

Dr. Jaume Bacardit
Curriculum Vitae
September 2017

Personal information

- Full name: Jaume Bacardit Penarroya
- Nationality: British/Spanish
- Year of birth: 1976
- Email: jaume.bacardit@newcastle.ac.uk
- Home page: <http://homepages.cs.ncl.ac.uk/jaume.bacardit/>
- Current affiliation: Reader in Machine Learning, School of Computing, Newcastle University

Academic qualifications

- 2004 PhD in Computer Science. Ramon Llull University, Barcelona, Spain. Thesis title: "Pittsburgh Genetics-Based Machine Learning in the Data Mining Era: Representations, generalization, and run-time". Grade: Summa Cum Laude. Citations in Google Scholar: 119.
- 2000 MEng in Computer Engineering. Ramon Llull University, Barcelona, Spain. Grade: A
- 1998 BEng in Computer Engineering. Ramon Llull University, Barcelona, Spain. Grade: B+

Employment history

- Since August 2017. Reader in Machine Learning at Newcastle University
- January 2014-July 2017 Senior Lecturer in Biodata Mining at Newcastle University
- August 2012-December 2013. Lecturer in Bioinformatics at the University of Nottingham. Permanent post at the School of Computer Science.
- January 2008-July 2012. Lecturer in Bioinformatics at the University of Nottingham. Five-years fixed term post jointly appointed between the schools of Computer Science and Biosciences
- March 2005-December 2007. Research Fellow at the University of Nottingham. Project: Robust Prediction with Explanatory Power for Protein Structure and Related Prediction Problems. Sponsor: Engineering and Physical Sciences Research Council of UK (grant GR/T07534/01)

Funding (as principal investigator or co-investigator)

Year	Title	Investigators	Awarding Body	Amount Awarded
2009	Bioinformatics, Systems and Synthetic Biology Plant Science Summer School (ESF-2482)	Prof. N. Krasnogor (PI), Dr. Jaume Bacardit, Prof. M. Bennett	European Science Foundation	€60435
2010	BBSRC CASE Studentship "Bioinformatics analysis of proteomic data from <i>in vitro</i> models of articular cartilage inflammation"	Dr. J. Bacardit (Main supervisor), Dr. A. Mobasher, Dr. D. Allaway	UK BBSRC/Waltham Centre for Pet Nutrition	£74410 from BBSRC, £20000 from Waltham
2010	Principled Application of Learning Classifier Systems to Large-Scale Challenging Datasets (LCSxLCD) - EP/H016597/1	Dr. Jaume Bacardit (PI)	UK EPSRC First Grant	£101,458
2011	CADMAD - Paving the Way for Future Emerging DNA-based Technologies: Computer-Aided Design and Manufacturing of DNA libraries	Newcastle investigators: Prof N. Krasnogor (PI), Dr. J. Bacardit (CI)	EU FP7	€3056308.
2012	D-BOARD - Novel Diagnostics and Biomarkers for Early Identification of Chronic	PI: Prof. A. Mobasher (University of Surrey) CI for Newcastle: Prof.	EU FP7	€5996451

	Inflammatory Joint Diseases	J. Loughlin, Dr. J. Bacardit, Dr. L. Reynard		
2015	apprOAch: Applied Public-Private Research enabling OsteoArthritis Clinical Headway	PI: Harrie Weinans (University Medical centre Utrecht) CI for Newcastle: Prof. J. Loughlin, Dr. J. Bacardit	EU Innovative Medicine Initiative	€15M (€7.5M from IMI and €7.5M provided as In-kind support by GSK, Merck-Serono and Servier)
2015	CRITiCaL - Combatting cRiminals In The Cloud	PI: Thomas Gross. CI: Prof. A. van Moorsel, Prof. P. Briggs, Dr. J. Bacardit, Dr. J. Yan, Dr. GE. Oxburgh, Dr. T. Ploetz	UK EPSRC (Contrails Call)	£2M
2016	Synthetic Portabolomics: Leading the way at the crossroads of the Digital and the Bio Economies	PI: Prof. N. Krasnogor. CI: P. Lord, A. Wipat, P. Zuliani, Y. Yuzenkova, R. Daniel, H. Murray, M. Kaiser, J. Bacardit, S. Woods, N. Zenkin	UK EPSRC Programme Grant	£4.3M from EPSRC, £2.5M Institutional, £0.5M Industrial support
2016	Knowledge Transfer Partnership with Orchard Information Systems Ltd	PI: Dr. S.Y. Coleman. CI: Dr. J. Bacardit	Innovate UK	£140,203.00
2017	Knowledge Transfer Partnership with Rain Data Ltd	PI: Dr. S.Y. Coleman. CI: Dr. J. Bacardit	Innovate UK	£132,522.00

