Aurélien Bellet

INRIA Lille - Nord Europe
40 avenue Halley, 59650 Villeneuve d'Ascq, France
☎ *+33 (0) 3 59 35 87 13*
✉ *aurelien [dot] bellet [at] inria [dot] fr*
http://researchers.lille.inria.fr/abellet/

Employment

2015–present **Research Scientist (CR2)**, *INRIA*, Lille, France.
Tenured researcher position. Member of the Magnet (MAchine learninG in information NETworks) research group, a joint team between INRIA and CRIStAL (University of Lille).

2014–2015 (12 months) **Postdoctoral Researcher**, *LTCI UMR CNRS 5141, Télécom ParisTech*, Paris, France.
Member of the STA (Statistics and Applications) group led by Stéphan Clémençon:
- Studied approximations of U-statistics by sampling, with applications to large-scale empirical risk minimization for learning to rank, metric learning and graph inference problems.
- Developed decentralized gossip algorithms for U-statistics estimation and optimization.
- Involved in the industrial chair "Machine Learning for Big Data" (see below).

2013–2014 (18 months) **Postdoctoral Researcher**, *University of Southern California*, Los Angeles, USA.
Member of the SMILE group led by Fei Sha:
- Developed new metric learning techniques that scale to high-dimensional data.
- Proposed distributed, communication-efficient algorithms for large-scale sparse learning.
- Involved in the IARPA-funded BABEL program on automatic speech recognition (see below).

Education

2009–2012 **Ph.D. in Computer Science and Engineering**, *Laboratoire Hubert Curien UMR CNRS 5516, Université Jean Monnet*, Saint-Etienne, France.
- Title: Supervised Metric Learning with Generalization Guarantees.
- Main contributions: an analytical framework for metric learning, and a family of efficient metric learning algorithms for data represented as numerical vectors or structured objects (strings, trees, etc) with theoretical guarantees on their performance.
- Supervisors: Marc Sebban and Amaury Habrard.
- Defense date: December 11, 2012.

2009–2010 **Postgraduate Diploma in Informatics**, *University of Edinburgh*, Edinburgh.
- Specialty: Learning from Data.
- Courses on statistical machine learning, pattern recognition, probabilistic graphical models, reinforcement learning, data mining, information retrieval, computer algebra and algorithmic game theory.

2007–2009 **MSc in Computer Science**, *Université Jean Monnet*, Saint-Etienne.
- Specialties: Machine Learning and Data Mining.
- **Graduate exchange student** 2008-2009 at McMaster University, Canada. Courses on mathematical optimization, finding patterns in strings and dynamic cognitive systems.

2004–2007 **BSc in Computer Science**, *Université Jean Monnet*, Saint-Etienne.

Awards

Ph.D. award My thesis received the best Ph.D. award from the French Association for Artifical Intelligence (AFIA) in 2013.

Involvement in Research Projects

2016-present **PAMELA**, description coming soon.

2016-present **GRASP**, description coming soon.

2016-present **ERC POC SOM**, description coming soon.

2016-present **LEGO**, an Inria@SiliconValley associate team between Inria Magnet and Fei Sha's group (University of California, Los Angeles). The project is about learning appropriate representations of structured texts and applying such representations to complex and structured prediction tasks in natural language processing..

2014-2015 **Machine Learning for Big Data**, a chair led by Stéphan Clémençon and funded by major companies such as Safran, PSA Peugeot Citroën, Criteo and BNP Paris-bas. Besides pursuing academic research on topics such as stochastic optimization and distributed/decentralized learning, I have been involved in collaborations with the industrial partners through the co-organization of a workshop and a data challenge, the co-supervision of internships, etc.

2013-2014 **BABEL**, a IARPA-funded speech recognition program on efficiently building keyword search systems for any new language from limited and noisy data. Within the Lorelei consortium involving 5 universities as well as IBM Research, my research group proposed (among other things) large-scale kernel methods that can match deep neural nets' performance in acoustic modeling.

2009-2013 **LAMPADA**, a project funded by the French National Research Agency (ANR) on learning algorithms, models and sparse representations for structured data.

