

miguelaraujo

machine learning and data mining

contact

+351 919117733

maraujo@cs.cmu.edu
http://www.maraujo.com

languages

portuguese native
english & spanish fluency
german basics

programming

C++, Java, Scala
Python, Matlab
Apache Spark

profile

Strong background in algorithms, data mining and machine learning. I've developed anomaly detection, pattern mining and prediction solutions in a variety of domains using telecommunication records, bank transactions, traffic-sensors measurements and social networks graphs. Numerous refereed articles, a best paper award and a patent pending.

professional: algorithms, machine learning, data analysis and mining, graph mining, recommendation systems, anomaly detection.

personal: traveling, go (board game), scientific reading, fantasy novels.

education

2012– May 2017 **Dual PhD** in Computer Science Carnegie Mellon University, Pittsburgh, USA
University of Porto, Portugal

Focus on mining patterns and anomalies in static or temporal graphs with applications to social network analysis, fraud detection and recommendation systems. Developed novel matrix and tensor decompositions methods to cluster, find patterns and forecast high-dimensional data such as phone call networks and movie recommendations.

2007–2012 **Masters** in Informatics and Computing Engineering University of Porto, Portugal

Can we model human decisions based on choices at decision points during a planning task? Developed inverse reinforcement learning techniques to identify taxi drivers' priorities when choosing their route within the city.

experience

2015–2015 **Feedzai** Lisbon, Portugal

Research Scientist internship (4 months)

Developed large-scale fraud-detection tools. Focused on the *point of compromise* detection problem: how can we automatically detect where credit and debit cards were stolen? How can we detect data breaches in tera-sized datasets?

One patent pending.

- Created sound techniques for *point of compromise* detection.
- Leveraged Scala and Apache Spark to mine datasets with billions of records.

research

2009–2010 **LIACC - Artificial Intelligence and Computer Science Lab.** Porto, Portugal

Undergraduate Researcher

Developed high-level abstractions to enable simultaneous testing of artificial intelligence solutions in multiple traffic simulators. Implemented artificial intelligence algorithms for the dynamic control of traffic lights in traffic simulators.

teaching

2014	Faculty of Sciences of the University of Porto <i>Teaching Assistant of the Algorithms course.</i>	Porto, Portugal
2013	Carnegie Mellon University <i>Teaching Assistant of the undergraduate Algorithms course.</i>	Pittsburgh, USA
2010–2011	Faculty of Engineering of the University of Porto <i>Undergraduate Teaching Assistant of the Algorithm Design and Analysis and Graphical Applications Laboratory courses.</i>	Porto, Portugal

cool things

2014	Web Scraper for Price Comparisons Engine Python web scraper using Scrapy that would automatically crawl, parse, and store product details from multiple online retailers.
2007–2016	Judge and Problem Setter in Programming Contests Judge and problem setter in the ICPC South-Western Regional (2014-2016) and in multiple editions of the Portuguese Olympiads in Informatics.

awards

2009,2010	Merit Scholarships Yearly awarded to 1 in 500 students.	Faculty of Engineering, University of Porto
2007–2011	ACM Inter-Collegiate Programming Contest - SWERC University representative in Portugal (2007), Germany (2008) and Spain (2009, 2010 (Bronze Medalist), 2011).	University of Porto
2006–2007	International Olympiads in Informatics Representative in the International Olympiads in Mexico (2006) and Croatia (2007).	

publications

Most relevant:

1. **BreachRadar: Automatic Detection of Points-of-Compromise**
Miguel Araujo, Miguel Almeida, Jaime Ferreira, Luis Silva, Pedro Bizarro. *SIAM International Conference on Data Mining (SDM)*. 2017. Houston, United States.
2. **FastStep: Scalable boolean matrix decomposition**
Miguel Araujo, Pedro Ribeiro, Christos Faloutsos. *Pacific-Asia Knowledge Discovery and Data Mining (PAKDD)*. 2016. Auckland, New Zealand.
3. **Com2: Fast Automatic Discovery of Temporal (Comet) Communities**
Miguel Araujo, Spiros Papadimitriou, Stephan Günnemann, Christos Faloutsos, et. al. *Pacific-Asia Knowledge Discovery and Data Mining (PAKDD)*. 2014. Taiwan. **(Student Paper Award)**
4. **Understanding Sequential Decisions via Inverse Reinforcement Learning**
Siyuan Liu, **Miguel Araujo**, Ramayya Krishnan, Emma Brunskill, Rosaldo Rossetti and João Barros. *IEEE Mobile Data Management Conference (MDM)*. 2013. Milan, Italy.