Teaching, supervision and administrative duties

- **PhD supervisions:** Maria Franco (2009-2012), Anna Swan (2011-2014), Nicola Lazzarini (2013-2016), Ossama Sameer Hamed Alshabrawy (2017-), Jake Cowton (2017-)
- **Postdoctoral supervision:** Xiaolei Xia (2010-2011), Pawel Widera (2015-), Bart Craenen (2016), Samuel Danso (2017-), Stephen Matthews (2017-)
- **Visiting students:** Julian Luengo, University of Granada, May-July 2010 // Nuria Macia, Ramon Llull University, July-October 2010 // Alvaro Garcia, Ramon Llull University, September 2010-March 2011 // Maria Martinez, University of Seville June-September 2011 // Alfonso Marquez, Pablo de Olavide University, September-December 2011 // Gualberto Asencio, Pablo de Olavide University, November 2011-February 2012 // Andreu Sancho, Ramon Llull University, April-July 2012 // Isaac Triguero, University of Granada, July-September 2013, Pablo Gutierrez Perez, University of Granada, April-June 2014, Ismael Rodriguez Fernandez, University of Santiago de Compostela, May-July 2014, Juan Carlos Gámez Granados, University of Cordoba, June-September 2016
- **PhD examiner for external students:** David Howard, University of the West of England, July 2010, Julian Luengo, University of Granada, January 2011, Nuria Macia, Ramon Llull University, October 2011, Alvaro Garcia, Ramon Llull University, April 2012, Jose Garcia Moreno-Torres, University of Granada, March 2013, Muhammad Iqbal, Victoria University of Wellington, January 2014, Javier Perez Rodriguez, University of Granada, March 2015, Sol Lim, Seoul National University, May 2015, Fabio Fabris, University of Kent, March 2017.
- **Teaching:** Bioinformatics Theory and Practice (2014-present), Bio-algorithms (2014-present), Biologically-Inspired Computing (2014-present), Comparative and Evolutionary Genomics (2014), Unix Software and Tools (2013), Data Mining Techniques and Applications (2011-2013), Operating Systems (2008), Developments in Digital Business (2009), Computer Systems and Architecture (2009), Genetic Analysis and Bioinformatics (2009-2012).

- **Administrative duties:** Careers officer, School of Computer Science, University of Nottingham (2008-2010), Library liaison, School of Computer Science, University of Nottingham (2010-2012), University of Nottingham HPC committee (2012-2013), Newcastle University HPC committee (2015-present), Newcastle University School of Computing Science HPC advocate (2015-present)

Professional memberships:

- Association for Computing Machinery (ACM) since 2005
- ACM Special Interest Group in Evolutionary Computation (SIGEVO) since 2005
- ACM Special Interest Group in High Performance Computing (SIGHPC) since 2011
- International Society for Computational Biology (ISCB) since 2013

Invited Talks (14)

