

Mokhtar Z. Alaya

ASSISTANT PROFESSOR · STATISTICS MACHINE LEARNING

Laboratory LMAC, University of Technology of Compiègne

Département Génie Informatique (GI), 57 avenue de Landshut, Compiègne Cedex, 60203 France

☎ (+33) 03 44 23 44 65 | ✉ elmokhtar.alaya@utc.fr | 🏠 <http://mzalaya.github.io/> | 📱 mzalaya | 🌐 mzalaya

Professional Experience

Assistant Professor

LABORATORY OF APPLIED MATHEMATICS OF COMPIÈGNE - LMAC

University of Technology of Compiègne

Sep. 2020 – present

Postdoctoral Researcher

LABORATORY OF IT, INFORMATION PROCESSING AND SYSTEMS - LITIS

University Rouen Normandy

Jan. 2019 - Aug. 2020

Postdoctoral Researcher

LABORATORY OF RANDOM MODELLING AND APPLICATIONS - MODAL'X

University Paris Nanterre

Oct. 2017 - Sep. 2018

Temporary Researcher and Teaching Assistant

LABORATORY OF RANDOM MODELLING AND APPLICATIONS - MODAL'X

University Paris Nanterre

Sep. 2016 - Aug. 2017

Temporary Researcher and Teaching Assistant

LABORATORY OF THEORETICAL AND APPLIED STATISTICS - LSTA

University Pierre and Marie Curie

Sep. 2015 - Aug. 2016

PhD. Candidate and Teaching Assistant

LABORATORY OF THEORETICAL AND APPLIED STATISTICS - LSTA

University Pierre and Marie Curie

Oct. 2012 - Aug. 2015

Education

Ph.D. in Statistics Machine Learning

TITLE: SEGMENTATION OF COUNTING PROCESSES AND DYNAMICAL SYSTEMS

University Pierre and Marie Curie

2012 - 2016

- Ph.D. defense in June, 27th 2016 under the supervision of Stéphane Gaïffas and Agathe Guilloux.
- **Comittee:** Pierre Alquier (Examinator, ENSAE), Sylvain Arlot (Examinator, Univ. Paris-Sud), Gérard Biau (Examinator, Univ. Pierre and Marie Curie), Stéphane Gaïffas (Supervisor, École Polytechnique), Agathe Guilloux (Supervisor, Univ. Pierre and Marie Curie), Erwan Le Pennec (Reviewer, École Polytechnique).

Master of Sciences in Statistics

MASTER THESIS: CHANGE-POINT DETECTION IN GAUSSIAN SIGNALS

University Pierre and Marie Curie

2011 - 2012

- Supervision: Stéphane Gaïffas and Agathe Guilloux

Master of Sciences in Probabilities

MASTER THESIS: POISSON ACCESS NETWORKS WITH SHADOWING - STATISTICAL MODELISATION

University Pierre and Marie Curie

2010 - 2011

- Supervision: Bartek Bartłomiej (INRIA-ENS) and Mohamed K. Karray (Orange Labs R&D).

Magisterium of Mathematics

MAGISTERIUM THESIS: BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS AND FINANCIAL MATHEMATICS

University Gabes Tunisia

2008 - 2010

- Supervision: Said Hamadène (University of Maine) et Ibtissem Hdhiri (University Gabes Tunisia)

Distinctions

- 2017 **Postdoctoral Fellowship Laureat of DIM Math Innov Program**, Fondation of Mathematical Sciences of Paris
- 2016 **Ph.D. Applied Mathematics with honors**, University Pierre and Marie Curie
- 2012 **Doctoral Scholarship**, University Pierre and Marie Curie
- 2012 **M.Sc. Statistics, rank-2^{sd}**, University Pierre and Marie Curie

Research

My research topics are machine learning mainly high-dimensional statistical learning, with a particular interest in sparse inference, matrix completion and survival analysis. Currently, I am interested in optimal transport techniques for machine learning applications.

KEY WORDS:

Machine Learning; Deep Learning; Optimal Transport; High-Dimensional Statistics; Sparse Inference; Change-Point Detection; Sur-

SUMMARY OF SOME WORKS:

Supervised Learning

- In the context of high-dimensional supervised learning, a classical (and often necessary) pre-processing consists in standardizing the columns of the matrix of explanatory variables. Another approach is to discretize them, for example by a binarization process called one-hot encoding. In [9], we were interested in combining the trick of one-hot encoding with a new penalization called binarsity. In each group of binary variables resulting from the one-hot encoding of a single raw variable, this penalty uses a variation-total regularization accompanied by an additional linear constraint to avoid collinearity within the groups. Theoretical results are obtained in the form of fast non-asymptotic oracle inequalities verified by generalized linear models, and the numerical performance of this approach is illustrated on several logistic regression datasets.
- In [8] we considered the problem of joint completion of matrices with multiple and heterogeneous data sources. Several cases are studied: a family of exponential laws of the matrix entries, then a family of exponential laws of the noise and a free distribution for the matrix entries. The estimation procedures are based on the minimization of the sum of a matching term and the penalty by the nuclear norm of the whole joint matrix. Theoretical estimation results lead to fast convergence speeds, which are illustrated by numerical experiments.