2009-2013 **PASCAL2**, a European Commission's ICT-funded Network of Excellence for pattern analysis, statistical modeling and computational learning.

Seminar and Invited Talks

2017–2018 Privacy-Preserving Algorithms for Decentralized Collaborative Machine Learning.
Seminar at Inria Sequel, Inria Multispeech, Naver Labs Europe and Alan Turing Institute. Invited talk at the Russian-French Workshop in Big Data and Applications and the EPFL-Inria Workshop.

2017 A Decentralized and Robust Protocol for Private Averaging over Highly Distributed Data.
Seminar at CRIStAL Lille (DaTinG department day). Invited talk at the DALI 2017 Workshop on Fairness and Privacy in Machine Learning.

2017 Decentralized Estimation and Optimization of Pairwise Functions.
Seminar at SIGMA (Centrale Lille).

2016–2017 Decentralized Collaborative Learning of Personalized Models over Networks.
Seminar at Inria Magnet and LPD Lab (EPFL). Invited talk at the Workshop on Distributed Machine Learning (Télécom ParisTech) and at the Journée Apprentissage et Interactions du GdR IA (UPMC).

2016-2017 U-Statistics in Machine Learning: Large-Scale Minimization and Decentralized Estimation.
Seminar at EURA NOVA, Inria Magnet, LaHC Saint-Étienne and Proba/Stat Lille.

2016 Metric Learning for Large-Scale Data.
Seminar Statistical Machine Learning (SMILE) in Paris.

2015–2016 Large-Scale Similarity and Distance Metric Learning.
Seminar at Criteo Labs Paris and Inria Magnet.

2015 Similarity and Distance Metric Learning with Applications to Computer Vision.
Tutorial at ECML/PKDD 2015 (with M. Cord).

2015 Scaling-up Empirical Risk Minimization: Optimization of Incomplete U-statistics.
Invited talk at International Workshop on Machine learning, Optimization and Big Data (MOD 2015).

2014–2015 The Frank-Wolfe Algorithm: Recent Results and Applications to High-Dimensional Similarity Learning and Distributed Optimization.
Seminar at LIP6 Paris 6, LaHC Saint-Étienne, Heudiasyc Compiègne, LIF Marseille, Inria Magnet, LIG Grenoble, CEREMADE Paris Dauphine, IMT Toulouse, LRI Inria/Paris Sud, AgroParisTech, Séminaire Parisien de Statistique and ENS/Inria Paris. Invited talk at ICML 2015 workshop "Greed is Great".

2014 Metric Learning (and incidentally some distributed optimization).
Seminar at LTCI (Télécom ParisTech), Paris.

2013 Tutorial on Metric Learning.
Course/tutorial at the Université Catholique de Louvain for the CIL Doctoral School.

2012–2013 Supervised Metric Learning with Generalization Guarantees.
Seminar at Machine Learning Group Louvain, ENS/Inria Paris, LIRIS Lyon, LITIS Rouen, LIG Grenoble and USC Machine Learning.

Teaching

2017 **Advanced Machine Learning**, *Cycle Master*, Télécom ParisTech.
Advanced machine learning methods : metric learning, large-scale kernel methods, graph-based learning. 18 hours of lectures and lab sessions (taught in French).

2017 **Certificat d'Études Spécialisées Data Scientist**, Télécom ParisTech.
Introduction to supervised learning, Support Vector Machines. 14 hours of lectures and lab sessions (taught in French).

2017 **Lecture on Learning Similarity and Distance Functions**, Pre-doc Summer School on Learning Systems.
Introduction to similarity and distance metric learning at a summer school targeting pre-doc students organized at ETH Zurich. 1.5 hours of lectures (taught in English).

2016 **Lecture on Graph Mining**, Allianz France.
Introduction to graph mining and machine learning in graphs targeted to Allianz professionals, as part of a series of 7 lectures on data science. 3 hours of lectures (taught in French).