- **2017:** June, "Inferring functional networks from rule-based machine learning models", invited seminar, School of Computer Science, University of Birmingham, UK
- **2015:** March, "The bi-directional feedback between big data mining and bioinformatics", invited seminar, School of Computing, University of Kent, Canterbury, UK // June 2015 "Machine Learning for Knowledge discovery from biological/biomedical data", University of Vic, Spain // July 2015 "Machine Learning for Knowledge discovery from omics/clinical OA data", University of Surrey
- **2014:** June "The bi-directional feedback between big data mining and bioinformatics", invited seminar, Centre for Scientific Computing, Warwick University, Coventry, UK
- **2012:** *February 2012* "Large scale data mining using Genetics-Based Machine Learning", **keynote talk**, MAEB-2012 conference, Albacete, Spain.
- **2010:** *May 2010* "Data Mining Protein Structures' Topological Properties to Enhance Contact Map Predictions", Bioinformatics and Genome Research Open Club, Weizmann Institute of Science, Israel
- **2009:** *June 2009* "GBML methods for large-scale datasets", Learning Classifier Systems Group, University of the West of England, Bristol, UK
- **2008:** *May 2008* "Memetic Pittsburgh Learning Classifier Systems", 1st Workshop on Knowledge Extraction based on Evolutionary Learning, Granada, Spain // *May 2008* "GBML methods for large-scale datasets", 1st Workshop on Knowledge Extraction based on Evolutionary Learning, Granada, Spain.
- **2007:** *December 2007* "Bioinformatics: Protein Structure Prediction using Evolutionary Rule Learning", School of Mathematics, University of La Laguna, Spain.
- **2006:** *December 2006* "Coordination number prediction using Learning Classifier Systems: Performance and interpretability", Structural Computational Biology Group of the Spain National Center for Oncology Research - CNIO, Madrid, Spain // *May 2006* "Pittsburgh Learning Classifier Systems for Protein Structure Prediction: Scalability and Explanatory Power", NCSA/ILLIAC Gathering on Evolutionary Learning, NIGEL2006, Urbana, IL, USA // *February 2006* "The GAssist Learning Classifier System", Learning Classifier Systems Group, University of the West of England, Bristol, UK.

Tutorials delivered (6)

- **2013:** *July 2013* "Large-Scale Data Mining using Genetics-Based Machine Learning", tutorial speaker, GECCO-2013 conference, Amsterdam, the Netherlands.
- **2012:** *August 2012* "Introduction to Bioinformatics", tutorial speaker, PPSN-2012 conference, Taormina, Italy // *July 2012* "Large-Scale Data Mining using Genetics-Based Machine Learning", tutorial speaker, GECCO-2012 conference, Philadelphia, USA
- **2011:** *July 2011* "Large-Scale Data Mining using Genetics-Based Machine Learning", tutorial speaker, GECCO-2011 conference, Dublin, Ireland.
- **2010:** *July 2010* "Large-Scale Data Mining using Genetics-Based Machine Learning", tutorial speaker, WCCI-2010 Conference, Barcelona, Spain
- **2009:** *July 2009* "Large-Scale Data Mining using Genetics-Based Machine Learning", tutorial speaker, GECCO-2009 conference, Portland, USA

Organising symposiums, workshops and conferences

- Co-chair, Evolutionary Computation, Machine Learning and Data Mining for Biology and Medicine (EvoBIO) track of the EvoApplications conference since 2016
- Co-organiser, Evolutionary Computation for Big Data and Big Learning workshop (ECBDL'14)
- Co-organiser, International Workshop on Learning Classifier Systems - IWLCS (2007-2011). Workshops co-chair. Genetic and Evolutionary Computation Conference (GECCO) in 2010 and 2011.
- Co-chair of the "Genetics-Based Machine Learning/Evolutionary Machine Learning" track, Genetic and Evolutionary Computation Conference (GECCO) in 2009, 2013 and 2014.
- Co-editor, special issue on "Metaheuristics for large-scale data mining", Memetic Computing, 2010.
- Co-chair, "Bioinformatics, Systems and Synthetic Biology" stream of the OR53 conference, 2011.
- Co-organiser, Plant Bioinformatics, Systems and Synthetic Biology Summer School, University of Nottingham, 2009

Book editing

- Co-editor of the book "Learning Classifier Systems; Revised Selected Papers of IWLCS2008 and IWLCS2009; Lecture Notes in Artificial Intelligence 6471, Springer 2010.
- Co-editor of the book "Learning Classifier Systems; Revised Selected Papers of IWLCS2006 and IWLCS2007; Lecture Notes in Artificial Intelligence 4998, Springer 2008.

