*13 citations*) (*document*)

[16] M. Monteiro, J. Séneca, L. Torgo, et al. "Environmental controls on estuarine nitrifying communities along a salinity gradient". In: *Aquatic Microbial Ecology* 80 (2) (2017), pp. 167-180.
(*2 citations*) (*document*)

[17] L. Torgo, B. Krawczyk, P. Branco, et al. "Learning with Imbalanced Domains: preface". In: *Proceedings of Machine Learning Research (PMLR)* 74 (2017),

pp. 1-6.
(*document*)

[18] P. Branco, L. Torgo, and R. P. Ribeiro. "SMOGN: a Pre-processing Approach for Imbalanced Regression". In: *Proceedings of Machine Learning Research (PMLR)* 74 (2017), pp. 36-50.
(*7 citations*) (*document*)

[19] N. Moniz, P. Branco, and L. Torgo. "Evaluation of Ensemble Methods in Imbalanced Regression Tasks". In: *Proceedings of Machine Learning Research (PMLR)* 74 (2017), pp. 129-140.
(*5 citations*) (*document*)

[20] N. Moniz, L. Torgo, M. Eirinaki, et al. "A Framework for Recommendation of Highly Popular News Lacking Social Feedback". In: *New Generation Computing* 35 (4) (2017), pp. 417-450.
(*4 citations*) (*document*)

[21] N. Moniz, P. Branco, and L. Torgo. "Resampling Strategies for Imbalanced Time Series Forecasting". In: *International Journal of Data Science and Analytics* 3.3 (2017), pp. 161-181.
(*10 citations*) (*document*)

[22] N. Moniz, L. Torgo, and J. Vinagre. "Data-driven relevance judgments for ranking evaluation". In: *CoRR* abs/1612.06136 (2016).
(*document*)

[23] L. Baía and L. Torgo. "A comparative study of approaches to forecast the correct trading actions". In: *Expert Systems* 34.1 (2016), pp. e12169-n/a.
(*2 citations*) (*document*)

[24] P. Branco, L. Torgo, and R. Ribeiro. "A Survey of Predictive Modeling on Imbalanced Domains". In: *ACM Comput. Surv.* 49.2-31 (2016).
(*257 citations*) (*document*)

[25] P. Branco, R. Ribeiro, and L. Torgo. "A UBL: an R package for Utility-based Learning". In: *CoRR* abs/1604.08079 (2016).
(*15 citations*) (*document*)

[26] N. Moniz and L. Torgo. "Socially Driven News Recommendation". In: *CoRR* abs/1506.01743 (2015).
(*1 citations*) (*document*)

[27] P. Branco, L. Torgo, and R. Ribeiro. "A Survey of Predictive Modelling under Imbalanced Distributions". In: *CoRR* abs/1505.01658 (2015).
(*document*)

[28] L. Torgo, P. Branco, R. P. Ribeiro, et al. "Re-sampling Strategies for Regression". In: *Expert Systems* 32.3 (2015), pp. 465-476.
(*63 citations*) (*document*)

[29] L. Torgo. "An Infra-Structure for Performance Estimation and Experimental Comparison of Predictive Models in R". In: *CoRR* abs/1412.0436 (2014).
(*29 citations*) (*document*) (*extra information*)

[30] J. Vanschoren, J. N. van Rijn, B. Bischl, et al. "OpenML: networked science in machine learning". In: *SIGKDD Explorations Newsletter* 15.2 (2013), pp. 49-60.
(*401 citations*) (*document*)

[31] B. Drury, L. Torgo, and J. J. Almeida. "Classifying News Stories with a Constrained Learning Strategy to Estimate the Direction of a Market Index". In: *IJCSA* 9.1 (2012), pp. 1-22.
(*13 citations*)

[32] M. Herrera, L. Torgo, J. Izquierdo, et al. "Predictive models for forecasting hourly urban water demand". In: *Journal of Hydrology* 387.1-2 (Jun. 2010), pp. 141-150.
(*321 citations*) (*document*)

[33] L. Torgo and R. P. Ribeiro. "Modelos de Previsão de Valores Extremos e Raros". In: *Boletim da Sociedade Portuguesa de Estat'istica* Primavera 2010 (2010), pp. 15-22.

[34] L. Torgo. "Deteção de fraude usando o R: um caso de estudo". In: *Boletim da Sociedade Portuguesa de Estatística* (2009).

[35] R. Ribeiro and L. Torgo. "A Comparative Study on Predicting Algae Blooms in Douro River, Portugal". In: *Ecological Modelling - Selected Papers from the 5th European Conference on Ecological Modelling* 212.1-2 (2008), pp. 86-91.
(*19 citations*) (*document*)

[36] A. Silva, A. Jorge, and L. Torgo. "Design of an end-to-end method to extract information from tables". In: *International Journal on Document Analysis and Recognition* 8.2-3 (2006), pp. 144-171.
(*86 citations*)

[37] L. Torgo and J. P. Costa. "Clustered Partial Linear Regression". In: *Machine Learning* 50.3 (2003), pp. 303-319.
(*13 citations*) (*document*)

[38] L. Torgo. "Thesis: Inductive learning to tree-based regression models.". In: *AI Commun.* 13.2 (2000), pp. 137-138.

[39] L. Torgo and J. Gama. "Regression using Classification Algorithms". In: *Intelligent Data Analysis* 1.4 (1997).
(*55 citations*) (*document*)

Full Papers at International Conferences with Peer Reviewing

[1] C. Bellinger, P. Branco, and L. Torgo. "The CURE for Class Imbalance". In:

Discovery Science. Ed. by P. Kralj Novak, T. Šmuc and S. Džeroski. Springer International Publishing, 2019, pp. 3-17. ISBN: 978-3-030-33778-0.
CORE rank: None
(*document*)

[2] V. Cerqueira, L. Torgo, and C. Soares. "Layered Learning for Early Anomaly Detection: Predicting Critical Health Episodes". In: *Discovery Science*. Ed. by P. Kralj Novak, T. Šmuc and S. Džeroski. Springer International Publishing, 2019, pp. 445-459. ISBN: 978-3-030-33778-0.
CORE rank: None
(*document*)

[3] I. Areosa and L. Torgo. "Explaining the Performance of Black Box Regression Models". In: *2019 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*. 2019, pp. 110-118.
CORE rank: None
(*document*)

[4] P. Branco and L. Torgo. "A Study on the Impact of Data Characteristics in Imbalanced Regression Tasks". In: *2019 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*. 2019, pp. 193-202.
CORE rank: None
(*document*)

[5] M. Etemad, A. Soares, S. Matwin, et al. "On Feature Selection and Evaluation of Transportation Mode Prediction Strategies". In: *Proceedings of the Workshops of the EDBT/ICDT 2019 Joint Conference, EDBT/ICDT 2019, Lisbon, Portugal, March 26, 2019*. Ed. by P. Papotti. Vol. 2322. CEUR Workshop Proceedings. CEUR-WS.org, 2019.
CORE rank: None
(*document*)

[6] I. Areosa and L. Torgo. "Visual Interpretation of Regression Error". In: *Progress in Artificial Intelligence*. Ed. by P. Moura Oliveira, P. Novais and L. P. Reis. Springer International Publishing, 2019, pp. 473-485. ISBN: 978-3-030-30244-3.
(*document*)

[7] M. Oliveira, N. Moniz, L. Torgo, et al. "Biased Resampling Strategies for Imbalanced Spatio-Temporal Forecasting". In: *2019 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*. 2019, pp. 100-109.
CORE rank: None
(*document*)

[8] N. Guimarães, Á. Figueira, and L. Torgo. "Contributions to the Detection of Unreliable Twitter Accounts through Analysis of Content and Behaviour". In: *Proceedings of the 10th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, IC3K 2018, Volume 1: KDIR, Seville, Spain, September 18-20, 2018*. 2018, pp. 90-99. DOI:

10.5220/0006932800900099.
CORE rank: None
(*document*)

[9] P. Branco, L. Torgo, and R. Ribeiro. "MetaUtil: Meta Learning for Utility Maximization in Regression". In: *Proceedings of the International Conference on Discovery Science, DS'18*. Springer, 2018, pp. 129-143.
CORE rank: None
(*1 citations*) (*extra information*)

