Igor Colin

igor.colin@gmail.com — https://igorcolin.github.io/

RESEARCH INTERESTS

Large-scale Machine Learning, Decentralized Optimization, Statistical Learning Theory, Graphs, Bandits.

EDUCATION

2016   **Ph.D. in Machine Learning,** Adapting Machine Learning Techniques to U-statistics.
Télécom ParisTech, Paris. Supervisors: Stephan Clémençon and Joseph Salmon

2013   **M.Sc. in Machine Learning,** Ecole Normale Supérieure, Cachan.

2013   **M.Sc. in Applied Mathematics and Computer Science,** Ecole des Ponts ParisTech.

2010   **B.Sc. in Mathematics,** Ecole des Ponts ParisTech.

PROFESSIONAL EXPERIENCE

2018   **Senior Research Engineer,** Huawei (ongoing).
Senior member in the Paris Machine Learning research team. Use of Machine Learning for projects related to wireless telecommunication networks. Supervision of 2 CIFRE Ph.D. students.

2017   **Post-Doctoral Position,** SIERRA — Inria, École Normale Supérieure (1 year).
Large-scale nonconvex optimization (portfolio optimization), in collaboration with AXA Global Direct. Supervisor: Alexandre d'Aspremont.

2013   **Independent Consulting,** (2 years).
Development of an application for predicting travel times and delays in public transportations.

2013   **Research Internship,** Technicolor (6 months).
Learning influence from cascades of information (EM on a mixture model, SGD, FISTA, submodular maximization). Supervisors: Nidhi Hegde and Francis Bach.

2012   **Internship,** Natixis (1 year).
Stock prediction and portfolio optimization using machine learning techniques (HMM, K-means, supervised learning). Supervisor: Adel Ben Haj Yedder.

INVITED TALKS

2018   **Stable Bounds on the Duality Gap of Finite Sum Problems,** Noah's Ark Forum
How to extend Shapley-Folkman's theorem for general constrained optimization.

2017   **Decentralized optimization and U-statistics,** MAGNET — Inria Lille
Decentralized optimization techniques and their extension to U-statistics.

2016   **Decentralized estimation of pairwise objectives,** Télécom ParisTech
Prensetation of decentralized estimation algorithms and their application on a gossip setting to Huawei's scientific board.

TEACHING

2017   **Practical Introduction to Machine Learning,** Université Paris Dauphine MASH (ongoing)
Introduction to Python and ML libraries (scikit-learn, pandas, *etc.*) and application to data challenges.

2017   **Introduction to Machine Learning,** Télécom Evolution (ongoing)
2-day formations aiming at introducing basic ML theory (ERM paradigm, VC dimension) and methods (KNN, LDA, Random Forests, SVM) to professionals with scientific backgrounds.

2016   **Teacher assistant,** Paris Diderot University Machine Learning master's
Introduction to convex optimization and first order methods for solving the ERM problem.

2016   **Teacher assistant,** Télécom ParisTech
Seminars about martingales, Markov chains, stopping times and ergodicity.

SKILLS

Languages	French (native); English (fluent); Spanish (knowledge).
OS	UNIX/Linux, Windows, Mac OS X.
Code	**Python**, Matlab, HTML/PHP/CSS, SQL, C++.
	Large-scale implementation of machine learning algorithms.

RESEARCH PAPERS

Parallel Contextual Bandits in Wireless Handover Optimization.
Igor Colin, Albert Thomas and Moez Draief. In ICDM 2018.
Study of contextual bandits methods when used in a parallel settings, where one pulls a batch of arms and observes an associated batch of rewards at each time step.

Stable Bounds on the Duality Gap of Finite Sum Minimization Problems.
Thomas Kerdreux, Igor Colin and Alexandre d'Aspremont.
Study of Shapley-Folkman theorem for establishing tighter dual-gap bounds in finite sum minimization problems. Establishment of tighter guarantees using a new vector-valued Bernstein's inequality.

An Approximate Shapley-Folkman Theorem.
Alexandre d'Aspremont and Igor Colin
Study of Shapley-Folkman's theorem and formulation of approximate versions using Hoeffding's inequality and Frank-Wolfe algorithm.

Decentralized Topic Modelling with Latent Dirichlet Allocation.
Igor Colin and Christophe Dupuy, in NIPS 2016, PMPML workshop.
Adaptation of the particular LDA topic model to decentralized optimization. Experiments on synthetic graph datasets, for both synchronous and asynchronous scenario.

Adapting Machine Learning Techniques to U-statistics.
Igor Colin, Ph.D. thesis.
Many machine learning methods are based on the standard ERM setting, where the objective is separable in the observations. Here, we study the case where the empirical risk depends on pair (or more) of observations and show how standard methods shoud be transposed.

Scaling-up Empirical Risk Minimization: Optimization of Incomplete U-statistics.
Stephan Clémençon, Aurélien Bellet and Igor Colin, in JMRL.
Study of a new sampling scheme for adapting stochastic optimization methods to objectives that depend on pair (or d-tuples) of observations.

Gossip Dual Averaging for Decentralized Optimization of Pairwise Functions.
Igor Colin, Aurélien Bellet, Joseph Salmon and Stephan Clémençon, in ICML 2016.
Extension of a decentralized optimization method for objectives that depend on pairs of observations, using an additional observation propagation step. Theoretical guarantees along with empirical results, for both synchronous and fully asynchronous settings.

Extending Gossip Algorithms to Distributed Estimation of U-statistics.
Igor Colin, Aurélien Bellet, Joseph Salmon and Stephan Clémençon, in NIPS 2015 (selected for spotlight).
Introduction of a new gossip algorithm for estimating a U-statistic of degree 2 in both synchronous and fully asynchronous settings.

MISCELLANEOUS

2016 **Workshop Organizer,** Télécom ParisTech, Paris
Workshop on Distributed Machine Learning.