Reviewing duties

- Member of the EPSRC Associate Peer Review College
- Member of the BBSRC Pool of Experts
- Grants reviewer for the French National Cancer Institute, Netherlands Organisation for Scientific Research, Swiss National Science Foundation, Leverhulme Trust, Royal Society and the Chile's Superior Council of the National Fund for Scientific & Technological Development.
- Editorial board member for PeerJ Computer Science and Progress in Artificial Intelligence.
- Reviewer for over 20 journals including: PLOS Computational Biology, PLOS One, Computers in Biology and Medicine, BMC Genomics, Information Sciences, IEEE Transactions on Evolutionary Computation, Memetic Computing, Applied Soft Computing, Soft Computing, IEEE Systems, Man and Cybernetics B, Neurocomputing, Evolutionary Computation Journal, WIREs Data Mining and Knowledge Discovery, Journal of Molecular Graphics and Modelling, Pattern Recognition Letters, BioSystems, Amino Acids, Artificial Intelligence, Knowledge and Information Systems, Data and Knowledge Engineering, Plant Physiology, Evolutionary Intelligence, Advances in Artificial Intelligence, Nature Communications, Big Data Journal.
- Program committee member of, among other conferences, Genetic and Evolutionary Computation Conference (since 2005), Spanish Conference on meta-heuristics and evolutionary and bio-inspired algorithms (since 2005), International Workshop on Learning Classifier Systems (since 2005), International Joint Conference on Neural Networks (since 2006), International Conference on Evolutionary Computation (CEC) (since 2008), International Workshop on Nature Inspired Cooperative Strategies for Optimization, NICO (since 2007), IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology (since 2006), Symposium on Applied Computing - SAC (since 2010), International Conference on Hybrid Artificial Intelligence Systems - HAIS (since 2009), The Ibero-American Conference on Artificial Intelligence (since 2008), Spanish Conference on Artificial Intelligence - CAEPIA (since 2009), International Work-Conference on Bioinformatics and Biomedical Engineering - IWBBIO (since 2013), Parallel Problem Solving from Nature - PPSN (since 2012), the AAAI Conference on Artificial Intelligence (since 2017).

Awards and competitions

- **2011**: Best paper award for the Machine Learning Track of the GECCO-2011 conference
- **2010**: Best paper award for the Machine Learning Track of the GECCO-2010 conference
- **2010**: Best *Ab-initio* method in the residue-residue contact prediction category of the 9th Community Wide Experiment on the Critical Assessment of Techniques for Protein Structure Prediction (CASP)
- **2007**: Bronze medal in the 'Humies Awards'. This prize is awarded to Evolutionary Computation applications that manage to obtain human-competitive results of high relevance and impact.

List of top 10 publications plus all publications in the last five years. Full list of publications is available at <http://homepages.cs.ncl.ac.uk/jaume.bacardit/publications.html>. Citations based on Google Scholar. Date of citation count: 7/9/2017. **Top 10 publications marked with blue background.**