Change-Point Detection

- To address the problem of detecting multiple change-points by explanatory variables in high dimension, we have introduced in [3] a prognostic method called binacox. The latter is based on the Cox model and combines one-hot coding with the binarsity penalty [9]. Theoretical guarantees are obtained for prediction and estimation via fast oracle inequalities in terms of Kullback-Leibler divergence. The statistical performance of the approach is evaluated on simulated and genetic cancer data, and compared with state-of-the-art methods for threshold detection in survival analysis. Binacox significantly outperforms existing univariate methods for detection in terms of detection and computation time.

Machine Learning with Optimal Transport

- In [7] we proposed a new algorithm SCREENKHORN to efficiently approximate the OT distance. The algorithm combines a pre-processing step to identify dual variables with an L-BFGS-B algorithm step constrained on the identified variables. The authors illustrated the effectiveness of SCREENKHORN on complex tasks such as dimensionality reduction and domain adaptation in a deep learning context.
- In [5] we considered the partial optimal transport problem, which arises when the considered distributions do not have the same total mass to be transported, a setting in which a classical EO is no longer valid. They considered the partial EO of Wasserstein and Gromov-Wasserstein for which they proposed an algorithm based on Frank-Wolfe optimization to compute the partial transport scheme. Empirically, partial OT is shown to be effective in positive unlabeled learning (PU-learning).
- In [1] we addressed the problem of aligning distributions whose supports do not necessarily lie in the same metric space. They considered a method of projecting the measured metric spaces onto a common Euclidean space, for which the EO is computed on the projected distributions. This leads to a robust Wasserstein sub-embedding distance (SERW). It is shown that the SERW distance defines a proper distance and behaves like the Gromov-Wasserstein distance. The performance of the new SERW distance is illustrated on real applications such as matching and alignment in text-image data.
- In [6] we developed a two-stage optimal transport approach that maps from a source distribution to a target distribution. The target has the particularity of presenting new classes not present in the source domain. The first step of the procedure aims at rejecting the samples from these new classes using an optimal transport scheme. The second step solves the change in the target (class ratio) still as an optimal transport problem. The optimization problem is solved in each step of the proposed method and this two-step approach leads to results that outperform recent literature.
- In [4] I contributed to the python library for optimal transport called POT¹. POT implements several key ideas in optimal transport for the statistical learning community. It contains implementations of a number of seminal EO works for machine learning, such as the Sinkhorn algorithm and Wasserstein barycenters, but also provides generic solvers that can be used to conduct new fundamental research. POT is open source with an MIT license.

Publications

-
- [1] Theoretical Guarantees for Bridging Metric Embedding and Optimal Transport** Neurocomputing
M. Z. ALAYA, M. BÉRAR, G. GASSO, A. RAKOTOMAMONJY 2022
- [2] Optimal Transport for Conditional Domain Matching and Label Shift** Machine Learning
A. RAKOTOMAMONJY, R. FLAMARY, G. GASSO, M. Z. ALAYA, M. BÉRAR, N. COURTY 2021
- [3] Automatic Cut-Points Estimation in High-Dimensional Cox Model** Biometrics
S. BUSSY, M. Z. ALAYA, A. GUILLOUX, A. S. JANNOT 2019
- [4] POT: Python Library for Optimal Transport in Machine Learning** Journal of Machine Learning Research
R. FLAMARY N. COURTY, A. GRAMFORT, M. Z. ALAYA, A. BOISBUNON, S. CHAMBON AND L. CHAPEL, A. CORENFLOS, K. FATRAS, N. FOURNIER, L. GAUTHERON, N. T.H. GAYRAUD, H. JANATI, A. RAKOTOMAMONJY, I. REDKO, A. ROLET, A. SCHUTZ, V. SEGUY, D. J. SUTHERLAND, R. TAVENARD, A. TONG, T. VAYER 2021
- [5] Partial Optimal Transport with Applications on Positive Unlabeled Learning** NeurIPS 2020 Proceedings
L. CHAPEL, M. Z. ALAYA, G. GASSO 2020

- [6] Open Set Domain Adaptation using Optimal Transport** ECML-PKD 2020 Proceedings
M. KECHAOU, R. HÉRAULT, M. Z. ALAYA, G. GASSO 2020
- [7] Screening Sinkhorn Algorithm for Regularized Optimal Transport** NeurIPS 2019 Proceedings
M. Z. ALAYA, M. BÉRAR, G. GASSO, A. RAKOTOMAMONJY 2019
- [8] Collective Matrix Completion** Journal of Machine Learning Research
M. Z. ALAYA, O. KLOPP 2019
- [9] Binarisity: a Penalization for One-Hot Encoded Features in Linear Supervised Learning** Journal of Machine Learning Research
M. Z. ALAYA, S. BUSSY, S. GAIFFAS, A. GUILLOUX 2019
- [10] Learning the Intensity of Time Events with Change-Points** IEEE Trans. on Information Theory
M. Z. ALAYA, S. GAIFFAS, A. GUILLOUX 2015