[10] M. Oliveira, L. Torgo, and V. S. Costa. "Evaluation procedures for forecasting with spatio-temporal data". In: *Proceedings of the ECML/PKDD'2018 Conference*. Springer, 2018, pp. ??-??
CORE rank: A
(*8 citations*) (*extra information*)

[11] V. Cerqueira, F. Pinto, L. Torgo, et al. "Constructive Aggregation and its Application to Forecasting with Dynamic Ensembles". In: *Proceedings of the ECML/PKDD'2018 Conference*. Springer, 2018, pp. ??-??
CORE rank: A
(*extra information*)

[12] N. Moniz and L. Torgo. "The Utility Problem of Web Content Popularity Prediction". In: *Proceedings of the 29th ACM Conference on Hypertext and Social Media, HT 2018*. 2018, pp. 82-86.
CORE rank: A
(*document*)

[13] Á. Figueira, N. Guimarães, and L. Torgo. "Current State of the Art to Detect Fake News in Social Media: Global Trendings and Next Challenges". In: *Proceedings of the 14th International Conference on Web Information Systems and Technologies, WEBIST 2018, Seville, Spain, September 18-20, 2018.* 2018, pp. 332-339. DOI: 10.5220/0007188503320339.
CORE rank: None
(*document*)

[14] V. Cerqueira, L. Torgo, M. Oliveira, et al. "Dynamic and Heterogeneous Ensembles for Time Series Forecasting". In: *IEEE International Conference on Data Science and Advanced Analytics (DSAA'2017)*. 2017, pp. 242-251.
CORE rank: None
(*document*)

[15] P. Branco, L. Torgo, R. P. Ribeiro, et al. "Learning Through Utility Optimization in Regression Tasks". In: *IEEE International Conference on Data Science and Advanced Analytics (DSAA'2017)*. 2017, pp. 30-39.
CORE rank: None
(*document*)

[16] V. Cerqueira, L. Torgo, J. Smailović, et al. "A Comparative Study of Performance Estimation Methods for Time Series Forecasting". In: *IEEE*

International Conference on Data Science and Advanced Analytics (DSAA'2017). 2017, pp. 529-538.
CORE rank: None
(*document*)

[17] V. Cerqueira, L. Torgo, F. Pinto, et al. "Arbitrated Ensemble for Time Series Forecasting". In: *Proceedings of the ECML/PKDD'2017 Conference*. Ed. by M. Ceci, J. Hollmén, L. Todorovski and C. Vens. Lecture Notes in Artificial Intelligence. Springer, 2017, pp. 478-494.
CORE rank: A (BEST STUDENT MACHINE LEARNING PAPER AWARD)
(*29 citations*)

[18] P. Branco, L. Torgo, and R. P. Ribeiro. "Exploring Resampling with Neighborhood Bias on Imbalanced Regression Problems". In: *Proceedings of 18th EPIA Conference on Artificial Intelligence (EPIA 2017)*. Ed. by E. Oliveira, J. Gama, Z. Vale and H. L. Cardoso. LNCS 10423. Springer, 2017, pp. 513-524.
CORE rank: B
(*3 citations*) (*document*)

[19] V. Cerqueira, L. Torgo, and C. Soares. "Arbitrated Ensemble for Solar Radiation Forecasting". In: *Proceedings of IWANN'2017*. Vol. 10305. LNCS. Springer, 2017, pp. 720-732.
CORE rank: B
(*5 citations*) (*document*)

[20] P. Branco, L. Torgo, and R. P. Ribeiro. "Relevance-based Evaluation Metrics for Multi-class Imbalanced Domains". In: *Advances in Knowledge Discovery and Data Mining - 21th Pacific-Asia Conference, PAKDD 2017, Jeju, South Korea, May 23-26, 2017, Proceedings*. Ed. by J. Kim, K. Shim, L. Cao, J. Lee, X. Lin and Y. Moon. Lecture Notes in Computer Science, vol 10234. Springer. 2017, pp. 698-710.
CORE rank: A
(*14 citations*) (*document*)

[21] N. Guimarães, L. Torgo, and A. Figueira. "Lexicon Expansion System for Domain and Time Oriented Sentiment Analysis". In: *Proceedings of the 8th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*. Vol. 1: KDIR. 2016.
CORE rank: C
(*document*)

[22] N. Moniz, L. Torgo, and M. Eirinaki. "Time-Based Ensembles for Prediction of Rare Events in News Stream". In: *IEEE International Conference on Data Mining Workshops, ICDM Workshops 2016, December 12-15, 2016, Barcelona, Spain*. 2016, pp. 1066-1073. DOI: 10.1109/ICDMW.2016.0154.
CORE rank: A*
(*2 citations*) (*document*)

[23] M. Oliveira, L. Torgo, and V. S. Costa. "Predicting Wildfires: Propositional

and Relational Spatio-Temporal Pre-Processing Approaches". In: *Proceedings of Discovery Science 2016*. LNAI. Springer, 2016.
CORE rank: None
(*3 citations*) (*document*)

[24] N. Moniz, P. Branco, and L. Torgo. "Resampling Strategies for Imbalanced Time Series". In: *Proceedings of DSAA 2016*. 2016.
CORE rank: None
(*12 citations*) (*document*)

[25] A. Martins, A. Dias, E. Silva, et al. "MarinEye - A tool for marine monitoring". In: *OCEANS 2016 - Shanghai*. 2016, pp. 1-7.
CORE rank: None
(*document*)

[26] L. Nezvaloví, L. Popelínsky, L. Torgo, et al. "Class-Based Outlier Detection: Staying Zombies or Awaiting for Resurrection?" In: *Proceeedings of IDA'2015*. Vol. 9385. Lecture Notes in Computer Science . Springer, 2015, pp. 193-204.
CORE rank: A
(*3 citations*) (*document*)

[27] L. Baia and L. Torgo. "Forecasting the Correct Trading Actions". In: *Proceeedings of 17th Portuguese Conference on Artificial Intelligence, EPIA 2015*. LNAI. Springer, 2015, pp. 560-571.
CORE rank: B
(*2 citations*) (*document*)

[28] M. Oliveira and L. Torgo. "Ensembles for Time Series Forecasting". In: *Proceedings of Asian Conference on Machine Learning (ACML'2014)*. Vol. 39. JMLR: Workshop and Conference Proceedings. 2014, pp. 360-370.
CORE rank: None
(*38 citations*) (*document*) (*extra information*)

[29] N. Moniz, L. Torgo, and F. Rodrigues. "Resampling approaches to improve news importance prediction". In: *Advances in Intelligent Data Analysis XIII (IDA'2014)*. Ed. by B. H.. Vol. 8819. LNCS. Springer, 2014, pp. 215-226.
CORE rank: A
(*7 citations*) (*document*)

[30] N. Moniz and L. Torgo. "Improvement of News Ranking through Importance Prediction". In: *Proceeding of KDD'2014 workshop NewsKDD - Data Science for News Publishing*. 2014.
CORE rank: C
(*3 citations*) (*document*) (*extra information*)

[31] L. Torgo, R. P. Ribeiro, B. Pfahringer, et al. "SMOTE for Regression". In: *Proceedings of EPIA'2013*. Springer, 2013.
CORE rank: B
(*63 citations*) (*document*) (*extra information*)

[32] J. van Rijn, B. Bischl, L. Torgo, et al. "OpenML: A collaborative Science Platform". In: *Proceedings of ECML/PKDD'2013*. Springer, 2013, pp. 645-649.
CORE rank: A
(*49 citations*)

[33] O. Ohashi and L. Torgo. "Spatial Interpolation Using Multiple Regression". In: *Data Mining (ICDM), 2012 IEEE 12th International Conference on*. 2012, pp. 1044-1049.
CORE rank: A*
(*11 citations*) (*document*) (*extra information*)

[34] O. Ohashi and L. Torgo. "Wind speed forecasting using spatio-temporal indicators". In: *ECAI 2012 - 20th European Conference on Artificial Intelligence*. Ed. by L. D. Raedt, C. Bessière, D. Dubois, P. Doherty, P. Frasconi, F. Heintz and P. J. F. Lucas. IOS Press, 2012, pp. 975-980.
CORE rank: A
(*32 citations*) (*document*)

