

researcher cv

andr  s r. masegosa

10/12/2020



personal details

address: Almer  a (Spain)

age: 40 years old

email: andresmasegosa@ual.es

phone number: +34 653 414 732

web: <https://andresmasegosa.github.io>

github: <https://github.com/andresmasegosa>

twitter: @andresmasegosa

research interests

Artificial Intelligence – Machine Learning – Probability Theory –
Uncertainty Modelling – Bayesian Statistics – Probabilistic
Programming – Variational Inference – PAC-Bayes theory

education

- 2009 - **PhD in Computer Science**. University of Granada, 2009. Score: A+. Dissertation on Supervised Classification Models. Applications to Genomics and Information Retrieval.
- 2008 - **Master in Soft Computing and Intelligent Systems**. University of Granada, 2008. Score: A+. Dissertation on Modelling Contextual Relevance to measure the impact of Context.
- 2007 - **Diploma of Advance Studies in Probabilistic Graphical Models for Artificial Intelligence and Data Mining**. University of Granada, 2007. Score: A+. Dissertation on a Semi-naive Bayes classifier with grouping of cases.
- 2005 - **Diploma of Advance Studies in Soft Computing and Intelligent Systems**. University of Granada. Dissertation on Gene Data Classification. Score: A.
- 2004 - **Computer Engineering Bachelor Degree**. University of Granada.

professional experience

- 2018 - **Assistant Professor** at the Department of Mathematics of University of Almería (Spain).
- 2016-2018 - **Senior researcher** at the Department of Mathematics of University of Almería (Spain).
- 2014-2016 - **Post-doc researcher** at the Department of Computer Science of NUTNU (Norway).
- 2009-2013 - **Post-doc researcher** at the Department of Computer Science of University of Granada (Spain).
- 2004-2009 - **PhD Student** at the Department of Computer Science

of University of Granada (Spain).

scholarships, grants

2020-2022 – National Research Grant. Explainable Machine Learning: A Probabilistic Approach. Competitive National Call by the Spanish Ministry of Science. Co-PI. Funded with 62,000€.

2019-2020 – José Castillejo Scholarship. Competitive National Call by the Spanish Ministry of Science. Funded with 15,000€ for a research visit at University of Copenhagen.

2017-2018 – Early-stage Researcher Grant. Probabilistic Programming Languages for the Development of Intelligent Applications on Large Volumes of Data. Competitive National Call by the Spanish Ministry of Science. PI. Funded with 172,000€.

2013-2013 – Young Researcher Scholarship. Local Call (University of Granada). Discretization of Expression Quantitative Trait Loci for Learning with Machine Learning Models. Funded with 3,000€.

2005-2009 – PhD Scholarship. Competitive National Call by the Spanish Ministry of Science. Funded with 96,000€.

research internships/visits

August 2019 – January 2020 - Department of Computer Science. University of Copenhagen (DK). Host: Dr. Yevgeny Seldin
Achievements: 2 conference papers accepted at NeurIPS 2020.

July-August 2017 and 2018 - Department of Computer Science. Technical University of Berlin (DE). Host: Dr. Shinichi Nakayima.
Achievements: 1 conference paper accepted at ICCVW'17.
1 conference paper rejected at AISTATS'19.

March 2014 - August 2016 - Department of Computer Science. Aalborg University (Denmark). Host: Dr. Thomas Nielsen. Achievements: 2 conference papers accepted at ICML'17 and ECAI'16. 2 journal papers at IJAR and IEEE-CIM.

July-August 2012 - Department of Computer Science. University of Utrecht. Host: Professor Linda C. van der Gaag. Achievements: 1 conference paper rejected at UAI'13. 1 journal paper at IJAR.

July-December 2006 - Department of Computer Science. University of Glasgow (UK). Host: Professor Joemon Jose. Achievements: 2 papers at two top international conferences on Information Retrieval: ECIR'07 and SIGIR'07.

invited talks, tutorials & seminars

November 2020 – **Learning under model misspecification.** Talk at the Alan Turing Institute (London).

October 2019 - **Probabilistic Programming with Deep Neural Networks.** Tutorial at the Autumn school on Machine Learning 2019 in Tbilisi (Georgia).

September 2019 - **A Bayesian approach for modelling of non-stationary data streams.** Talk at Department of Computer Science in the University of Copenhagen (Denmark).

June 2019 - **Variational Inference in Probabilistic Programming.** Tutorial at the Probabilistic AI Summer Schools 2019 in Trondheim (Norway).

August 2018 - **Bayesian Modelling of Concept Drift.** Talk at the Machine Learning Group in the Technical University of Berlin.

June 2018 - **Probabilistic Machine Learning with the AMIDST Toolbox.** Tutorial at the Geilo Winter School 2018 (Norway).

November 2016 - **AMIDST Toolbox - Scalable Probabilistic Machine Learning**. Talk at the Apache Big Data Europe 2016 in Seville (Spain).

