# GILLES LOUPPE

*Academic resume*

## PERSONAL INFORMATION


Born in Belgium, 26 April 1987

*email* g.louppe@nyu.edu

*scholar* F_77d4QAAAA

*orcid* 0000-0002-2082-3106

*github* @glouppe

*phone* +41 (0) 77 99 53 433


## RESEARCH INTERESTS AND OBJECTIVES

As a researcher, my far ambition is to make artificial intelligence a cornerstone of the modern scientific method. Using particle physics as a testbed, my present research interests circle around how to use or design new machine learning algorithms to approach data-driven scientific problems in new and revolutionizing ways. With this goal in mind, my topics of current research include methods for simulator-based likelihood-free inference, algorithms to handle systematic uncertainties in inference models, and developments in sequential model-based optimization.

## WORK EXPERIENCE

*2015–Present* Postdoctoral Associate · *New York University (USA), CERN (Switzerland)*

Machine learning research for high energy physics.

*2014–2015* Marie-Curie COFUND Research Fellow · *CERN (Switzerland)*

Machine learning and data analysis on text and bibliographic data.

*2010–2014* F.R.S.-FNRS Research Fellow · *University of Liège (Belgium)*

Fundamental and applied research in machine learning and data analysis. Expertise in tree-based methods.

Teaching assistant (Introduction to Algorithmics, Data Structures and Algorithms, Machine Learning).

## EDUCATION

*2010–2014* PhD in Computer Sciences · *University of Liège (Belgium)*

Thesis: Understanding Random Forests – From Theory to Practice.

*2008–2010* Master in Computer Sciences · *University of Liège (Belgium)*

Thesis: Collaborative Filtering – Scalable approaches using Restricted Boltzmann Machines.
Erasmus student at the Royal Institute of Technology (KTH), Sweden.
*Summa cum laude.*

*2005–2008* Bachelor in Computer Sciences · *University of Liège (Belgium)*

*Summa cum laude.*

## PAPERS

*2016*

[22] Learning to Pivot with Adversarial Networks. Gilles Louppe, Michael Kagan, Kyle Cranmer. [PDF, Code]

[21] Experiments using machine learning to approximate likelihood ratios for mixture models. Kyle Cranmer, Juan Pavez, Gilles Louppe, W. K. Brooks. [PDF]

[20] Approximating Likelihood Ratios with Calibrated Discriminative Classifiers. Kyle Cranmer, Juan Pavez, Gilles Louppe. [PDF, Code]

[19] Random subspace with trees for feature selection under memory constraints. Antonio Sutera, Clia Chatel, Gilles Louppe, Louis Wehenkel, Pierre Geurts. [PDF]

[18] Context-dependent feature analysis with random forests. Antonio Sutera, Gilles Louppe, Vn Anh Huynh-Thu, Louis Wehenkel, Pierre Geurts. [PDF]

[17] Visualizatoin of Publication Impact. Eamonn Maguire, Javier Martin Montull, Gilles Louppe. [PDF, Code]

[16] Collaborative analysis of multi-gigapixel imaging data using Cytomine. Raphael Maree, Loic Rollus, Benjamin Stevens, Renaud Hoyoux, Gilles Louppe, Remy Vandaele, Jean-Michel Begon, Philipp Kainz, Pierre Geurts, Louis Wehenkel. [PDF, Code]

*2015*

[15] Pitfalls of evaluating a classifiers performance in high energy physics applications. Gilles Louppe, Tim Head. [Notebook]

[14] Ethnicity sensitive author disambiguation using semi-supervised learning. Gilles Louppe, Hussein Al-Natsheh, Mateusz Susik, Eamonn Maguire. [PDF, Code]

[13] Collaborative analysis of gigapixel images using Cytomine. Rmy Vandaele, Raphal Mare, Pierre Geurts, Loc Rollus, Benjamin Stvens, Renaud Hoyoux, Jean-Michel Begon, Gilles Louppe, Louis Wehenkel. Acta Stereologica, July, 2015. [PDF]

[12] Scikit-learn: Machine Learning Without Learning the Machinery. Gael Varoquaux, Lars Buitinck, Gilles Louppe, Olivier Grisel, Fabian Pedregosa, Andreas Mueller. GetMobile: Mobile Computing and Communications 19 (1), 29-33, 2015. [PDF]