[35] B. Drury, G. Dias, and L. Torgo. "A Contextual Classification Strategy for Polarity Analysis of Direct Quotations from Financial News". In: *RANLP*. Ed. by G. Angelova, K. Bontcheva, R. Mitkov and N. Nicolov. RANLP 2011 Organising Committee, 2011, pp. 434-440.
CORE rank: C
(*9 citations*)

[36] B. Drury, L. Torgo, and J. Almeida. "Classifying news stories to estimate the direction of a stock market index". In: *Information Systems and Technologies (CISTI), 2011 6th Iberian Conference on*. 2011, pp. 1-4.
CORE rank: None
(*15 citations*) (*document*)

[37] B. Drury, L. Torgo, and J. J. Almeida. "Guided Self Training for Sentiment Classification". In: *Proceedings of International Conference On Recent Advances in Natural Language Processing (RANLP 2011) - ROBUS workshop*. RANLP 2011 Organising Committee, 2011.
CORE rank: C
(*document*)

[38] L. Torgo and E. Lopes. "Utility-Based Fraud Detection". In: *IJCAI 2011, Proceedings of the 22nd International Joint Conference on Artificial Intelligence*. Ed. by T. Walsh. AAAI Press. IJCAI/AAAI, 2011, pp. 1517-1522.
CORE rank: A*
(*12 citations*) (*document*) (*extra information*)

[39] L. Torgo and O. Ohashi. "2D-interval predictions for time series". In: *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD'2011*. Ed. by C. A. é, J. Ghosh and P. Smyth. ACM, 2011, pp. 787-794.
CORE rank: A*

(*7 citations*) (*document*) (*extra information*)

[40] O. Ohashi, L. Torgo, and R. P. Ribeiro. "Interval Forecast of Water Quality Parameters". In: *ECAI 2010 - 19th European Conference on Artificial Intelligence*. Ed. by H. Coelho, R. Studer and M. Wooldridge. IOS Press, 2010, pp. 283-288.
CORE rank: A
(*3 citations*) (*document*)

[41] L. Torgo, W. Pereira, and C. Soares. "Detecting Errors in Foreign Trade Transactions: Dealing with Insufficient Data". In: *14th Portuguese Conference on Artificial Intelligence, EPIA 2009*. Ed. by L. e. al Lopes. LNAI - 5816. Springer, 2009.
CORE rank: B
(*3 citations*) (*document*)

[42] L. Torgo and R. P. Ribeiro. "Precision and Recall for Regression.". In: *Discovery Science*. Ed. by J. Gama, V. S. Costa, A. M. Jorge and P. Brazdil. Vol. 5808. Lecture Notes in Computer Science. Springer, 2009, pp. 332-346.
CORE rank: None
(*41 citations*) (*document*)

[43] R. Ribeiro and L. Torgo. "Utility-based performance measures for regression". In: *Proceedings of the 3rd Workshop on Evaluation Methods for Machine Learning, in conjunction with the 25th International Conference on Machine Learning (ICML 2008)*. 2008.
CORE rank: A*
(*2 citations*) (*document*)

[44] L. Torgo. "Resource-bounded Fraud Detection". In: *Progress In Artificial Intelligence, Proceedings of the 13th Portuguese Conference on Artificial Intelligence Workshops (EPIA 2007)*. Ed. by J. Neves, M. F. Santos and J. Machado. Vol. 4874. Lecture Notes in Artifical Intelligence. Springer, Dec. 2007, pp. 449-460.
CORE rank: None
(*15 citations*) (*document*)

[45] L. Torgo and R. Ribeiro. "Utility-based Regression". In: *Proceedings of the 11th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD 2007)*. Ed. by K. JN, K. J. de Mántaras RL, S. Matwin, D. Mladenic and A. Skowron. Vol. 4702. Lecture Notes in Artificial Intelligence. Springer, 2007, pp. 597-604.
CORE rank: A
(*42 citations*) (*document*)

[46] R. Ribeiro and L. Torgo. "Rule-based Prediction of Rare Extreme Values". In: *Proceedings of the 9th International Conference on Discovery Science (DS'2006)*. LNAI. Springer, 2006.
CORE rank: None

(*5 citations*) (*document*)

[47] L. Torgo and R. Ribeiro. "Predicting Rare Extreme Values". In: *Proceedings of the 10th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD'2006)*. Ed. by W. Ng. Lecture Notes in Artificial Intelligence 3918. Springer, 2006.
CORE rank: A
(*5 citations*) (*document*)

[48] R. Ribeiro and L. Torgo. "A Comparative Study on Predicting Algae Blooms in River Douro, Portugal". In: *Proceedings of the V European Conference on Ecological Modelling (ECEM-2005)*. 2005.
CORE rank: None
(*19 citations*) (*document*)

[49] L. Torgo. "Regression Error Characteristic Surfaces". In: *Proceedings of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2005)*. Ed. by R. Grossman, R. Bayardo, K. Bennett and J. Vaidya. ACM Press, 2005, pp. 697-702.
CORE rank: A*
(*27 citations*) (*document*)

[50] L. Torgo. "The TNT Financial Trading System: a midterm report". In: *Proceedings of the Workshop on Data Mining for Business at ECML/PKDD 2005*. 2005.
CORE rank: A
(*2 citations*) (*document*)

[51] L. Torgo and J. Marques. "Adapting Peepholing to Regression Trees". In: *Proceedings of the 12th EPIA*. LNAI. Springer, 2005.
CORE rank: B
(*document*)

[52] A. Loureiro, L. Torgo, and C. Soares. "Outlier Detection using Clustering Methods: a data cleaning application". In: *Proceedings of KDNet Symposium on Knowledge-based Systems for the Public Sector*. 2004.
CORE rank: None
(*173 citations*) (*document*)

[53] R. P. Ribeiro and L. Torgo. "Predicting Harmful Algae Blooms.". In: *EPIA*. Ed. by F. Moura-Pires and S. Abreu. Vol. 2902. Lecture Notes in Computer Science. Springer, 2003, pp. 308-312.
CORE rank: B
(*10 citations*) (*document*)

[54] A. C. e Silva, A. Jorge, and L. Torgo. "Automatic Selection of Table Areas in Documents for Information Extraction.". In: *EPIA*. Ed. by F. Moura-Pires and S. Abreu. Vol. 2902. Lecture Notes in Computer Science. Springer, 2003, pp. 460-465.

CORE rank: B
(*8 citations*)

[55] A. Silva, A. Jorge, and L. Torgo. "Selection of Table Areas for Information Extraction". In: *Proceedings of the 3rd International Workshop in Document Analysis and its Applications (DLIA 2003)*. 2003.
CORE rank: None
(*2 citations*) (*document*)

[56] L. Torgo and R. Ribeiro. "Predicting Outliers". In: *Proceedings of Principles of Data Mining and Knowledge Discovery (PKDD'03)*. Ed. by N. Lavrac, D. Gamberger, L. Todorovski and H. Blockeel. 2838 LNAI. Springer, 2003, pp. 447-458.
CORE rank: A
(*19 citations*) (*document*)

[57] L. Torgo. "Computationally Efficient Linear Regression Trees". In: *Classification, Clustering and Data Analysis: recent advances and applications (Proc. of IFCS 2002)*. Ed. by K. Jajuga, A. Sokolowski and H. Bock. Studies in Classification, data analysis, and knowledge organization. Springer, 2002, pp. 409-415.
CORE rank: None
(*14 citations*) (*document*)

[58] P. Almeida and L. Torgo. "The Use of Domain Knowledge in Feature Construction for Financial Time Series Prediction". In: *Proceedings of the Portuguese AI Conference (EPIA'01)*. Ed. by P. Brazdil and A. Jorge. LNAI 2258. Springer, 2001, pp. 116-129.
CORE rank: B
(*4 citations*)

[59] L. Torgo. "A study on end-cut preference in least squares regression trees". In: *Proceedings of the Portuguese AI Conference (EPIA 2001)*. Ed. by P. Brazdil and A. Jorge. LNAI 2258. Springer, 2001, pp. 104-115.
CORE rank: B
(*7 citations*) (*document*)

[60] L. Torgo. "Efficient and Comprehensible Local Regression". In: *Proceedings of the 4th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2000)*. Ed. by Terano, Liu and Chen. LNAI 1805. Springer, 2000, pp. 376-379.
CORE rank: A
(*2 citations*) (*document*) (*extra information*)