July 2012 - **Interactive Learning of Bayesian Networks**. Talk at the Department of Information and Computing Sciences in the University of Utrecht (Netherlands).

research publications

Bibliometrics (Google Scholar).

- 25 papers in JCR indexed journals. 13 of them in Q1 journals.
- 38 papers in international conferences. 5 of them in Core A* conferences (SIGIR, UAI, ICML and NeurIPS). 3 of them in Core A conferences (ECIR, ECAI, IDA). 9 of them in Core C conferences (ECSQARU, IPMU, INEX).
- Number of citations: 675. H-index: 14. i10-index: 19.

List of the 10 most relevant publications.

Masegosa, A. R. (2020). Learning under model misspecification: Applications to variational and ensemble methods. Advances in Neural Information Processing Systems, **NeurIPS**, 33. [Core A* conference]

Masegosa, A. R., Lorenzen, S., Igel, C., & Seldin, Y. (2020). Second order PAC-Bayesian bounds for the weighted majority vote. Advances in Neural Information Processing Systems, **NeurIPS**, 33. [Core A* conference]

Cózar, J., Cabañas, R., Salmerón, A., & **Masegosa, A. R.** (2020). InferPy: Probabilistic modeling with deep neural networks made easy. **Neurocomputing**, 415, 408-410. [JCR Position 28/137 – Computer Science, AI]

Masegosa, A. R., Ramos-López, D., Salmerón, A., Langseth, H., & Nielsen, T. D. (2020). Variational Inference over Nonstationary Data Streams for Exponential Family Models. **Mathematics**, 8(11), 1942. [JCR Position 28/328 - Mathematics]

Masegosa, A. R., Nielsen, T. D., Langseth, H., Ramos-López, D., Salmerón, A., & Madsen, A. L. (2017). Bayesian models of data streams with hierarchical power priors. In Proceedings of the 34th International Conference on Machine Learning, **ICML**, 70. [Core A* conference]

Masegosa, A. R., Martinez, A. M., & Borchani, H. (2016). Probabilistic graphical models on multi-core CPUs using Java 8. **IEEE Computational Intelligence Magazine**, 11(2), 41-54. [JCR Position 6/133 – Computer Science, AI]

Masegosa, A. R. (2014). Stochastic discriminative EM. In Proceedings of the Thirtieth Conference on Uncertainty in Artificial Intelligence. **UAI**. (pp. 573-582). [Core A* conference]

Abellán, J., & **Masegosa, A. R.** (2012). Bagging schemes on the presence of class noise in classification. **Expert Systems with Applications**, 39(8), 6827-6837. [JCR Position 13/79 – Operation Research and Management]

Cano, A., **Masegosa, A. R.**, & Moral, S. (2011). A method for integrating expert knowledge when learning Bayesian networks from data. **IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)** 41(5), 1382-1394. [JCR Position 10/111 - Computer Science]

Abellan, J., & **Masegosa, A. R.** (2010). An ensemble method using credal decision trees. **European journal of operational research**, 205(1), 218-226. [JCR Position 6/75 – Operation Research and Management]

research projects

06/2020 – 05/2023 - **Explainable Machine Learning: A probabilistic Approach**. Funder: Spanish Ministry of Science. Co-PI.

01/2017 – 12/2019 - **Probabilistic Programming Language for the Development of Intelligent Applications on Large Volumes of Data**. Funder: Spanish Ministry of Science. PI.

01/2014 – 12/2016 - **AMIDST: Analysis of massive data streams.** Funder: European Commission. FP7-ICT. Post-Doc.

01/2013–12/2015 – **Probabilistic Graphical Models for Scalable Data Analytics,** Funder: Spanish Ministry of Science, Research Team.

01/2013–12/2013 - **Discretization of Expression Quantitative Trait Loci for Learning with Machine Learning Models,** Funder: University of Granada, PI.

01/2010–12/2012 - **Data Mining with Probabilistic Graphical Models,** Funder: Spanish Ministry of Science, Research Team.

01/2008–12/2010 - **Clados: Detection of Structural Anomalies of the Genome Using MGPs,** Funder: Regional Government of Andalucía, Research Team.

06/2009–12/2011 - **Multimodal Interaction in Pattern Recognition and Computer Vision,** Funder: Spanish Ministry of Science, Post-Doc.

01/2007–12/2009 - **Design of New Algorithms in Probabilistic Graphical Models,** Funder: Spanish Ministry of Science, PhD Student.

01/2004–12/2006 - **Adaptive Learning of Probabilistic Graphic Models,** Funder: Spanish Ministry of Science, PhD Student.

industrial projects

01/2004–12/2006 - **Development of Modules 7 and 8 i-Apus Project: Advanced Police Unification Software.** Company: Granatel S.L., Research Team, 64,112€.