2016 **Advanced Machine Learning**, *Cycle Master*, Télécom ParisTech.
Advanced machine learning methods : distributed optimization, metric learning, large-scale kernel methods, graph-based learning. 21 hours of lectures and lab sessions (taught in French) and organization of a data challenge with the company Morpho on the topic of rank aggregation for facial recognition.

2016 **Certificat d'Études Spécialisées Data Scientist**, Télécom ParisTech.
Introduction to supervised learning, Support Vector Machines. 10 hours of lectures and lab sessions (taught in French).

2015 **Advanced Machine Learning**, *Cycle Master*, Télécom ParisTech.
Advanced machine learning methods : stochastic approximation, distributed algorithms, structured prediction, large-scale kernel methods. 15 hours of lab sessions (taught in French) and organization of a data challenge with the company Morpho on the topic of metric learning for facial recognition.

2015 **Machine Learning and Data Mining**, *Cycle Master*, Télécom ParisTech.
Introduction to machine learning and data mining : Python programming, k nearest neighbors, perceptron, classification and regression trees, large margin and kernel methods, ensemble learning and unsupervised learning. 24 hours of lab sessions, taught in French.

2012 **Pattern Recognition**, *Master Erasmus-Mundus CIMET (Color in Informatics and Media Technology)*, Université Jean Monnet – Saint-Etienne.
Introduction to convex optimization, support vector machines (SVM) and their applications. 20 hours of lectures, tutorials and lab sessions, taught in English. In charge of the course unit.

2010-2012 **Design and Analysis of Algorithms**, *Master Erasmus-Mundus CIMET (Color in Informatics and Media Technology)*, Université Jean Monnet – Saint-Etienne.
Introduction to `C` programming, divide-and-conquer algorithms, dynamic programming, graph algorithms. 66 hours of lab sessions with student projects, taught in English.

2012 **Object-Oriented Programming**, *Licence professionnelle ATII par alternance*, IUT de Saint-Etienne.
Introduction to object-oriented programming and Java. 21 hours of lectures, tutorials and lab sessions taught in French to students in cooperative education (work-based learning). In charge of the course.

2011 **Introduction to Office Tools**, *Licence Sciences Technologie Santé 1ère année*, Université Jean Monnet – Saint-Etienne.
Introduction to information technology : Internet, word processing, HTML, spreadsheet and presentation applications, databases. 12 hours of lab sessions, taught in French.

Other Professional Activities

Research supervision
- Ph.D. student: Robin Vogel (2017-, CIFRE thesis).
- Master's thesis: Maxime Flauder (ENS Paris, 2015), Robin Vogel (ENSAE, 2016), Pierre Dellenbach (École Polytechnique, 2016), Paul Vanhaesebrouck (École Polytechnique, 2016 - best internship award), Thibault Liétard (ENS Rennes, 2016).

Software Publicly available open-source implementations of GESL and SCML, two learning algorithms proposed in my ECML/PKDD 2011 and AAAI 2014 papers.

Reviewer
- Journals: JMLR, MLJ, IEEE TPAMI, IEEE TKDE, IEEE CYB, Neurocomputing, SIGPRO.
- Conferences: ICML 2015 2016 2017, NIPS 2016 2017, AISTATS 2017, ECML/PKDD 2011 2013 2014, IJCAI 2013 2016.
- Workshops: MOD 2015 2016 2017.

Publications

Books and Book Chapters

[1] A. Bellet, A. Habrard, and M. Sebban. *Metric Learning*. Morgan & Claypool Publishers, 2015.

Articles in International Journals

[2] A. Bellet, J. F. Bernabeu, A. Habrard, and M. Sebban. Learning Discriminative Tree Edit Similarities for Linear Classification — Application to Melody Recognition. *Neurocomputing*, 214:155–161, 2016.

[3] S. Clémençon, I. Colin, and A. Bellet. Scaling-up Empirical Risk Minimization: Optimization of Incomplete U-statistics. *Journal of Machine Learning Research*, 17(76):1–36, 2016.

[4] A. Bellet and A. Habrard. Robustness and Generalization for Metric Learning. *Neurocomputing*, 151(1):259–267, 2015.

