Marine Le Morvan

✉ marine.le-morvan@inria.fr

Main research fields

Machine learning, Computational biology, genomics, missing values.

Research experience

may 2019 – now **Postdoctoral researcher, Machine Learning**, *Parietal group - INRIA*, Saclay.
- Developing methods for supervised learning in the presence of missing values.
- Collaboration with Gaël Varoquaux, Julie Josse, Erwan Scornet.

2014 – 2018 **PhD candidate, Machine Learning and Computational Biology**, *Mines Paristech*, Paris.
- Devoloped a scalable optimization algorithm for sparse linear models with interactions. Applied it to trait prediction from genome wide genetic variations.
- Proposed a new data normalization method that is coupled to the subsequent prediction task. Applied it to gene expression data.
- Proposed a new representation of cancer mutation data, based on gene-gene interaction graphs, that improves patient stratification and prognosis.
- Advisors: Jean-Philippe Vert, Andrei Zinovyev.

2016 **Simons Institute for the Theory of Computing**, *UC Berkeley*, *3 months*.
Participation upon invitation to the spring 2016 program: *Algorithmic challenges in genomics*.

2014 **R&D intern, Machine Learning**, *Ariana pharma*, Paris, *5 months*.
Developed a predictive algorithm using tumorous molecular and cellular data for the selection of a personalized treatment in oncology.

2013 **Research intern, Biomolecular engineering**, *David Savage laboratory, Energy Biosciences Institute*, Berkeley, *5 months*.
- Constructed genetically encoded fluorescent biosensors for monitoring the presence of acyl-CoAs using the bacterial transcription factor FadR.
- Work congratulated by the Ecole Polytechnique jury of research internships.

Education

2010 – 2014 **Ecole Polytechnique**, Palaiseau.
- MSc in Engineering (Diplome d'ingenieur de l'Ecole Polytechnique).
- Cross disciplinary courses, centered around applied mathematics and biology.
- Double degree with Mines ParisTech biotech program.

2007 – 2010 **Classes preparatoires aux grandes ecoles**, *Lycee Saint Genevieve*, Versailles.
Intensive undergraduate program preparing for the national competitive entrance examination of top French Grandes Ecoles. Mathematics and physics program.

Languages and programming skills

- Languages: French (native), English (fluent), Spanish (advanced)
- Programming languages: Python, R, C++

Publications et Pre-prints

- Le Morvan, Marine, Julie Josse, Thomas Moreau, Erwan Scornet, and Gaël Varoquaux. **NeuMiss networks: differentiable programming for supervised learning with missing values.** *Advances in Neural Information Processing Systems*, 33, 2020a

- Le Morvan, Marine, Nicolas Prost, Julie Josse, Erwan Scornet, and Gael Varoquaux. **Linear predictor on linearly-generated data with missing values: non consistency and solutions**. In Silvia Chiappa and Roberto Calandra, editors, *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pages 3165–3174. PMLR, 2020b

- Marine Le Morvan and Jean-Philippe Vert. **WHInter: A Working set algorithm for High-dimensional sparse second order Interaction models**. In *Proc. 35th Int. Conf. Mach. Learn.*, pages 3635–3644. PMLR, 2018

- Marine Le Morvan, Andrei Zinovyev, and Jean-Philippe Vert. **NetNorM: Capturing cancer-relevant information in somatic exome mutation data with gene networks for cancer stratification and prognosis**. *PLoS Comput. Biol.*, 13(6):e1005573, 2017

- Marine Le Morvan and Jean-Philippe Vert. **Supervised Quantile Normalisation**. *ArXiv e-prints*, 2017

Talks

- **NeurIPS Paris meetup, Paris, France, Décembre 2020**
  NeuMiss networks: differentiable programming for supervised learning with missing values.

- **Séminaire de Probabilités et Statistiques de l'IMAG, Université de Montpellier, France, Décembre 2020**
  NeuMiss networks: differentiable programming for supervised learning with missing values.

- **Séminaire Palaisien, Orsay, France, Novembre 2020**
  NeuMiss networks: differentiable programming for supervised learning with missing values.

- **Centre International de Rencontres Mathématiques (CIRM), Luminy, France, Mars 2020**
  Linear predictor on linearly-generated data with missing values: non consistency and solutions.

- **Laboratoire Paul Painlevé, Université de Lille, France, Janvier 2020**
  Linear predictor on linearly-generated data with missing values: non consistency and solutions.

- **Laboratoire de l'accélérateur linéaire seminar, Université Paris sud, Orsay, France, November 2019**
  Cancer stratification and prognosis from mutations using gene networks.

- **International Conference on Machine Learning (ICML) talk, Stockholm, Sweden, July 2018**
  WHInter: A Working set algorithm for High-dimensional sparse second order Interaction models.

- **PhD defense, Institut Curie, Paris, July 2018**
  Learning from genomic data: efficient representations and algorithms.

- **Statistics seminar P6-P7, Jussieu, Paris, May 2018**
  WHInter: A Working set algorithm for High-dimensional sparse second order Interaction models.

- **Laplace reading group, Ecole Nationale Supérieure (ENS) Paris, March 2018**
  Scaling up the LASSO with interaction features.

- **Tech talk, Google Zurich, Nov. 2017**
  Scaling up the LASSO with interaction features.

- **Algorithmic Challenges in Genomics program, Simons Institute for the Theory of Computing, Berkeley, California, Feb. 2016**
  Survival time prediction from mutation profiles using gene networks.

- **Spotlight, Machine Learning Summer School, Kyoto, Japan, Aug. 2015**
  Representing cancer mutations using gene networks.

Extracurricular Activities

Horse-riding Galop 7 (advanced), Running

2011-2012 Treasurer of X-MicroFinance, a microcredit student association. Participated in the the 2011 campaign in Guatemala where 50.000 euros were lent.