[11] Solar Energy Prediction: An International Contest to Initiate Interdisciplinary Research on Compelling Meteorological Problems. Amy McGovern, David John Gagne II, Lucas Eustaquio, Gilberto Titericz Junior, Benjamin

Lazorthes, Owen Zhang, Gilles Louppe, Peter Prettenhofer, Jeffrey Basara, Thomas Hamill, David Margolin. Bulletin of the American Meteorological Society, 2015. [PDF]

*2014* [10] Understanding Random Forests: From Theory to Practice. Gilles Louppe. PhD thesis, University of Liège, 2010. [PDF, Code]

[9] Simple connectome inference from partial correlation statistics in calcium imaging. Antonio Sutera, Arnaud Joly, Vincent Francois-Lavet, Zixiao Aaron Qiu, Gilles Louppe, Damien Ernst, Pierre Geurts. [PDF, Code]

[8] Exploiting SNP Correlations within Random Forest for Genome-Wide Association Studies. Vincent Botta, Gilles Louppe, Pierre Geurts, Louis Wehenkel. PLoS ONE 9(4), 2014. [PDF, Code]

[7] A hybrid human-computer approach for large-scale image-based measurements using web services and machine learning. Raphael Marée, Loic Rollus, Benjamin Stevens, Gilles Louppe, et al. 11th IEEE International Symposium on Biomedical Imaging, Beijing, China, 2014. [PDF]

*2013* [6] Understanding variable importances in forests of randomized trees. Gilles Louppe, Louis Wehenkel, Antonio Sutera, Pierre Geurts. NIPS, Lake Tahoe, United States, 2013. [PDF, Code]

[5] API design for machine learning software: experiences from the scikit-learn project. Lars Buitinck, Gilles Louppe, Mathieu Blondel, et al. ECML-PKDD 2013 Workshop: Languages for Data Mining and Machine Learning, Pragues, Czech Republic, 2013. [PDF, Code]

*2012* [4] Ensembles on Random Patches. Gilles Louppe, Pierre Geurts. ECML-PKDD 2012, Bristol, UK, 2012. [PDF, Code]

*2011* [3] Learning to rank with extremely randomized trees. Pierre Geurts, Gilles Louppe. JMLR: Workshop and Conference Proceedings, 14, 49-61, 2011. [PDF]

*2010* [2] A zealous parallel gradient descent algorithm. Gilles Louppe, Pierre Geurts. Learning on Cores, Clusters and Clouds workshop, NIPS, Vancouver, Canada, 2010. [PDF, Code]

[1] Collaborative filtering: Scalable approaches using restricted Boltzmann machines. Gilles Louppe. Master's thesis, University of Liège, 2010. [PDF, Code]

### TALKS

*2016* [32] Learning to Pivot with Adversarial Networks. ATLAS ML Forum, CERN, Switzerland November 17, 2016. [Materials]

[31] Series of Lectures on Machine Learning. Machine Learning and Data Science in Physics, Barcelona October 17-21, 2016. [Materials]

[30] Learning to generate with adversarial networks. US ATLAS Physics Support, Software and Computing meeting, Chicago, USA August 3, 2016. [Materials]

[29] Learning to generate with adversarial networks. ATLAS ML Forum, CERN, Switzerland July 21, 2016. [Materials]

[28] Learning to generate with adversarial networks. DS @ HEP at the Simons Foundation, New York, USA July 5-7, 2016. [Materials]

[27] Learning to generate with adversarial networks. Software Tech Forum, CERN, Switzerland June 27, 2016. [Materials]

[26] Approximating likelihood ratios with Calibrated Classifiers. Invited lecture at the 2nd MLHEP summer school, Lund, Sweden June 22, 2016. [Materials]

[25] Robust and Calibrated Classifiers with Scikit-Learn. Zurich ML meetup, Switzerland April 13, 2016. [Materials]

[24] Approximating likelihood ratios with Calibrated Classifiers. ETH, Zurich, Switzerland April 13, 2016. [Materials]

[23] Approximating likelihood ratios with Calibrated Classifiers. ATLAS ML workshop, CERN, Switzerland March 29-31, 2016. [Materials]

[22] An introduction to Bayesian Optimization. ATLAS ML workshop, CERN, Switzerland March 29-31, 2016. [Materials]