[5] A. Bellet, A. Habrard, E. Morvant, and M. Sebban. Learning A Priori Constrained Weighted Majority Votes. *Machine Learning*, 97(1–2):129–154, 2014.

[6] A. Bellet, A. Habrard, and M. Sebban. Good edit similarity learning by loss minimization. *Machine Learning*, 89(1):5–35, 2012.

[7] A. Bellet, M. Bernard, T. Murgue, and M. Sebban. Learning state machine-based string edit kernels. *Pattern Recognition*, 43(6):2330–2339, 2010.

Articles in Peer-Reviewed International Conferences

[8] A. Bellet, R. Guerraoui, M. Taziki, and M. Tommasi. Personalized and Private Peer-to-Peer Machine Learning. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*. 2018.

[9] P. Vanhaesebrouck, A. Bellet, and M. Tommasi. Decentralized Collaborative Learning of Personalized Models over Networks. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*. 2017.

[10] I. Colin, A. Bellet, J. Salmon, and S. Clémençon. Gossip Dual Averaging for Decentralized Optimization of Pairwise Functions. In *International Conference on Machine Learning (ICML)*. 2016.

[11] Z. Lu, D. Guo, A. Bagheri Garakani, K. Liu, A. May, A. Bellet, L. Fan, M. Collins, B. Kingsbury, M. Picheny, and F. Sha. A Comparison Between Deep Neural Nets and Kernel Acoustic Models for Speech Recognition. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2016.

[12] G. Papa, A. Bellet, and S. Clémençon. On Graph Reconstruction via Empirical Risk Minimization: Fast Learning Rates and Scalability. In *Annual Conference on Neural Information Processing Systems (NIPS)*. 2016.

[13] A. Bellet, Y. Liang, A. Bagheri Garakani, M.-F. Balcan, and F. Sha. A Distributed Frank-Wolfe Algorithm for Communication-Efficient Sparse Learning. In *SIAM International Conference on Data Mining (SDM)*. 2015.

[14] S. Clémençon, A. Bellet, O. Jelassi, and G. Papa. Scalability of Stochastic Gradient Descent based on "Smart" Sampling Techniques. In *INNS Conference on Big Data (INNS-BigData)*. 2015.

[15] I. Colin, A. Bellet, J. Salmon, and S. Clémençon. Extending Gossip Algorithms to Distributed Estimation of U-statistics. In *Annual Conference on Neural Information Processing Systems (NIPS)*. 2015. **Selected for spotlight presentation (4% acceptance rate)**.

[16] K. Liu, A. Bellet, and F. Sha. Similarity Learning for High-Dimensional Sparse Data. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*. 2015.

[17] G. Papa, S. Clémençon, and A. Bellet. SGD Algorithms based on Incomplete U-statistics: Large-Scale Minimization of Empirical Risk. In *Annual Conference on Neural Information Processing Systems (NIPS)*. 2015.

[18] Y. Shi, A. Bellet, and F. Sha. Sparse Compositional Metric Learning. In *AAAI Conference on Artificial Intelligence*. 2014.

[19] A. Bellet, A. Habrard, and M. Sebban. Similarity Learning for Provably Accurate Sparse Linear Classification. In *International Conference on Machine Learning (ICML)*. 2012.

[20] A. Bellet, A. Habrard, and M. Sebban. An Experimental Study on Learning with Good Edit Similarity Functions. In *IEEE International Conference on Tools with Artificial Intelligence (ICTAI)*. 2011.

[21] A. Bellet, A. Habrard, and M. Sebban. Learning Good Edit Similarities with Generalization Guarantees. In *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD)*. 2011. **Among the 10 papers selected for an extended version in *Machine Learning Journal***.

Articles in Peer-Reviewed International Workshops

[22] A. Bellet, R. Guerraoui, M. Taziki, and M. Tommasi. Personalized and Private Peer-to-Peer Machine Learning. In *NIPS Workshop on Machine Learning on the Phone and other Consumer Devices*. 2017.