01/2006–04/2006 - **Software for commercial sectorization and route planning,** Company: Gobile S.L. Reserch Team, 14,500€.

software projects



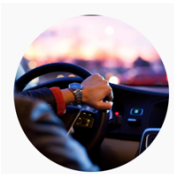
InferPy: Deep Probabilistic Modelling with Tensorflow Made Easy

Lead Developer and Researcher
<https://inferpy.readthedocs.io/>



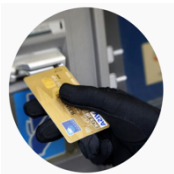
AMIDST: a Java toolbox for scalable probabilistic machine learning

Lead Developer and Researcher
<http://www.amidsttoolbox.com>



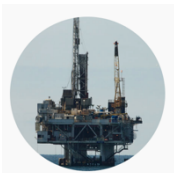
Early recognition of traffic manoeuvre intentions

Developer and Researcher
<http://project.amidsttoolbox.com>



Risk prediction in credit operations

Developer and Researcher
<http://project.amidsttoolbox.com>



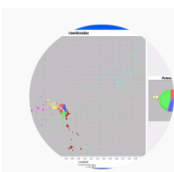
Real time pattern recognition in drilling logs

Developer and Researcher
<http://project.amidsttoolbox.com>



Crime prediction using data mining methods.

Lead Developer and Researcher
<https://andresmasegosa.github.io/projects/>



Sales Force Designing using an Artificial Intelligence based approach.

Lead Developer and Researcher
<https://andresmasegosa.github.io/projects/>



Elvira, a Java tool to construct probabilistic models-driven based decision support systems.

Developer and researcher
<http://leo.ugr.es/elvira>

services to the community

I regularly review for international conferences such as NeurIPS, ICML, ICLR, ECML and PGM, and for journals like IJAR, Neurocomputing, Applied Soft-Computing, Information Sciences, Expert Systems with Applications, etc.

teaching experience

I have teaching experience in three different universities: University of Granada (UGR), where I got my PhD; University of Aalborg (AAU) where I did a post-doc; and University of Almería (UAL) where I've been principal investigator and assistant professor during the last four years.

Courses

- Computing in Environmental Sciences degree. UGR, 2006. 5,5 ECTS¹. 30 participants, in Spanish. Role: tutoring practical exercises.
- Computing and Data Bases in Library Science degree. UGR, 2007 and 2008. 5,5 ECTS each year. Around 30 participants each year, in Spanish. Role: tutoring practical exercises.
- Intelligent Systems in Telecommunications Engineering. UGR, 2010. 8 ECTS. 60 participants, in Spanish. Role: Course supervisor, lectures and tutoring practical exercises.
- Data Bases 101 in Computer Science degree. UGR, 2010. 3 ECTS. 50 participants, in Spanish. Role: Course supervisor, lectures and tutoring practical exercises.
- Statistics 101 in Computer Science degree. UAL, 2018 and 2019. 5,5 ECTS each year. Around 80 participants each year, in Spanish. Role: Course co-supervisor, lectures and tutoring practical exercises.

¹ The stated ECTS credits refer to the students working load (1 ECTS=27 student learning hours) assigned to the part of the course I was responsible for.

- Statistics 101 in Economics degree. UAL, 2018 and 2019. 3 ECTS each year. 40 participants each year, in Spanish. Role: Tutoring practical exercises.
- Reliability and Risk Management in Computer Science degree. UAL, 2018. 5 ECTS. 10 participants, in Spanish. Role: Course supervisor, lectures and tutoring practical exercises.
- Statistics 101 in Environmental Sciences degree. UAL, 2019. 2,5 ECTS. 60 participants, in Spanish. Role: tutoring practical exercises.
- Statistics 101 in Tourism degree. UAL, 2019. 5,5 ECTS. 80 participants, in Spanish. Role: Course supervisor, lectures and tutoring practical exercises.

Supervision

- Predicting the Outcome of League of Legends Matches using Machine Learning on Big Data. Master Thesis. Aalborg University, 2015. Unique Supervisor. Supervision using the Problem based Learning approach.
- Optimizing Item Shopping for League of Legends using Machine Learning. Aalborg University, 2016. Master Thesis. Unique Supervisor. Supervision using the Problem based Learning approach.
- Missing Data Modelling with Variational Inference, Master project. University of Almería, 2019. Master Thesis. Co-supervisor.

Examinator/Censor

- Member of the Assessment committee of 3 PhD students: University of Granada (02/11/2018 and 24/6/2020) and University of Castilla La Mancha, Albacete, Spain (3/05/2019).
- Member of the Assessment committee of Bachelor Final Projects in June and September 2020.

training in pedagogy

- “Teaching in Multicultural Class” in 2005, 25 hours. This course aimed to foster the reflection around the presence of students in the class with very different backgrounds.
- “Pedagogical Adaptation Course” in 2005, 180 hours. This course is required by Spanish authorities to teach in high schools and vocational training centers. During this course, I received general training in Psychopedagogy. It was completed with teaching practices in a local high school.
- “Introduction to Problem Based Learning” in 2015 at Aalborg University (AAU), 10 hours. This introductory course was addressed to those supervising groups of students working on master/bachelor theses. The course provided training in the problem based learning methodology, which is the main teaching approach at AAU. I applied this methodology when supervising two master thesis projects.