Mohamed El Amine Seddik

Ph.D. Student at CEA & Centralesupélec
Random Matrix Theory, Machine Learning and Deep Learning

Age: 26
26 Cours Pierre Vasseur
91120 Palaiseau, France
+33 (0) 6 15 30 21 46
melaseddik@gmail.com
melaseddik.github.io

Education

Nov. 2017 - curr. **Ph.D. in Computer-Science**, *CEA LIST and Centrale-Supelec (University of Paris-Saclay)*, France, Thesis: *Random Matrix Theory for AI: from Theory to Practice*.

July 2019 **Gretsi Summer School**, *Peyresq*, France, Courses: Differential and Riemannian Geometry, Manifolds and Information Geometry.

June 2018 **Data Science Summer School (DS3)**, *Ecole Polytechnique*, France, Tutorials: Representing and Comparing Probabilities with Kernels, Optimal Transport and Machine Learning.

Sep. 2016 - mar. 2017 **Third Year of Engineering School in Data-Science**, *Telecom ParisTech; Ecole Polytechnique and ENS Cachan*, Paris, France, Courses: Machine Learning, Deep Learning, Reinforcement Learning, Graphical Models, Optimization, Random Matrix Theory and Image Processing (Masters DATASCIENCE and MVA).

Jan. 2016 - Mai 2016 **Second Year of Engineering School**, *Ecole de Technologie Superieure de Montreal*, Montreal, Canada, Courses: Artificial Intelligence, Image Processing, Probability and Statistics, Wireless Communication Systems and Projects Management.

Sep. 2014 - Jun. 2015 **First Year of Engineering School**, *IMT Lille-Douai - University of Lille*, Villeneuve d'Ascq, France, Courses: Information Theory, Signal Processing, Algorithmic, Parametric Estimation, Measure Theory and Probability Theory.

2012-2014 **Two-year highly selective classes to prepare for the competitive exams to the "Grandes Ecoles" on Mathematics and Physics (MPSI-MP*)**, *Lycee Pothier*, Orleans, France.

2011-2012 **First Year of University Technology Degree (DUT)**, *Iut de Chartres*, France.

Scientific & Computer Skills

Maths Random Matrix Theory, Linear Algebra, Functional Analysis, Probability and Statistics, Statistics in High-dimensions, Stochastic Processes, Optimization.

Machine-Learning Kernel methods, Logistic regression, Boltzmann Machine, Deep Learning (Multi-Layer Perceptron, Convolutionnal and Recurrent Neural Networks), Transfer-Learning, Domain Adaptation, Dim. reduction, Sparsity, Clustering, others.

OS Windows, Linux, Contiki, Riot, OpenWSN.

Languages MATLAB, PYTHON (CERTIFIED), C/C++, JAVA, JESS, PROLOG, VHDL, LATEX.

Softwares Labview, Simulink, KiCad, Arduino, Visual Studio, Photoshop, Inkscape.

Libraries Machine-Learning: Scikit-Learn, LibLINEAR, LibSVM.

Deep-Learning and Computer-Vision: TensorFlow, Keras, PyTorch, OpenCV.

Languages French (bilingual), Arabic (bilingual), English (fluent), German and Spanish (basics).

References Contact Details

Romain Couillet, Full Professor at CentraleSupélec, University of ParisSaclay.
Webpage: `https://romaincouillet.hebfree.org/`

Laurent Clavier, Full Professor at IMT-Lille-Douai, University of Lille.
Webpage: `http://pagesperso.telecom-lille.fr/clavier/`

Nathalie Mitton, Scientific head of the FUN research group at Inria Lille-Nord Europe.
Webpage: `http://researchers.lille.inria.fr/~mitton/`

Publications

2020

ICML'20 Submitted ***A Random Matrix Analysis of Learning with α-Dropout***, MEA.Seddik, R.Couillet, M.Tamaazousti, Workshop "The art of learning with missing values".

NeurIPS'20 Submitted ***The Unexpected Deterministic and Universal Behavior of Large Softmax Classifiers***, MEA.Seddik, C.Louart, R.Couillet, M.Tamaazousti.

NeurIPS'20 Submitted ***Neural Networks Classify through the Class-wise Means of their Representations***, MEA.Seddik, M.Tamaazousti.

ICIP'20 ***Lightweight Neural Networks from PCA & LDA Based Distilled Dense Neural Networks***, MEA.Seddik, H.Essafi, A.Benzine, M.Tamaazousti.

ICML'20 ***Random Matrix Theory Proves that Deep Learning Representations of GAN-data Behave as Gaussian Mixtures***, MEA.Seddik, C.Louart, M.Tamaazousti, R.Couillet.

2019

Gretsi'19 ***Why do Random Matrices Explain Learning? An Argument of Universality Offered by GANs***, MEA.Seddik, M.Tamaazousti, R.Couillet.

ICASSP'19 ***Kernel Random Matrices of Large Concentrated Data: The Example of GAN-generated Images***, MEA.Seddik, M.Tamaazousti, R.Couillet.

ICLR'19 ***A Kernel Random Matrix-Based Approach for Sparse PCA***, MEA.Seddik, M.Tamaazousti, R.Couillet.

Neuro computing'19 ***Generative Collaborative Networks for Single Image SuperResolution***, MEA.Seddik, M.Tamaazousti, J.Lin.

SCIA'19 ***Deep Multi-class Adversarial Specularity Removal***, J.Lin, MEA.Seddik, M.Tamaazousti, A.Bartoli, Y.Tamaazousti.