[23] T. L. Van, A. Bellet, and J. Ramon. Decentralized and Privacy-Aware Learning of Traversal Time Models. In *ECML/PKDD workshop on Data Mining with Secure Computation*. 2017.

[24] P. Dellenbach, J. Ramon, and A. Bellet. A Decentralized and Robust Protocol for Private Averaging over Highly Distributed Data. In *NIPS workshop on Private Multi-Party Machine Learning*. 2016.

[25] A. Bellet, Y. Liang, A. Bagheri Garakani, M.-F. Balcan, and F. Sha. Distributed Frank-Wolfe Algorithm: A Unified Framework for Communication-Efficient Sparse Learning. In *ICML workshop on New Learning Frameworks and Models for Big Data*. 2014.

[26] A. Bellet, A. Habrard, and M. Sebban. Good Similarity Learning for Structured Data. In *NIPS workshop "Beyond Mahalanobis: Supervised Large-Scale Learning of Similarity"*. 2011.

Articles in Peer-Reviewed French Conferences

[27] P. Dellenbach, J. Ramon, and A. Bellet. Un protocole décentralisé et robuste pour le calcul de moyenne sous contrainte de vie privée. In *Conférence Francophone sur l'Apprentissage Automatique (CAp)*. 2017.

[28] P. Vanhaesebrouck, A. Bellet, and M. Tommasi. Apprentissage collaboratif et décentralisé de modèles personnels sur un réseau. In *Conférence Francophone sur l'Apprentissage Automatique (CAp)*. 2017.

[29] I. Colin, A. Bellet, J. Salmon, and S. Clémençon. Un algorithme de Gossip pour l'optimisation décentralisée de fonctions sur les paires. In *Conférence Francophone sur l'Apprentissage Automatique (CAp)*. 2016.

[30] A. Bellet, A. Habrard, E. Morvant, and M. Sebban. Vote de majorité a priori contraint pour la classification binaire : spécification au cas des plus proches voisins. In *Conférence Francophone sur l'Apprentissage Automatique (CAp)*. 2013.

[31] A. Bellet, A. Habrard, and M. Sebban. Apprentissage de bonnes similarités pour la classification linéaire parcimonieuse. In *Conférence Francophone sur l'Apprentissage Automatique (CAp)*. 2012.

[32] A. Bellet, A. Habrard, and M. Sebban. Apprentissage Parcimonieux à partir de Fonctions de Similarité d'Édition (ϵ, γ, τ)-Good. In *Conférence Francophone sur l'Apprentissage Automatique (CAp)*. 2011.

[33] A. Bellet, M. Bernard, T. Murgue, and M. Sebban. Apprentissage de noyaux d'édition de séquences. In *Conférence Francophone sur l'Apprentissage Automatique (CAp)*. 2009. **Best paper award**.

Technical Reports

[34] A. Bellet, R. Guerraoui, M. Taziki, and M. Tommasi. Fast and Differentially Private Algorithms for Decentralized Collaborative Machine Learning. Technical report, arXiv:1705.08435, 2017.

[35] A. May, A. B. Garakani, Z. Lu, D. Guo, K. Liu, A. Bellet, L. Fan, M. Collins, D. Hsu, B. Kingsbury, M. Picheny, and F. Sha. Kernel Approximation Methods for Speech Recognition. Technical report, arXiv:1701.03577, 2017.

[36] W. Zheng, A. Bellet, and P. Gallinari. A Distributed Frank-Wolfe Framework for Learning Low-Rank Matrices with the Trace Norm. Technical report, arXiv:1712.07495, 2017.

[37] Z. Lu, A. May, K. Liu, A. B. Garakani, D. Guo, A. Bellet, L. Fan, M. Collins, B. Kingsbury, M. Picheny, and F. Sha. How to Scale Up Kernel Methods to Be As Good As Deep Neural Nets. Technical report, arXiv:1411.4000, 2014.

[38] A. Bellet, A. Habrard, and M. Sebban. A Survey on Metric Learning for Feature Vectors and Structured Data. Technical report, arXiv:1306.6709, 2013.