[61] L. Torgo. "Partial Linear Trees". In: *Proceedings of the 17th International Conference on Machine Learning (ICML 2000)*. Ed. by P. Langley. Morgan Kaufmann Publishers, 2000, pp. 1007-1014.
CORE rank: A*
(*12 citations*) (*document*)

[62] L. Torgo and J. P. Costa. "Clustered Multivariate Regression". In: *Data*

Analysis, Classification, and Related Methods. Ed. by Kiers, Rasson, Groenen and Schader. Springer, 2000.
CORE rank: None

[63] L. Torgo and J. P. Costa. "Clustered Partial Linear Regression". In: *Proceedings of the Fifth International Workshop on Multistrategy Learning (MSL-2000)*. Ed. by R. Michalski and P. Brazdil. 2000.
CORE rank: None
(*13 citations*) (*document*)

[64] L. Torgo and J. P. Costa. "Clustered Partial Linear Regression". In: *Proceedings of the 11th European Conference on Machine Learning (ECML 2000)*. Ed. by R. Mantaras and E. Plaza. LNAI 1810. Springer, 2000, pp. 426-436.
CORE rank: A
(*9 citations*)

[65] L. Torgo. "Predicting the Density of Algae Communities using Local Regression Trees". In: *Proceedings of the European Congress on Intelligent Techniques and Soft Computing (EUFIT'99)*. 1999.
CORE rank: None
(*9 citations*) (*document*)

[66] J. Gama, L. Torgo, and C. Soares. "Dynamic Discretization of Continuous Attributes.". In: *IBERAMIA*. Ed. by H. Coelho. Vol. 1484. Lecture Notes in Computer Science. Springer, 1998, pp. 160-169.
CORE rank: None
(*41 citations*)

[67] L. Torgo. "A Comparative Study of Reliable Error Estimators for Pruning Regression Trees". In: *Proceedings of the Iberoamericam Conference on AI (IBERAMIA-98)*. Ed. by H. Coelho. 1998.
CORE rank: None
(*16 citations*) (*document*)

[68] L. Torgo. "Error Estimates for Pruning Regression Trees". In: *Proceedings of the 10th European Conference on Machine Learning*. Ed. by C. Nedellec and C. Rouveirol. LNAI 1398. Springer Verlag, 1998.
CORE rank: A
(*14 citations*) (*document*)

[69] L. Torgo. "Functional Models for Regression Tree Leaves". In: *Proceedings of the 14th International Conference on Machine Learning*. Ed. by D. Fisher. Morgan Kaufmann Publishers, 1997.
CORE rank: A*
(*142 citations*) (*document*)

[70] L. Torgo. "Kernel Regression Trees". In: *Poster papers of the European Conference on Machine Learning (ECML-97)*. 1997.

CORE rank: A
(*34 citations*) (*document*)

[71] L. Torgo and J. Gama. "Search-based Class Discretization". In: *Proceedings of the European Conference on Machine Learning (ECML-97)*. LNAI 1224. Springer, 1997.
CORE rank: A
(*32 citations*) (*document*)

[72] L. Torgo and J. Gama. "Regression by Classification". In: *Advances In Artificial Intelligence, Proceedings of the 13th Brazilian Symposium on Artificial Intelligence (SBIA 1996)*. Ed. by D. L. Borges and C. A. A. Kaestner. Vol. 1159. Lecture Notes in Artificial Intelligence. Springer, Oct. 1996, pp. 51-60.
CORE rank: None
(*55 citations*) (*document*)

[73] L. Torgo. "Applying Propositional Learning to Time Series Prediction". In: *Workshop on Statistics, Machine Learning and Knowledge Discovery in Databases*. Ed. by Y. et all Kodratoff. 1995.
CORE rank: None
(*6 citations*) (*document*)

[74] L. Torgo. "Data Fitting with Rule-based Regression". In: *Proceedings of the 2nd International Workshop on Artificial Intelligence Techniques (AIT 1995)*. Ed. by J. Zizka and P. Brazdil. 1995.
CORE rank: None
(*27 citations*) (*document*)

[75] L. Torgo. "Controlled Redundancy in incremental Rule Learning". In: *European Conference on Machine Learning (ECML'93)*. Ed. by P. Brazdil. LNAI 667. Springer-Verlag, 1993, pp. 185-195.
CORE rank: A
(*53 citations*) (*document*)

[76] L. Torgo. "Rule Combination in Inductive Learning". In: *European Conference on Machine Learning (ECML'93)*. Ed. by P. Brazdil. LNAI 667. Springer-Verlag, 1993, pp. 384-389.
CORE rank: A
(*14 citations*) (*document*)

[77] P. Brazdil, M. Gams, S. Sian, et al. "Learning in Distributed Systems and Multi-Agent Environments". In: *Machine Learning: EWSL-91 (European Working Session on Learning)*. Ed. by Y. Kodratoff. Vol. 482. Lecture Notes in Artificial Intelligence. Springer, 1991, pp. 412-423.
CORE rank: None
(*70 citations*) (*document*)

[78] L. Torgo and M. Kubat. "Knowledge Integration and Forgetting". In: *Proceedings of the Czechoslovak Conference on Artificial Intelligence*. 1991.

CORE rank: None
(*4 citations*) (*document*)

Other Publications

[1] L. Torgo. "Uma Breve Introdução à Data Science (in portuguese)". In: *Boletim da APDIO* 57 (2017), pp. 9-11.
(*document*)

[2] C. Magalhães, A. Mucha, F. Carvalho, et al. "Development of an autonomous system for integrated marine monitoring". In: *Proceedings of 41st CIESM Congress*. 2016.
(*document*)

[3] C. Magalhães, J. Séneca, C. Leite, et al. "Distribution and Environmental Controls on Marine Nitrogen Biogeochemical Functions". In: *Proceedings of 41st CIESM Congress*. 2016.
(*document*)

[4] C. Magalhães, C. Lee, M. Monteiro, et al. "Everything is not everywhere: Antarctica Dry Valleys as an extreme counter example". In: *Proceedings of XXXIV SCAR*. 2016.

[5] J. Séneca, C. Magalhães, M. Monteiro, et al. "Distribution of prokaryotic communities and NifH gene diversity in the extrem Darwin Mountains, Abtarctica". In: *Proceedings of XXXIV SCAR*. 2016.

[6] M. Monteiro, J. Séneca, L. Torgo, et al. "The impact of environmental changes on nitrifyng communities from the Dry Valleys of Antarctica". In: *Proceedings of XXXIV SCAR*. 2016.

[7] C. Bartilotti, A. Santos, R. Marques, et al. "Presenting the MarinEye project – Development and validation of a prototype for multitrophic oceanic monitoring". In: *Proceedings of ASC 2016- ICES Anual Science Meeting*. 2016.

[8] C. Magalhães, A. Mucha, F. Carvalho, et al. "Development of an autonomous system for integrated marine monitoring". In: *Proceedings of XIX Iberian Symposium on Marine Biology Studies*. 2016.
(*document*)

[9] J. N. van Rijn, V. Umaashankar, S. Fischer, et al. "A RapidMiner extension for Open Machine Learning". In: *RapidMiner Community Meeting and Conference, 2013*. 2013.
(*7 citations*)

[10] L. Torgo. "Inductive Learning of Tree-based Regression Models". PhD thesis. Faculty of Sciences, University of Porto, 1999.

