06.40.69.65.65
Montpellier, France
tanguy.lefort@umontpellier.fr

Tanguy Lefort

Ph.D student in statistics

GitHub: tanglef
https://tanglef.github.io

Skills

Tools and Languages	Python, Git, R, LaTeX, JavaScript, HTML, CSS
Statistics	Machine learning algorithms (logistic regression, trees, SVM, dimension reduction, KNN, neural networks like CNN for image classification ...) Convex Optimization (descent methods) with ℓ_1 and ℓ_2 penalty
Communication	Documentation of a project for users. Interactive web applications using shiny/dash/flask.
Languages	English, French, Spanish

Experience

Ph.D student **10-2021 — Present**
Institut Montpelliérain Alexander Grothendieck (IMAG) and INRIA Montpellier *Montpellier, France*

- Under the supervision of Benjamin Charlier (IMAG-CNRS), Joseph Salmon (IMAG-CNRS) and Alexis Joly (INRIA)
- Uncertain labels in classification and expert feedback
- This aims to provide a theoretical framework and new algorithms to control and improve the quality identification in the context of a large-scale cooperative system such as Plantnet. We build on information theory and recent advances in the understanding of deep learning. A crucial point is the intrinsic ambiguity of the data. Whatever the annotation, the images themselves usually contain an intrinsic ambiguity. We focus on problems of data ambiguity in cooperative annotation.

Missions during Ph.D **10-2021 — Present**
IMAG *Montpellier, France*

- Numerical part of the course on Convex Optimization for undergraduate mathematics students. We introduce standard techniques in the field. Codes are in Python and written with Joseph Salmon.
- Teaching assistant for an undergraduate first year biology course in mathematical general concepts: set theory, probabilities, conditional probabilities, random variables with classical distributions.
- Open source collaboration coding mission on the Benchopt library. I participated in the website rendering results visualizations, using HTML, CSS, JavaScipt and Plotly with Python. This also improved local plotting data visualizations.

Intern / Master's thesis **03-2021 — 08-2021**
IMAG *Montpellier, France*

- Master's thesis internship in High dimensional optimization for penalized linear models with interactions using graphics card computational power.
- Supervised by Benjamin Charlier and Joseph Salmon.
- Benchmarking descent methods on linear models with ℓ_1 and ℓ_2 penalties with first order interactions. Adapting strategies for GPU handlers.
- Application to genomics dataset.

Intern **07-2020 — 08-2020**
IMAG *Montpellier, France*

- Participating in the PyKeOps library under the supervision of Benjamin Charlier
- Rewrite and adapt Scipy's Fortran conjugate gradient routine for symbolical matrices in PyKeOps.
- Benchmarking using Ridge-Tikhonov regularization.
- Introduction to optimal transport problems and batch-handling for the GeomLoss package.

Talks

- High dimensional optimization for penalized linear models with interactions using graphics card computational power, at Probability and statistics (EPS) team seminar - Univ. Montpellier (content from my master's thesis internship) 11 2021
- Introduction to neural network with Joseph Salmon, at ML-MTP seminar - Univ. Montpellier. (session 0 for reading group on Deep Learning: a statistical viewpoint) 10 2021
- Paper club Ridge Regularization: an Essential Concept in Data Science by Trevor Hastie with Florent Bascou, at ML-MTP seminar - Univ. Montpellier 04 2021

Education

Master Biostatistics, *Montpellier France* 2019—-2021 Theoretical and applied statistics and probabilities. Classification algorithms. Survival analysis and modelization of populations.
Bachelor in mathematics, *Dijon France* 2016—-2019 Bachelor with honours. First two years in mathematics and informatics. Then specialized in applied mathematics. Bachelors final project on skeletonization algorithm to fulfill a gamma-ray surgery.