Year	Title	Citations
2017	N. Lazzarini and J. Bacardit. RGIFE: a ranked guided iterative feature elimination heuristic for the identification of biomarkers. BMC Bioinformatics 18:322, 2017. IF: 2.448	--
	S. Baron, N. Lazzarini and J. Bacardit. Characterising the Influence of Rule-Based Knowledge Representations in Biological Knowledge Extraction from Transcriptomics Data. Proceedings of EvoApplications 2017, LNCS 10199, p.125-141, Springer, 2017	--
	A Garcia-Piquer, J Bacardit, A Fornells and E Golobardes. Scaling-up multiobjective evolutionary clustering algorithms using stratification. Pattern Recognition Letters, 93:69-77, 2017. IF: 1.995	--
2016	N. Lazzarini, P. Widera, S. Williamson, R. Heer, N. Krasnogor and J. Bacardit. Functional networks inference from rule-based machine learning models. BioData Mining 2016, 9:28. IF: 1.64.	1
	P.D. Gutiérrez, M. Lastra, J. Bacardit, J.M. Benítez and F. Herrera. GPU-SME-kNN: Scalable and Memory Efficient kNN and Lazy Learning using GPUs. Information Sciences, 330:385-402. IF: 3.364	4
	M. Franco, and J. Bacardit. Large-scale experimental evaluation of GPU strategies for evolutionary machine learning. Information Sciences, 330:385-402. IF: 3.364	2
2015	F. Eduati et al. (119 authors). Prediction of human population responses to toxic compounds by a collaborative competition. Nature Biotechnology, 33,933-940. IF: 43.113	26
	I. Triguero, S. del Río, V. López, J. Bacardit, J.M. Benítez and F. Herrera. ROSEFW-RF: The winner algorithm for the ECBDL'14 big data competition: An extremely imbalanced big data bioinformatics problem. Knowledge-Based Systems, (2015) 87:69-79. IF: 3.325	30
	A.L. Swan, D.J. Stekel, C. Hodgman, D. Allaway, M.H Alqahtani, A. Mobasheri and J. Bacardit. A machine learning heuristic to identify biologically relevant and minimal biomarker panels from omics data. BMC Genomics, 16(Suppl 1):S2. IF: 3.867	12
	I. Triguero, D. Peralta, J. Bacardit, S. Garcia and F. Herrera. MRPR: A MapReduce Solution for Prototype Reduction in Big Data Classification. Neurocomputing journal 150(A):331-345. IF: 2.005	82
	María Martínez-Ballesteros, Jaume Bacardit, Alicia Troncoso and José C. Riquelme. Enhancing the scalability of a genetic algorithm to discover quantitative association rules in large-scale datasets. Integrated Computer-Aided Engineering, 22(1):21-39. IF: 4.981	9
2014	J. Bacardit, P. Widera, N. Lazzarini and N. Krasnogor. Hard Data Analytics Problems Make for Better Data Analysis Algorithms: Bioinformatics as an Example. Big Data Journal, 2(3):164-176	4
	Alkurashi MM, May ST, Kong K, Bacardit J, Haig D, Elsheikha HM. Susceptibility to experimental infection of the invertebrate locusts (<i>Schistocerca gregaria</i>) with the apicomplexan parasite <i>Neospora caninum</i> .. PeerJ 2:e674	---
	D.J. Gibbs, J. Bacardit, A. Bachmair, and M.J. Holdsworth. The eukaryotic N-end rule pathway: conserved mechanisms and diverse functions. Trends in Cell Biology, 24(10):603-611. IF: 12.007	65
	I. Triguero, D. Peralta, J. Bacardit, S. Garcia and F. Herrera. A Combined MapReduce-Windowing Two-Level Parallel Scheme for Evolutionary Prototype Generation. Proceedings of the IEEE World Congress on Computational Intelligence 2014, pp. 3036—3043	5
	J. Blakes, O. Raz, U. Feige, J. Bacardit, P. Widera, T. Ben-Yehezkel, E. Shapiro and N. Krasnogor. A heuristic for maximizing DNA reuse in synthetic DNA library	11