[21] An introduction to machine learning with Scikit-Learn. Heavy Flavour Data Mining workshop, Zurich, Switzerland February 18, 2016. [Materials]

*2015*

[20] Pitfalls of evaluating a classifiers performance in high energy physics applications. ALEPH workshop, NIPS, Montral, Canada. December 11, 2015. [Materials]

[19] An introduction to machine learning with Scikit-Learn. Data Science at LHC, Switzerland. November 12, 2015. [Materials]

[18] Classification with a control channel: Don't cheat yourself! CERN, Switzerland. October 5, 2015. [Materials]

[17] Scikit-Learn tutorial. AstroHack Week, New York, USA. September 30, 2015. [Materials]

[16] Understanding Random Forests. CERN, Switzerland. September 21, 2015. [Materials]

[15] An introduction to Machine Learning with Scikit-Learn. CERN, Switzerland. April 23, 2015. [Materials]

[14] Tree models with Scikit-Learn: Great learners with little assumptions. PyData, Paris, France. April 5, 2015. [Materials]

[13] Machine Learning for Author Disambiguation. CERN, Switzerland. March 3, 2015. [Materials]

*2014*

[12] Bias-variance decomposition in Random Forests. Paris Machine Learning Meetup 4 (saison 2), Paris, France. December 9, 2014. [Materials]

[11] Scikit-Learn in Particle Physics. Data Science Academic software: From scikit-learn and scikit-image to domain science, Paris, France. November 18, 2014. [Materials]

[10] Understanding Random Forests: From Theory to Practice. Liège, Belgium. October 9, 2014. [Materials]

[9] Accelerating Random Forests in Scikit-Learn. EuroScipy, Cambridge, UK. August 29, 2014. [Materials]

[8] Gradient Boosted Regression Trees in Scikit-Learn. PyData, London, UK. February 23, 2014. [Materials]

[7] Forecasting Daily Solar Energy Production Using Robust Regression Techniques. Atlanta, USA. February 5, 2014. [Materials]

*2013* [6] Scikit-Learn: Machine Learning in the Python ecosystem. NIPS Workshop on Machine Learning Open Source Software, Lake Tahoe, USA. December 10, 2013. [Materials]

[5] Understanding variable importances in forests of randomized trees. NIPS, Lake Tahoe, USA. December 8, 2013. [Materials]

[4] Scikit-Learn, or why I joined an open source software project. University of Liège, Belgium. October 30, 2013. [Materials]

*2012* [3] Ensembles on Random Patches. ECML, Bristol, UK. September 25, 2012. [Materials]

*2011* [2] Large-scale machine learning for collaborative filtering. Groupe de contact FNRS Calcul Intensif, University of Liège, Belgium. April 28, 2011.

*2010* [1] A zealous parallel gradient descent algorithm. NIPS Workshop on Learning on Cores, Clusters and Clouds, Whistler, Canada. December 11, 2010. [Materials]

## ACADEMIC SUPERVISION

*2016–Present* Manoj Kumar · *New York University (USA)*

Junior data scientist at the Center for Data Science. Sequential model-based optimization.

*2014–2015* Mateusz Susik · *University of Warsaw (Poland)*

CERN intern. Author disambiguation with supervised learning.

*2014–2015* Hussein Al-Natsheh · *University of Lyon 2 (France)*

CERN intern. Author disambiguation with supervised learning.

*2014–2015* Joseph Boyd · *EPFL (Switzerland)*

CERN intern. Text mining with Machine Learning.

## OPEN SOURCE SOFTWARE

*2016–Present* Scikit-Optimize · `https://scikit-optimize.github.io`

Founder of Scikit-Optimize, a Python library for sequential model-based optimization.

*2011–Present* Scikit-Learn · `https://scikit-learn.org`

Core developer of Scikit-Learn, a machine learning library written in Python.

## MISCELLANEOUS

*Awards*

2016 · Best paper award for *Ethnicity sensitive author disambiguation using semi-supervised learning* at KESW'2016

2015 · AIM Prize for best PhD's thesis

2010 – 2014 · F.R.S.-FNRS research fellow scholarship

2010 · Melchior Salier Award for best Master's Thesis

2010 · Baudouin Elleboudt Award for best Master's Thesis

*Languages*

FRENCH · Mothertongue

ENGLISH · Fluent

DUTCH · Basic

December 19, 2016