4.2 Projects

Project Coordination

- [*2018 – 2019*]
  Online Observatory of Economic Activity through Digital Means
  Private funding: ASAE, Ministry of Economy, Portugal - Principal investigator (PI)
  Budget: 68 kEur
- [*2018 – 2021*]
  Nitrolimit - Life at the Edge: Define the Boundaries of the Nitrogen Cycle in the Extreme Antarctic Environments
  Portuguese Science Foundation (FCT), 02/SAICT/2017, Co-Principal investigator (PI) (PI: Catarina Magalhaes)
  Budget: 238.5 kEur
- [*2015 – 2018*]
  Parfois Product/Shop Sales Forecasting for Supporting Logistics Decisions
  Private funding: Parfois - Principal investigator (PI)
  Budget: 100 kEur
- [*2015 – 2016*]
  News Summarizer
  Private funding: SkimIT - Principal investigator (PI)
  Budget: 35 kEur
- [*2011 – 2014*]
  e-Policy, Engineering the POlicy-making LIfe CYcle
  EC 7th Framework Programme Theme ICT-2011-7 (EC), Pr. Nr. 288147 - Local coordinator (PI of Portuguese Partner)
  Budget (local partner): 295 kEur; Global budget: 3 MEur
- [*2008 – 2011*]
  MORWAQ, Monitoring and Predicting Water Quality Parameters
  Portuguese Science Foundation (FCT), PTDC/EIA/68489/2006 - Principal investigator (PI)
  Budget: 47 kEur
- [*2008 – 2011*]
  oRANKI, Resource-bounded Outlier Detection
  Portuguese Science Foundation (FCT), PTDC/EIA/68322/2006 - Principal investigator (PI)
  Budget: 48,5 kEur
- [*2008*]
  Exploratory Analysis of *Sonae Distribuição* Employees Survey

(2008)
MBA Consultores, private funding - Principal investigator (PI)

- [*2007 – 2008*]
  Monitoring water quality parameters
  Águas do Douro e Paiva, SA, private funding - Principal investigator (PI)
- [*2004–2006*]
  MODAL, Models for Predicting Algae Blooms in River Douro
  Portuguese Science Foundation (FCT), POSI/2000/SRI/40949 - Principal investigator (PI)
  Budget: 33 kEur
- [*2005*]
  Exploratory Analysis of *Sonae Distribuição* Employees Survey (2005)
  MBA Consultores, private funding - Principal investigator (PI)
- [*2003 – 2004*]
  Development of a System for Automatic Intraday Trading in Stock Markets
  private funding - Principal investigator (PI)

Project Participation

- [*2016–2018*]
  CORAL - Sustainable Ocean Exploitation: Tools and Sensors
  FEDER, Portugal 2020, Norte 2020 - Leader of the data analysis team
  Global budget: 2.3 MEur
- [*2015–2017*]
  MarinEye - a prototype for multitrophic ocean monitoring
  EEA Grants, project PT02-0037 - leader of the workpackage on data analysis
  Global budget: 373 kEur
- [*2016–2018*]
  Reminds - Relevance Mining and Detection System
  Portuguese Science Foundation - UT Austin/Portugal Program
  Global budget: 187 kEur
- [*2015–2018*]
  FOTOCATGRAF - Graphene-based semiconductor photocatalysis for a safe and sustainable water supply: advanced technology for emerging pollutants removal
  Portuguese Science Foundation - UT Austin/Portugal Program - 2014, project 137424 - leader of the data analysis team
  Global budget: 200 kEur

- [*2014–2015*]
  OpenML
  EC Harvest Pascal Network
- [*2011 – 2012*]
  PRODUTECH-PSI, New Products and Services for the Transformation Industry
  Compete - Portugal 2020
  Global budget: 12.5 MEur
- [*2008 – 2011*]
  Rank!, Development of methods for predicting item scheduling
  Portuguese Science Foundation (FCT), PTDC/EIA/81178/2006
- [*2000 – 2005*]
  Sol-Eu-Net
  European Community (EC), IST-1999-11495
  Global budget: 3 MEur
- [*2000 – 2001*]
  Tsam, Knowledge extraction from financial time series for risk managment
  Portuguese Science Foundation (FCT), POSI/SRI/34329/99
- [*1998 – 2002*]
  METAL, A Meta-Learning Assistant for Providing User Support in Machine Learning and Data Mining
  European Community (EC), ESPRIT 26.357
- [*1997 – 2000*]
  ECO, Knowledge Extraction from Databases
  Portuguese Science Foundation (FCT), Praxis XXI
- [*1991 – 1994*]
  Statlog
  European Community (EC), Esprit Project 5170
- [*1989 – 1992*]
  ECOLES
  European Community (EC), Esprit II 3059

4.3 Prizes

- [*2017*]
  Co-author (author was Vitor Cerqueira) of the paper that won the *Best Student Machine Learning Paper Award* given by the Machine Learning Journal at the European Conference on Machine Learnig (ECML/PKDD'2017)

- [*2017*]
  Supervisor of the PhD thesis of Nuno Moniz entitled "Prediction and Ranking of Highly Popular Web Content" that was awarded the 2nd place in the Fraunhofer Portugal Challenge 2017 competition in the category of PhD theses
- [*2006*]
  Co-author (author was Rita Ribeiro) of the paper that won the *Best Student Paper Award* given by Yahoo! Research Labs at the Discovery Science (DS'06) international conference
- [*1999*]
  Runner-up winner at the 3rd International Competition "Protecting rivers and streams by monitoring chemical concentrations and algae communities", organized by ERUDIT in conjunction with COIL, the cluster of four European Research Networks (ERUDIT, EvoNet, MLNet and NeuroNet)

4.4 Thesis Supervision

Post-Doctoral Fellows

Finished

- Colin Bellinger
  - Title: Class Imbalance and Learning from Rare Cases
  - 2018
- Paula Branco
  - Title: Utility-based Predictive Analytics
  - 2019

Ph.D.'s

Ongoing

- Mohammad Etemad
  - Title: Segmentation Algorithms for Trajectory Data
  - Supervisor: Stan Matwin; co-supervisor: Luis Torgo
  - PhD on Computer Science, Faculty of Computer Science, Dalhousie University
  - Start: 2019
- Mariana Oliveira
  - Title: Predictive Analytics for Dependent Data
  - Supervisor: Luis Torgo; Co-Supervisor: Vitor Santos Costa
  - MAPi PhD Program, Universities of Aveiro, Minho and Porto
  - Start: 2017
- Nuno Guimarães

  - Title: Analyzing and Developing Veracity Indicators for Building an Automatic Detector of False Online News
  - Supervisor: Álvaro Figueira; co-supervisor: Luis Torgo
  - PhD on Computer Science, Faculty of Sciences, University of Porto
  - Start: 2017

Finished

- Vitor Cerqueira
  - Title: Ensembles for Time Series Forecasting
  - Supervisor: Luis Torgo; Co-Supervisor: Carlos Soares
  - PhD Program Faculty of Engineering/UPorto
  - December/2019
- Paula Branco
  - Title: Utility-based Predictive Analytics
  - Supervisor: Luis Torgo; Co-Supervisor: Rita Ribeiro
  - MAPi PhD Program, Universities of Aveiro, Minho and Porto
  - September/2018
- Nuno Moniz
  - Title: Prediction and Ranking of Highly Popular Web Content
  - Supervisor: Luis Torgo
  - PhD on Computer Science, Faculty of Sciences, University of Porto
  - July/2017
  - Awarded the 2nd place in the Fraunhofer Portugal Challange
- Brett Drury
  - Title: A Text Mining System for Evaluating the Stock Market's Response To News
  - Supervisor: Luis Torgo; Co-Supervisor: José João Almeida (Univ. Minho)
  - MAPi PhD Program, Universities of Aveiro, Minho and Porto
  - April/2013
- Orlando Ohashi
  - Title: Spatio-Temporal Prediction Methods
  - Supervisor: Luis Torgo
  - MAPi PhD Program, Universities of Aveiro, Minho and Porto
  - December/2012
- Rita Ribeiro
  - Title: Utility-based Regression
  - Supervisor: Luis Torgo
  - PhD on Computer Science, Faculty of Sciences, University of Porto
  - September/2011
- Pedro Almeida
  - Title: Previsão do Comportamento de Séries Temporais Financeiras com Apoio de Conhecimento Sobre o Domínio
  - Supervisor: Luis Torgo
  - Doutoramento em Engenharia Informática, Universidade da Beira Interior

