Melissa Ailem

PhD

45 rue des Saints-Pères
75006, Paris, France
+33674745979
melissa.ailem@inria.fr
Age: 26
Nationality: French

Work experience

2017 – 2018 **Post-doctoral researcher - University of Southern California (Los Angeles, USA) and INRIA (Lille, France)** .

- **Team:** MAGNET team (INRIA) and Fei Sha's group (USC).

- **Title of the postdoc:** Transfer and multi-modal learning of word representations.

2016 – 2017 **Teaching and Research Assistant - Department of Statistic and Business Intelligence, University of Paris Descartes, Paris, France**.

- **Team:** Machine Learning for Data Science (MLDS).

- **Current research topics:** word embedding, mixture models, matrix factorization, clustering and text mining.

- **Taught classes (192 hours/year):** Advanced Databases, Business Analytics, Introduction to R for Data Visualisation, Python programming for Data Analysis.

2015 – 2016 **Teaching and Research Assistant - Department of Computer Science and Mathematics, University of Paris Descartes, Paris, France**.

Taught classes (192 hours/year): Exploratory and multivariate data analysis - R, Numeration and mathematical logic, Logical Programming - Prolog, Advanced Topics in Linux Operating System, Functional Programming - OCaml.

2013 – 2016 **PhD Student - Machine Learning for Data Science (MLDS) group, Computer Science Laboratory of Paris Descartes University, Paris, France**.

Subject: Sparsity-sensitive Diagonal Co-clustering Algorithms for the Effective Handling of Text Data.

Supervisors: Mohamed Nadif and François Role.

Examining committee: Massih-Reza Amini, Celine Robardet, François Yvon and Mohamed Quafafou.

Completed research in the field of: Text Mining, Biomedical Text Mining, Unsupervised learning, Graph approaches and Probabilistic Mixture Models for co-clustering.

Software: Python package for co-clustering (Coclust 0.2.1) available at : https://pypi.python.org/pypi/coclust

Feb., 2013 - Sep., 2013 **Internship, Biomedical Text Mining - INSERM, Paris, France**.
Text Mining approaches to identify multi-causal disease-susceptibility genes in the area of genome wide association studies.

Skills

Language **English: fluent | French, Kabyle and Arabic: native speaker**

Machine Learning **Unsupervised Learning, Text Mining, Mixture Models, Clustering, Co-clustering, Embedding**

Programming **Python, R, Shell, SQL, Ocaml, Prolog**

Awards, Grants and Scholarships

2017 **Postdoctoral Grant**, *Inria@SiliconValley Postdoctoral Fellowship.*

2017 **Simon Régnier Award**, *Société Francophone de Classification.*

2015 **Grant**, *SIGIR conference.*

2013 **Thesis Scholarship**, *AAP Sorbonne Paris Cité.*

Education

2013 – 2016 **PhD in Computer Science, specialization in Data Science**
Computer Science Laboratory of Paris Descartes University - Paris, France
Supervisors: Mohamed Nadif and François Role.
Subject: Sparsity-sensitive Diagonal Co-clustering Algorithms for the Effective Handling of Text Data.
Start date: October 01, 2013.
Defense date: November 18, 2016.

2012 – 2013 **MSc in Computer Science, specialization in Machine Learning**
Paris Descartes University - Paris, France.
- **Rank:** 1/33 - **Average grade:** 17/20

2011 – 2012 **MSc in Computer Science**
Paris Descartes University - Paris, France.
- **Rank:** 1/95 - **Average grade:** 15/20

2008 – 2011 **BSc in Mathematics and Computer Science - Algeria.**
BSc Project: Development of Natural Language Processing (NLP) approaches : Word-Sense Disambiguation (WSD).

2007 – 2008 **Algerian Baccalaureate, scientific section (Algerian equivalent to A levels) - Algeria.**

Publications

International publications

- A. Salah, M. Ailem and M. Nadif – "Word Co-occurrence Regularized Non-Negative Matrix Tri-Factorization for Text Data Co-clustering". Accepted in AAAI'2018.

- A. Salah, M. Ailem and M. Nadif – "A Way to Boost Semi-NMF for Document Clustering". Accepted in CIKM'2017.

- M. Ailem, A. Salah and M. Nadif – "Non-negative Matrix Factorization meets Word Embedding", SIGIR'2017, p. 1081–1084.

- M. Ailem, F. Role and M. Nadif – "Model-based co-clustering for the effective handling of sparse data", *Pattern Recognition Journal* **72** (2017), p. 108-122.

- M. Ailem, F. Role and M. Nadif – "Sparse Poisson Latent Block Model for Document Clustering", *IEEE Transactions on Knowledge and Data Engineering Journal (TKDE)* **29** (2017), p. 1563-1576.

- M. Ailem, F. Role and M. Nadif – "Graph Modularity Maximization as an Effective Method for Co-clustering Text Data", *Knowledge-Based Systems Journal (KBS)* **109** (2016), p. 160–173.

- M. Ailem, F. Role, M. Nadif and F. Demenais – "Unsupervised Text Mining for Assessing and Augmenting GWAS Results", *Journal of biomedical informatics (JBI)* **60** (2016), p. 252–259.

- M. Ailem, F. Role and M. Nadif – "Co-clustering Document-Term Matrices by Direct Maximization of Graph Modularity", CIKM'2015, p. 1807–1810.

French publications

- M. Ailem, F. Role, M. Nadif and F. Demenais – "Modèles vectoriels de documents pour la fouille de textes bio-médicaux : application à la validation d'études d'associations pan-génomiques (gwas)", SFC'2014, **21** (2014), p. 215–218.

- M. Ailem, F. Role and M. Nadif – "Modèle de Blocs Latents Poissonnien contraint pour la classification de Documents", SFC'2017.

Seminar

2016 **Seminary ISIS @AgroParisTech**, *Machine learning methods and applications to health.*

2015 **Seminary @INSERM**, *Biomedical Text Mining approaches.*

Extracurricular activities

Sports: **Swimming, Running, Football.**

Others: **Cinema, Dance, Programming.**