TPAMI'19 ***Learning More Universal Representations for Transfer-Learning***.
(Journal) Y.Tamaazousti, H.Le Borgne, C.Hudelot, MEA.Seddik, M.Tamaazousti

2018

WCNC'18 ***From Outage Probability to ALOHA MAC Layer Performance Analysis in Distributed WSNs***, MEA.Seddik, V.Toldov, L.Clavier, N.Mitton.

Patents

2020 **Statistically robust transfer learning schemes**.

2020 **Distillation methods for neural networks**.

2019 **Image super-resolution using generative collaborative networks**.

2018 **Time consistent estimation of scene illumination using improved sequential Monte Carlo estimate with reduced variance**.

Detailed Experience

Ph.D. Thesis

Oct. 2017 - current **Random Matrix Theory for Big-Data and Machine Leaning**, *CEA LIST and University of Paris-Saclay*.

Supervisors Mohamed Tamaazousti (CEA LIST) and Romain Couillet (Centrale-Supelec)

Description The bigdata era has driven the recent development of new algorithms and methods, often based on elementary principles allowing to handle large amounts of data. However, these large dimensional data impair the behavior of traditional methods that deserve to be revisited under the eye of more elaborate tools and methods. A better understanding of these methods in the bigdata regime indeed induces possibilities of improvements, thereby leading to the development of more efficient algorithms. The Random Matrix framework provides a powerful tool to understand and analyse the behavior of simple data models (such as the mixture of Gaussians model) in the large dimensional setting, which is naturally the case in the BigData paradigm. My PhD thesis will aim at going beyond the simple models hypothesis, to develop new methods that are more appropriate to practical datasets (structured data, images, etc.).

Tutorials & Talks

2019 **Kernel Random Matrices of Large Concentrated Data: the Example of GAN-Generated Images**, Ecole Normale Superieure (ENS Ulm), Paris, France.

Description Talk at the weekly Golosino seminar of ENS on: Kernel Random Matrices of Large Concentrated Data: the Example of GAN-Generated Images.

2019 **Tutorial on Random Matrix Advances in Large Dimensional Statistics, Machine Learning and Neural Nets**, *by R. Couillet, MEA. Seddik, and Malik Tiomoko*, EUSIPCO 2019, Coruna, Spain, 2019.

Description Kernel random matrices and applications (analysis and improvement) to sampling, kernels and sparse covariance estimation, spectral clustering, semi-supervised learning; random feature maps, neural networks; universality considerations.

Dec. 2018 **Talk at ICTP about Random Matrix Theory for Big-Data and Machine Learning**, *by MEA. Seddik*, Machine Learning Landscape - a workshop at the interface of string theory, machine learning and energy landscapes, Trieste, Italie, 2018, `http://indico.ictp.it/event/8784/session/27/contribution/183`.

Teaching

2019 - curr. **Introduction to Machine Learning**, *Centrale-Supelec*.

Description Teaching assistant of the "Introduction to Machine Learning" course (2nd year students at Centrale-Supelec).

2018 - curr. **Python Coding Weeks**, *Centrale-Supelec*.

Description Implementation of machine learning methods and tools on python.

2018 - curr. **Practical Signal-Processing**, *Centrale-Supelec*.

Description Signal Processing tutorials for first grader engineering students. Temporal Representation, Frequency Representation, Fourrier Transform, Signal Filtering, Sampling, Audio and Image Signals Analysis etc.

Master's Thesis

2014 - 2015 **Random Matrix Theory for Sparse Principal Components Analysis in High-dimension**, *CEA LIST*.

Instructors Mohamed Tamaazousti (CEA LIST) and Romain Couillet (Centrale-Supelec)

Description Consistent sparse principal components estimation using *Random Matrix Theory* and *Concentration of Measure* in the high-dimensional setting: (i) Non-linear deformation of the empirical covariance matrix; (ii) Asymptotic equivalent; (iii) Exploitation of the sparsity property of the principal components. [Matlab, LaTeX]

Projects & Internships

2017 **Variance Reduction in Monte Carlo Methods**, *Telecom ParisTech*.

Instructor Francois Portier (Telecom ParisTech)

Bibliography on Monte Carlo methods, variance reduction using functional control variables and reduction of dimension for the multi-dimensional case. [Linear Algebra, Probability and Statistics, Matlab]

2017 **Human Activities Classifation using Kernel Methods**, *Telecom ParisTech*.

Application of machine learning methods (in particular, kernel methods) to the problem of physical humain activities classification, using heterogeneous data from position sensors. [Python, Scikit-Learn]

2016 **Multi-Target Tracking**, *ETS Montreal*.

Instructor Rita Noumeir (ETS Montreal)

Bibliography and implementation of a multi-target tracking algorithm, silhouette detection, SVM classification and identification of targets. [Python, Matlab]

2016 **Neural Networks**, *ETS Montreal*.

Implementation of a document prediction and classification algorithm based on neural networks. Implementation from scratch of the Back-propagation algorithm. [Python, Scikit-Learn]

2015 **MAC Layer Performance Analysis of Distributed Wireless Networks**, *Inria*.

Instructor Nathalie Mitton (Inria Lille)

Study and analysis of the ALOHA MAC layer performances in distributed wireless sensor networks (theory, simulations and experimentations). [C/C++, Matlab,

Labview, Simulink, Contiki, OpenWSN]

2015 **Non-Gaussian Interference Modelling in Wireless Sensor Networks.**, *CNRS*.

Instructor Laurent Clavier (IRCICA CNRS & IEMN Lille)

Non-Gaussian Interference Modelling in Wireless Sensor Networks using Poisson Points Processes. [Matlab, Stochastic Processes, Probability, Information Theory]