- April/2003

M.Sc.'s

Finished

- Inês Areosa
  - Title: Visual Tools for Understanding Regression Performance
  - Supervisor: Catarina Magalhães; Co-Supervisors: Pedro Duarte and Luis Torgo
  - Master of Science Degree in Aerospace Engineering, IST, Portugal
  - Nov/2019
- Antonio Gaspar Goncalves de Sousa
  - Title: Arctic microbiome and N-functions during the winter-spring transition
  - Supervisor: Catarina Magalhães; Co-Supervisors: Pedro Duarte and Luis Torgo
  - Masters on Molecular and Celular Biology, ICBAS, University of Porto
  - Nov/2017
- Carlos Leite
  - Title: Domain Oriented Biclustering Validation
  - Supervisor: Luis Torgo; Co-Supervisor: Catarina Magalhães
  - Masters on Computer Science, Faculty of Sciences, University of Porto
  - 30/Nov/2016
  - Grade: 19 out of 20
- Nuno Guimarães
  - Title: Lexicon Expansion System for Domain and Time Oriented Sentiment Analysis
  - Supervisor: Luis Torgo; Co-Supervisor: Álvaro Figueira
  - Masters on Computer Science, Faculty of Sciences, University of Porto
  - 28/Nov/2016
  - Grade: 18 out of 20
- Mariana Oliveira
  - Title: Propositional and Relational Approaches to Spatio-Temporal Data Analysis
  - Supervisor: Luis Torgo; Co-Supervisor: Vitor Santos Costa
  - Masters on Computer Science, Faculty of Sciences, University of Porto
  - October/2015
  - Grade: 20 out of 20
- Luís Baía
  - Title: Actionable Forecasting and Activity Monitoring: applications to financial trading
  - Supervisor: Luis Torgo
  - Masters in Engineering Mathematics, Faculty of Sciences, University

of Porto
  - August/2015
  - Grade: 20 out of 20
- Paula Branco
  - Title: Re-sampling Approaches for Regression Tasks under Imbalanced Domains
  - Supervisor: Luis Torgo; Co-supervisor: Rita Ribeiro
  - Masters in Computer Science, Faculty of Sciences, University of Porto
  - September/2014
  - Grade: 19 out of 20
- Fernando Correia
  - Title: SunPet – Real-time Sun Exposure Monitorization using Smartphones
  - Supervisor: Luís Rosado (Fraunhofer AICOS); Co-Supervisor: Luis Torgo
  - Masters in Network and Information Systems Engineering, Faculty of Sciences, University of Porto
  - 2014
- João Cepêda
  - Title: Telecommunication Fraud Detection Using Data Mining techniques
  - Supervisor: Carlos Soares (FEUP/UPorto); Co-Supervisor: Luis Torgo
  - Master in Electrical and Computers Engineering, Faculty of Engineering, University of Porto
  - June/2014
- Pedro Coelho
  - Title: Multi-Topic Sentiment Analysis
  - Supervisor: Luis Torgo
  - Masters in Computer Science, Faculty of Sciences, University of Porto
  - 2013
- Hélia Costa
  - Title: Estudo comparativo de abordagens ao problema de débito de transações bancárias em contas com saldo insuficiente
  - Supervisor: Luis Torgo
  - Masters in Engineering Mathematics, Faculty of Sciences, University of Porto
  - September/2012
- Raquel Santos
  - Title: Modelos de Regressão para a Previsão de Vendas e de Clientes
  - Supervisor: Luis Torgo; Co-Supervisor: Luis Marques (SONAE)
  - Masters in Engineering Mathematics, Faculty of Sciences, University of Porto
  - 2010
- Pedro Duarte
  - Title: Service-Oriented Architectures

  - Supervisor: Paulo Martins (Critical); Co-Supervisor: Luis Torgo
  - Masters in Network and Information Systems Engineering, Faculty of Sciences, University of Porto
  - 2010
- Clara Gonçalves
  - Title: Modelos de Regressão com Análise Classificatória
  - Supervisor: Joaquim Pinto da Costa ; Co-Supervisor: Luis Torgo
  - Masters in Engineering Mathematics, Faculty of Sciences, University of Porto
  - 2005
- Jorge Barbosa
  - Title: Métodos para lidar com Mudanças de Regime em Séries Temporais Financeiras
  - Supervisor: Luis Torgo
  - Master in Data Analysis and Decision Support Systems, Faculty of Economics, University of Porto
  - 2005
- Joana Marques
  - Title: Um estudo sobre a eficiência computacional da construção de árvores de regressão
  - Supervisor: Luis Torgo
  - Masters in Artificial Intelligence and Computation, Faculty of Economics, University of Porto
  - 2004
- Rita Ribeiro
  - Title: Modelos de Previsão de Fenómenos Raros
  - Supervisor: Luis Torgo
  - Masters in Artificial Intelligence and Computation, Faculty of Economics, University of Porto
  - 2003
- Ana Silva
  - Title: Extracção da Informação de Tabelas Contidas em Texto - uma aplicação a Relatórios de Contas em Empresas Portuguesas
  - Supervisor: Alipio Jorge; Co-supervisor: Luis Torgo
  - Master in Data Analysis and Decision Support Systems, Faculty of Economics, University of Porto
  - 2002
- Mário Oldemiro
  - Title: Técnicas de Inteligência Artificial Aplicadas à Previsão de Séries Temporais Financeiras
  - Supervisor: Luis Torgo; Co-supervisor: Pavel Brazdil
  - Master in Data Analysis and Decision Support Systems, Faculty of Economics, University of Porto
  - 2002
- César Rocha
  - Title: Algoritmo Recursivo dos Mínimos Quadrados para Regressão

Linear Local
  - Supervisor: Luis Torgo
  - Masters in Statistics, Faculty of Sciences, University of Porto
  - 2001
- Sílvia Amorim
  - Title: A escolha do número de classes no método de classificação das k-Médias
  - Supervisor: Joaquim Pinto da Costa; Co-supervisor: Luis Torgo
  - Masters in Statistics, Faculty of Sciences, University of Porto
  - 2001

4.5 Organization of Events

- [*2018*]
2nd International Workshop on Learning with Imbalanced Domains: Theory and Applications, European Conference on Machine Learning, ECML'2018, workshop co-chair
- [*2018*]
International Workshop on Cost Sensitive Learning, to take place at SIAM International Conference on Data Mining, San Diego, USA, 3-5 May 2018, workshop co-chair
- [*2017*]
1st International Workshop on Learning with Imbalanced Domains: Theory and Applications, European Conference on Machine Learning, ECML'2017, workshop co-chair
- [*2015*]
25th European Conference on Machine Learning, ECML'2015, workshop chair
- [*2008*]
18th COMPSTAT Symposium of the IASC-ERS, COMPSTAT'08, local organizing committee member
- [*2005*]
9th European Conference on Principles and Practice of Knowledge Discovery, PKDD'2005, program co-chair, local organization committee member and webmaster
- [*2005*]
16th European Conference on Machine Learning, ECML'2005, local organization committee member and webmaster
- [*2003*]
14th European Conference on Machine Learning, ECML'2003, workshop chair

- [*2003*]
  7th European Conference on Principles and Practice of Knowledge Discovery, PKDD'2003
- [*2003*]
  International Workshop on Data Mining and Adaptive Modelling Methods for Economics and Managment, local organization committee member
- [*2001*]
  Workshop on Artificial Intelligence for Financial Time Series Analysis, program Chair, local organization committee member, and webmaster

4.6 Scientific Reviewing

Academic Juris

- [2017]
  Ph.D. Thesis of Pedro Saleiro - *Entity-Specific Text Mining for Online Reputation Monitoring*, University of Porto, Portugal
- [2017]
  Ph.D. Thesis of Davi D'Andréa Baccan - *Contributions of Computational Cognitive Modeling to the Understanding of the Financial Markets*, University of Coimbra, Portugal
- [2015]
  Ph.D. Thesis of Vinay Uday Prabhu - *Network Aided Classification and Detection of Data*, Carnegie Mellon University / MAP-I Doctoral Program, Pittsburgh, USA
- [2013]
  Ph.D. Thesis of Ricardo Nuno Taborda Campos - *Disambiguating Implicit Temporal Queries for Temporal Information Retrieval Applications*, PhD on Computer Science, Faculty of Sciences, University of Porto, Portugal
- [2012]
  Ph.D. Thesis of Nuno Constantino Castro - *Time Series Motif Discovery*, MAP-I Doctoral Program, University of Minho, Portugal
- [2012]
  MSc. Thesis of Nuno Moniz - *Bridging the gap between closed and open data, System proposal for the Portuguese Legislation*, Masters on Computer Engineering, specialization in Networks, Architectures and Systems , ISEP, Portugal

