Gustavo Lacerda *Data Science, Machine Learning, Statistics*

CONTACT INFORMATION

Gustavo Lacerda guscv@optimizelife.com

www.optimizelife.com *Voice:* 917-655-8707
github.com/gusl

EDUCATION

2010 – 2014: ***PhD candidate in Statistics, Columbia University (left with Master's)*** Research in non-parametric factor analysis, subspace dictionary learning. **Core Courses:** Mathematical Stat I, II, III; Applied Stat I, II; Analysis & Probability I; Stochastic Processes. **Electives:** Bayesian Computation; Statistical Computing; Non-parametric Bayes; Sparsity Seminar; Bayesian Models for Machine Learning.

2008 – 2010: ***MSc in Computer Science, University of British Columbia*** Thesis about inferring genetic clusters from relational data. Coursework included Machine Learning I and II; Empirical Algorithmics; Computer Vision.

Expertise: graphical models, dimension reduction, sparsity, model selection.

EMPLOYMENT HISTORY

- **Summer 2013 : Google (Mountain View, CA) – Software Engineering Intern.** Exploratory statistical analysis of the webtables corpus. Wrote Flume jobs (protobufs, SSTables) to extract data, accompanied progress on borg. Worked with sharded files in the Colossus File System. Data analysis with R.
- **2010 – 2014: Columbia University, Statistics Department (New York, NY) – Teaching Assistant. Graduate:** Data Mining; Stochastic Processes; Stat Inference **Undergrad:** Intro to Stat; Intro to Stat Reasoning.
- **2008 – 2009: University of British Columbia (Vancouver, British Columbia) – Teaching Assistant:** Models of Computation, Introduction to Artificial Intelligence, Introduction to Relational Databases.
- **2006 – 2008: Carnegie Mellon University (Pittsburgh, PA) – AI Programmer for the SimStudent Project.** Developing, maintaining a system that learns production rule models of student skills from behavior data; designing search algorithms; modeling problem-solving domains like stoichiometry.
- **2006: Cadence Design Systems (Munich, Germany) – Lisp Programming Intern.** Worked in the custom development team: developing, debugging, testing and re-releasing packages.
- **2001 – 2002: Amazon Technologies (Woburn, MA) – Software Engineer.** Web programming with JSP / JBuilder: front-end and back-end; data conversion, automatic report generation.
- **Summer 2000: Bell Labs, Lucent Technologies (Murray Hill, NJ) – Research Intern.** – Research intern at the Mathematics of Communications department. Implemented variations of the Viterbi algorithm to decode linear error-correcting codes over simulated noisy channels.

LANGUAGES

R, Python, C++, Julia, Matlab, Java, Common Lisp.

Fluent in English, Portuguese, Dutch. **Passable** French, German, Spanish.

Publications

- **Gustavo Lacerda**, John Paisley – "Beta Process Subspace Analysis" (in progress)
- **Gustavo Lacerda** – "Identification of gene modules using a generative model for relational data", UBC Master's thesis (2010), supervised by Jennifer Bryan.
- **Gustavo Lacerda**, Peter Spirtes, Joseph Ramsey, Patrik Hoyer – Discovering Cyclic Causal Models by Independent Components Analysis. *Uncertainty in Artificial Intelligence (UAI-2008). (plenary talk)*
- Patrik Hoyer, Aapo Hyvärinen, Richard Scheines, Peter Spirtes, Joseph Ramsey, **Gustavo Lacerda**, Shohei Shimizu – Causal discovery of linear acyclic models with arbitrary distributions. *Uncertainty in Artificial Intelligence (UAI-2008)*
- Noboru Matsuda, William W. Cohen, Jonathan Sewall, **Gustavo Lacerda**, and Kenneth R. Koedinger – Why Tutored Problem Solving may be better than Example Study: Theoretical Implications from a Simulated-Student Study. *In Proceedings of the International Conference on Intelligent Tutoring Systems 2008*
- Noboru Matsuda, William W. Cohen, Jonathan Sewall, **Gustavo Lacerda**, and Kenneth R. Koedinger – Evaluating a Simulated Student using Real Students Data for Training and Testing, *In Artificial Intelligence in Education*
- Noboru Matsuda, William W. Cohen, Jonathan Sewall, **Gustavo Lacerda**, and Kenneth R. Koedinger – Predicting Students' Performance with SimStudent: Learning Cognitive Skills from Observation. *In International Conference on User Modeling 2007.*
- S. Fissaha Adafre, W.R. van Hage, J. Kamps, **G. Lacerda de Melo**, and M. de Rijke – The University of Amsterdam at CLEF 2004.