	assembly. ACS synthetic biology 3 (8), 529-542. IF: 4.987	
	G.A Khoury, A. Liwo, F. Khatib, H. Zhou, G. Chopra, J. Bacardit, L. Bortot, R.A. Faccioli, X. Deng, Y. He, P. Krupa, J. Li, M.A. Mozolewska, A.K. Sieradzan, J. Smadbeck, T. Wirecki, S. Cooper, J. Flatten, K. Xu, D. Baker, J. Cheng, A.C.B. Delbem, C.A. Floudas, C. Keasar, M. Levitt, Z. Popovic, H.A. Scheraga, J. Skolnick, S.N. Crivelli and Foldit Players. WeFold: A Coopetition for Protein Structure Prediction. PROTEINS: Structure, Function, and Bioinformatics 82(9):1850-1868, 2014. IF: 2.627	30
	A. Garcia-Piquer, A. Fornells, J. Bacardit, A. Orriols-Puig and E. Golobardes. Large-Scale Experimental Evaluation of Cluster Representations for Multiobjective Evolutionary Clustering. IEEE Transactions on Evolutionary Computation, 18(1):36-53, 2014	24
2013	A.L. Swan, K.L. Hillier, J.R. Smith, D. Allaway, S. Liddell, J. Bacardit and A. Mobasher. Analysis of mass spectrometry data from the secretome of an explant model of articular cartilage exposed to pro-inflammatory and anti-inflammatory stimuli using machine learning. BMC Musculoskeletal Disorders, 14:349, 2013. IF: 1.898	13
	J. Bacardit and X. Llorà. Large-scale data mining using genetics-based machine learning. WIREs Data Mining and Knowledge Discovery, 3: 37-61. Contribution: lead author, 80%. IF: 1.358	58
	A.L. Swan, A. Mobasher, D. Allaway, S. Liddell and J. Bacardit. Application of Machine Learning to Proteomics Data: Classification and Biomarker Identification in Postgenomics Biology. OMICS: A Journal of Integrative Biology, 17(12): 595-610, 2013 IF: 2.730 (Please note, this journal has nothing to do with the Omics Publishing Group)	34
	M. Franco, N. Krasnogor and J. Bacardit. GAssist vs. BioHEL: Critical Assessment of Two Paradigms of Genetics-based Machine Learning. Soft Computing, 17(6):953-981, June 2013. IF: 1.304	11
	D.A. Calian and J. Bacardit. Integrating memetic search into the BioHEL evolutionary learning system for large-scale datasets. Memetic Computing, 5(2):95-130, June 2013. IF: 1.000	5
2012	J. Bacardit, P. Widera, A. Márquez, F. Divina, J.S. Aguilar-Ruiz and Natalio Krasnogor. Contact map prediction using a large-scale ensemble of rule sets and the fusion of multiple predicted structural features. Bioinformatics 28 (19): 2441-2448. IF: 5.323	33
	E. Glaab, J. Bacardit, J.M. Garibaldi and N. Krasnogor. Using Rule-Based Machine Learning for Candidate Disease Gene Prioritization and Sample Classification of Cancer Gene Expression Data. PLoS ONE 7(7):e39932. IF: 3.730	43
2011	Functional network construction in Arabidopsis using rule-based machine learning on large-scale data sets. GW Bassel, E Glaab, J Marquez, MJ Holdsworth, J Bacardit. The Plant Cell 23 (9), 3101-3116. IF: 8.987	58
2010	Speeding Up the Evaluation of Evolutionary Learning Systems using GPGPUs. M. Franco, N. Krasnogor and J. Bacardit. In Proceedings of the 12th Annual Conference on Genetic and Evolutionary Computation (GECCO2010), 1039-1046, ACM Press, 2010.	51
2009	Improving the scalability of rule-based evolutionary learning. J Bacardit, EK Burke, N Krasnogor. Memetic Computing 1 (1), 55-67. IF: 1.000	68
	Alcalá-Fdez, J., Sánchez, L., García S., del Jesus, M.J., Ventura, S., Garrell, J.M., Otera, J., Romero, C., Bacardit, J., Rivas, V.M., Fernández, J.C. and Herrera, F. KEEL: A Software Tool to Assess Evolutionary Algorithms for Data Mining Problems. Soft Computing Journal, 13(3):307-318, 2009. IF: 1.328.	788

Overall publication count

Citations: 2513

H-index: 26

Type of publication	Number
Journal papers	34
Refereed Conference Papers	31
Book Chapters	6
Edited books	2
Total	73