- [2011]
  Ph.D. Thesis of Rui Barbosa - *Agents in the Market Place*, University of Minho, Portugal
- [2008]
  Ph.D. Thesis of Pedro Rafael de Ruiz Graça - *Aprendizagem Interactiva em Sistemas Multi-Agente*, University of Lisbon, Portugal
- [2008]
  Ph.D. Thesis of Anneleen Van Assche - *Improving the Applicability of Ensemble methods in Data Mining*, Katholieke Universiteit Leuven, Belgium
- [2007]
  Ph.D. Thesis of Pedro Gabriel Dias Ferreira - *Sequence Pattern Mining in Biochemical Data*, University of Minho, Portugal
- [2005]
  Ph.D. Thesis of Kwok Pan Pang - *Statistics for Structural Break Detection and Their Application to Forecasting and Statistical Process Control*, Monash University, Australia
- [2004]
  MSc. Thesis of Susana Pereira - *Análise de Séries Temporais no Domínio das Telecomunicações Móveis*, Masters on Statistics and Information Managment, ISEGI, New University of Lisbon, Portugal
- [2003]
  Ph.D. Thesis of Vitor Lobo - *Ship Noise Classification, a contribution to prototype based classifier design*, New University of Lisbon, Portugal
- [2003]
  MSc. Thesis of Raul Moisão - *Modelo Predictivo, Baseado em Redes Neuronais, para Previsão em Séries Temporais com Origem em Sistemas Financeiros*,New University of Lisbon, Portugal

Research projects

- [2017]
  European Comission, Expert Reviewer (EX2017D300375), Review of 4 proposals for the H2020-MSCA-IF-2017 (Horizon 2020 Marie Sklodowska-Curie Actions - Individual Fellowships)
- [2016]
  KU Leuven, Belgium. *Review of one project proposal.*
- [2011 – 2014]
  FIRST - Large scale information extraction and integration infrastructure for supporting financial decision making. EC Seventh Framework Programme, project nr. 257928. *Member of the Advisory Board.*

- [2013]
  Czech Science Foundation – GACR. *Review of one project proposal.*
- [2011]
  Czech Science Foundation – GACR. *Review of two projects proposals.*
- [2010]
  Czech Science Foundation – GACR. *Review of one project proposal.*

Editorial Boards of Journals

- *Intelligent Data Analysis*, IOS Press.
  Member of the editorial board.

Journals

- *Data Mining and Knowledge Discovery*, Springer.
  Reviewing of 8 submissions.
- *Journal of Machine Learning Research.*
  Reviewing of 2 submissions.
- *Machine Learning Journal*, Kluwer Academic Publishers.
  Reviewing of 11 submissions.
- *IEEE Transactions on Knowledge and Data Engineering*, IEEE Computer Society.
  Reviewing of 2 submissions.
- *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE Computer Society.
  Reviewing of 1 submission.
- *Journal of Artificial Intelligence Research*, Morgan Kaufmann.
  Reviewing of 1 submission.
- *Decision Support Systems*, Elsevier.
  Reviewing of 2 submissions.
- *Neural Computing and Apllications*, Springer.
  Reviewing of 1 submission.
- *Neural Networks*, Elsevier.
  Reviewing of 1 submission.
- *Intelligent Data Analysis*, Elsevier Science.
  Reviewing of 1 submission.
- *Expert Systems*, Wiley.
  Reviewing of 1 submission.

- *International Journal of Human-Computer Studies*, Elsevier Science.
  Reviewing of 1 submission.
- *AI Communications*, IOS Press.
  Reviewing of 1 submission.
- *International Journal on Artificial Intelligence Tools*, World Scientific.
  Reviewing of 2 submission.

International Conferences

- *KDD, ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*
  2009 (PC member), 2007 (PC member)
- *ICML, International Conference on Machine Learning*
  2016 (PC member), 2011 (PC member), 2010 (PC member), 2009 (Area Chair), 2008 (Area Chair), 2004 (PC member)
- *ECML, European Conference on Machine Learning*
  2017 (Area Chair), 2016 (PC member), 2015 (Area Chair), 2014 (PC member), 2013 (PC member), 2012 (Area Chair), 2011 (Area Chair), 2010 (PC member), 2009 (PC member), 2008 (PC member), 2007 (Area Chair), 2006 (Area Chair), 2005 (Area Chair), 2004 (PC member), 2003 (PC member)
- *ICDM, IEEE International Conference on Data Mining*
  2007 (PC member), 2006 (PC member), 2005 (PC member), 2004 (PC member)
- *PKDD, European Conference on Principles and Practice of Knowledge Discovery in Databases*
  2017 (Area Chair), 2016 (PC member), 2015 (Area Chair), 2014 (PC member), 2013 (PC member), 2012 (Area Chair), 2011 (Area Chair), 2010 (PC member), 2009 (PC member), 2008 (PC member), 2007 (Area Chair), 2006 (Area Chair), 2005 (Program Chair), 2004 (PC member), 2003 (PC member)
- *PAKDD, Pacific-Asia Conference on Knowledge Discovery and Data Mining*
  2009 (PC member), 2008 (PC member), 2007 (PC member)
- *DS, International Conference on Discovery Science*
  2010 (PC member), 2009 (PC member), 2008 (PC member), 2007 (PC member)
- *AAAI, Conference on Artificial Intelligence*
  2017 (PC member), 2015 (PC member), 2014 (PC member)

- *IJCAI, International Joint Conference on Artificial Intelligence*
  2017 (PC member), 2013 (Senior PC member), 2011 (Senior PC member)
- *ECAI, European Conference on Artificial Intelligence*
  2014 (Senior PC member), 2012 (PC member), 2010 (PC member)
- *NIPS, Annual Conference on Neural Information Processing Systems*
  2014 (PC member)
- *ACML, Asian Conference on Machine Learning*
  2012 (PC member)
- *UseR, The R User Conference*
  2013 (PC member)
- *SAC, ACM Symposium on Applied Computing*
  2005 (PC member)
- *ADMA, International Conference on Advanced Data Mining and Applications*
  2008 (PC member), 2007 (PC member), 2006 (PC member), 2005 (PC member)
- *IBERAMIA, Iberoamerican Conference on Artificial Intelligence*
  2002 (PC member), 2000 (PC member), 1998 (PC member)
- *EPIA, Portuguese AI Conference*
  2005 (PC member), 2003 (PC member), 2001 (PC member)
- *SBIA, Brazilian Symposium on Artificial Intelligence*
  2008 (PC member), 2004 (PC member)
- *ENIA, Brazilian Meeting on Artificial Intelligence*
  2011 (PC member), 2007 (PC member)

4.7 Software

- [2016]
  DMwR2 - an R package with functions and data for the 2nd edition of "Data Mining with R"
  GitHub project page
- [2016, co-author]
  UBL - an R package for utility-based predictive analytics
  GitHub project page
- [2014]
  performanceEstimation - an R package for estimating the performance of predictive models
  GitHub project page

- [2010]
  DMwR - an R package with functions and data for the 1st edition of "Data Mining with R"
- [2005]
  TNT - an autonomous trading system for financial markets
- [2001]
  CLRT - Clustered regression models
- [1999]
  RT - Tree-based regression models
- [1997]
  C library for propositional learning (in conjuntion with João Gama)
- [1996]
  RECLA - Regression through classification
- [1996]
  KERTI - Kernel regression trees
- [1996]
  EcoTerme - Calculus of thermic behaviour of buildings
- [1995]
  R^2 - regression rules learner
- [1995]
  YAP-Prolog library for propositional learning
- [1993]
  YAILS - Incremental learning of classification rules
- [1991]
  INTEG - Knowledge integration system

4.8 Visits to Research Labs

- [Jan/2018] One week visit to the Jozef Stefan Institute, Ljubljana, Slovenia. Host: Prof. Igor Mozetic
- [Jan/2017] One week visit to the Jozef Stefan Institute, Ljubljana, Slovenia. Host: Prof. Igor Mozetic
- [Sep-Dec/2015]
  Four months visit to the Weka research lab at University of Waikato, New Zealand. Host: Prof. Bernhard Pfahringer.
- [July/2012]
  Two weeks visit to the Text Analysis and Machine Learning (TAMALE) research lab at University of Ottawa, Canada. Host: Prof. Stan Matwin.