Preprints

- [1] Adversarial Semi-Supervised Domain Adaptation for Semantic Segmentation** arXiv
M. KCHAOU, M. Z. ALAYA, R. HERAULT, G. GASSO 2022
- [2] Statistical and Topological Properties of Gaussian-Smoothed Sliced Probability Divergences** arXiv
A. RAKOTOMAMONJY, M. Z. ALAYA, G. GASSO, M. BÉRAR 2022
- [3] Heterogenous Wasserstein Discrepancy for Incomparable Distributions** arXiv
A. RAKOTOMAMONJY, M. Z. ALAYA, G. GASSO, M. BÉRAR 2022
- [4] High-Dimensional Time-Varying Aalen and Cox Models** arXiv
M. Z. ALAYA, T. ALLART, A. GUILLOUX, S. LEMLER 2018

Service of Review

- 2020- International Conference on Neural Information Processing Systems - NeurIPS
- 2020- International Conference on Artificial Intelligence and Statistics - AISTATS
- 2022- International Conference on Learning Representations - ICLR
- 2020- Journal of Machine Learning Research - JMLR
- 2021- IEEE Transactions on Pattern Analysis and Machine Intelligence - PAMI
- 2022- Neurocomputing
- 2022- Biometrical Journal

Manuscripts

- Segmentation of Counting Processes and Dynamical Systems** University Pierre and Marie Curie
DOCTORAL THESIS 2016
- Change-Point Detection in Gaussian Signals** University Pierre and Marie Curie
MASTER THESIS 2012
- Poisson Access Networks with Shadowing - Statistical Modelisation and Statistical Inference** University Pierre and Marie Curie
MASTER THESIS 2011
- Backward Stochastic Differential Equations and Financial Mathematics** University Gabes Tunisia
MAGISTERIUM THESIS 2010

Talks

- MLMDA Group - Borelli Center ENS-Paris-Sacaly** Paris-Sacaly
SEMINAR June, 2022
- Laboratory LMM - University Le Mans** Zoom
SEMINAR Mars, 2021

NeurIPS CONFERENCE	Zoom December, 2020
Data Sciences UTC SEMINAR	Zoom November, 2020
Laboratory CMAP - École polytechnique SEMINAR	Zoom June, 2020
Laboratory CREST - ENSAI Rennes SEMINAR	Zoom May, 2020
NeurIPS CONFERENCE	Vancouver, Canada December, 2019
Summer School on Applied Harmonic Analysis and Machine Learning WORKSHOP	Genoa, Italy September, 019
GDR ISIS/MIA MEETING WORKSHOP	CNRS, Paris July, 2019
Startup LumenAI SEMINAR	Paris September, 2018
4th International Society for NonParametric Statistics CONFERENCE	Salerno, Italy June, 2018
50ème Journées Françaises de Statistique CONFERENCE	Paris Saclay May, 2018
Laboratory Modal'X SEMINAR	University Paris Nanterre May, 2017
48ème Journées Françaises de Statistique CONFERENCE	Montpellier June, 2016
47ème Journées Françaises de Statistique CONFERENCE	Lille June, 2015
46ème Journées Françaises de Statistique CONFERENCE	Rennes June, 2014
Laboratory LSTA SEMINAR	University Pierre and Marie Curie Mars, 2015

Teaching

Machine Learning LECTURER	University of Technology of Compiègne Spring 2021, 2022
Multivariate Functions and Applications LECTURER	University of Technology of Compiègne Spring 2022
Probability and Applications TEACHING ASSISTANT	University of Technology of Compiègne Autumn 2021
Linear Algebra and Applications TEACHING ASSISTANT	University of Technology of Compiègne Autumn 2020, 2021, 2022
Statistics for Psychology TEACHING ASSISTANT, LICENCE 1 AND 2	University Paris Nanterre 2016 - 2017
Mathematical for Economics TEACHING ASSISTANT, LICENCE 1	University Paris Nanterre 2016 - 2017

Certificate of Computer Sciences

TEACHING ASSISTANT, LICENCE 2 ET 3

University Paris Nanterre

2016 - 2017

Times Series

TEACHING ASSISTANT, MASTER 1

University Pierre and Marie Curie

2015 - 2016

Mathematical Statistics

TEACHING ASSISTANT, LICENCE 3

University Pierre and Marie Curie

2015 - 2016

Multivariate Regression

TEACHING ASSISTANT, MASTER 1

University Pierre and Marie Curie

2013 - 2014

Linear Algebra and Geometry

TEACHING ASSISTANT, LICENCE 2

University Pierre and Marie Curie

2012 - 2014

Probabilities and Statistics

LECTURER, 3-TH YEAR MECHANICAL ENGINEERING

University Pierre and Marie Curie

2012 - 2013

Skills

Scientific Programming

Python (2/3), Jupyter Notebook, R, Matlab, C++

Librairies Python

Scikit-Learn, Pandas, Numpy, Scipy, Matplotlib, Bokeh, Cython, Pytorch, TensorFlow, MFlow,

Operations Systems

GNU/Linux, Mac OS X, Git, PyCharm, CLion

Typography

TeX, Microsoft Office, LibreOffice