- [May/2012]
  Two weeks visit to the Weka research lab at University of Waikato, New Zealand. Host: Prof. Bernhard Pfahringer.
- [Jun/2010]
  One month visit to the Department of Informatics, University degli Studi di Bari, Italy. Host: Prof. Donato Malerba
- [Fev-Jul/2008]
  Six months visit to the Weka research lab at University of Waikato, New Zealand. Host: Prof. Bernhard Pfahringer.
- [Mar-Aug/2004]
  Five months visit to the Stern Business School of the University of New York. Hosts: Profs. Foster Provost and Vasant Dhar
- [1994]
  Three months visit to the University of São Paulo, campus São Carlos, Brazil. Host: Prof. Carolina Monard

4.9 Invited Seminars

- [Sep/2019] *Adressing the Data Revolution]*, Engineers Nova Scotia Annual Conference, Halifax, Canada.
- [Jun/2018] *Predictive Analytics and the Ocean*, H2O Conference, Oceans Week, Halifax, Canada.
- [Jan/2018] *Arbitrage of Forecasting Experts*, Jozef Stefan Institute, Slovenia.
- [July/2017] *Data Pre-processing Methods for Forecasting with Spatio-Temporal Data*, invited talk at the international conference Data Science, Statistics and Visualization, Lisbon, Portugal
- [June/2017] *Handling Imbalanced Regression Tasks through Utility- based Regression*, invited seminar at Université de Fribourg, Fribourg, Switzerland
- [May/2017] *An Infra-Structure for Performance Estimation and Experimental Comparison of Predictive Models in R*, invited talk at SER, Niteroi, Rio de Janeiro, Brazil
- [Jan/2017] *Resampling Approaches for Handling Imbalanced Regression Tasks*, Jozef Stefan Institute, Slovenia.
- [Dec/2015]
  An Infra-Structure for Performance Estimation and Experimental Comparison of Predictive Models in R, University of Waikato, New Zealand
- [Sep/2015]
  Feature Engineering for Handling Spatial and Spatio-Temporal Forecasting, University of Waikato, New Zealand

- [Jun/2015]
The R Language - programming for data analysis, Join 2015, Braga, Portugal
- [May/2015]
Data Mining aplicado à Previsão de Blooms de Algas, Workshop Aquacultura 2015, Porto, Portugal
- [May/2015]
Data Science - what, why and how?, Porto Tech Hub, Porto, Portugal
- [Mar/2015]
An Infra-Structure for Performance Estimation and Experimental Comparison of Predictive Models in R, Porto R Users Group (PRUG), Porto, Portugal
- [Jan/2015]
An Infra-Structure for Performance Estimation and Experimental Comparison of Predictive Models in R, LIAAD Seminars, INESC Tec, Portugal
- [Nov/2014]
Monitoring and Forecasting Rare Events, Workshop INESC/CIIMAR, INESC Tec, Portugal
- [Jun/2014]
Dynamic Documents in R, DCC talks, FCUP/UPorto, Portugal
- [Jan/2014]
Spatio-temporal data mining and extreme behavior data mining, LIAAD Open-Day, INESC Tec, Portugal
- [Jul/2013]
Spatial Interpolation using Multiple Regression, University of Konstanz, Germany
- [Feb/2013]
Data Mining para a Deteção de Fraude, INESC Tec, Portugal
- [May/2012]
Modeling Deviations from Expected Behavior - two case studies, University of Waikato, New Zealand
- [Out/2011]
Modeling Deviations from Expected Behavior - two case studies, Back2Basics Seminar Series, Faculty of Engineering, University of Porto, Portugal
- [Out/2011]
Modeling Deviations from Expected Behavior - two case studies, Thought Leader Speaker Series, eBay Research Labs, San Jose, USA

- [Jul/2011]
Modelos de Previsão para Sistemas Dinâmicos Complexos, Seminários em Engenharia de Sistemas, University of Minho, Portugal
- [Jun/2010]
Resource-bounded Outlier Detection using Clustering Methods, Department of Informatics, University degli Studi di Bari, Italy.
- [Jan/2009]
Using Data Mining for Resource-aware Fraud Detection, Workshop on Data Mining for the Banking System, Faculty of Economics, University of Porto, Portugal
- [Mar/2008]
Utility-based Regression - recent developments, University of Waikato, New Zealand
- [Jan/2008]
Utility-based Regression - recent developments, Katholieke Universiteit Leuven, Belgium
- [Jan/2007]
Predicting Rare Extreme Values - recent developments, Solomon Seminars, Josef Stefan Institute, Slovenia
- [Jan/2006]
Non-Uniform Cost Surfaces for Predicting Rare Extreme Values, Solomon Seminars, Josef Stefan Institute, Slovenia
- [Jan/2006]
Regression Error Characteristic Surfaces, Solomon Seminars, Josef Stefan Institute, Slovenia
- [Sep/2004]
An autonomous trading system, International Summer School on Data Analysis, Instituto Superior de Gestão, Portugal
- [Jan/2004]
An intraday Autonomous Trading System, Faculty of Economics, University of Porto, Portugal
- [Nov/2003]
Mining DNA microarray data: techniques and applications, Instituto de Biologia Molecular e Celular, Porto, Portugal
- [Jun/2003]
Models for Predicting Water Quality, Jornadas em Informática (JOIN'03), University of Minho, Portugal
- [Out/2002]
Artificial Intelligence: from fiction to reality, 2 Ciclo de Conferências em Cibercultura, Guarda, Portugal

4.10 Service to the Community

- [since 1996]
  Repository of Regression Data Sets.

5 Management Activities

- [2014 – . . .]
  Founder and CEO of KNOYDA, a company devoted to training and consulting in data science
- [Sep/2013 – Sep/2015]
  Director of the Integrated Master's program on Network and Information Systems Engineering, Faculty of Sciences, University of Porto
- [Apr/2010 – Apr/2012]
  Director of the Integrated Master's program on Network and Information Systems Engineering, Faculty of Sciences, University of Porto
- [Nov/2010 – Sep/2011]
  Member of the commission responsible for creating the Evaluation Regulation of the teaching staff of the Faculty of Sciences of the University of Porto
- [2008–2011]
  Member of board of the Observatory of Economy and Management of Fraud
- [2008]
  Founding partner of the Observatory of Economy and Management of Fraud
- [2008]
  Member of the commission responsible for the creation of a new Post-Graduation course at the University of Porto Business School, entitled Fraud Management.
- [2005–2008]
  Member of the ECML-PKDD Steering Committee
- [2005–2009]
  Webmaster of the Faculty of Economics of the University of Porto,
- [2005 – . . .]
  Webmaster of the Modelling Dynamic Systems subgroup of interest of LIAAD
- [2001-2005]
  Member of the commission responsible for preparing the proposal for a

new degree at the Faculty of Economics of the University of Porto on Economics and Management of Information.

- [1999–2004]
  Organizer (together with Mário Florido and Luís Paulo Reis) of the Seminars of LIACC
- [1999–2005]
  Member of the Directive Board of the Portuguese Association of Artificial Intelligence (APPIA)

6 Other

General Public Publications

- [Dec/2017] Invited Article (in Portuguese) on the *Boletim da APDIO*, vol. 57, entitled *Uma Breve Introdução à Data Science*
- [Jun/2012]
  Article on the *Visão* magazine entitled *Coitadinho do Doente* in the section *Silêncio da Fraude*
- [Feb/2012]
  Article on the *Visão* magazine entitled *A Fraude das Patentes de Software* in the section *Silêncio da Fraude*

Affiliations

- Member of the Association for Computing Machinery (ACM)
- Member of the International Machine Learning Society
- Member of the Portuguese Association of Artificial Intelligence

Contacts

- *Institutional Address*:
  Goldberg Computer Science Building
  6050 University Avenue
  Halifax NS
  B3H 1W5
  Canada

- *Email contacts*:
  ltorgo *(at)* dal *(dot)* ca
  ltorgo *(at)* dcc *(dot)* fc *(dot)* up *(dot)* pt
  ltorgo *(at)* inesctec *(dot)* pt
- *Phone*:
  +1 902 494 2845
- *Web page*:
  https://web.cs.dal.ca/~ltorgo/