

Benjamin Muller

Experiences

- 2018-Present **PhD student**, *INRIA the French national research institute for the digital sciences*, Paris.
Domain transfer for non canonical data supervised by Benoit Sagot and Djamel Seddah
- 2018-Present **Teaching assistant**, *Paris-Descartes university and ENSAE*, Paris.
Guiding master's students in their research project in applied statistics. Teaching Master's Students on the broad topic of "Initiation to Research in Statistics"
- 2018 **NLP research engineer**, *INRIA*, Paris, 6 months.
Syntactic parsing for User Generated Content. Parsing 60+ languages from raw data using our own ELMoLex model implemented in pytorch (CoNLL 2018 Shared Task)
- 2017 **NLP research engineer**, *Riminder, Facebook startup garage (1st batch)*, Paris, 6 months.
Design, implement, and evaluate multilingual word embedding model, character based word embedding, and Hierarchical attention mechanisms for Classification using tensorflow and keras
- 2016-2018 **Data Scientist**, *Freelance*.
Designing, implementing and training machine learning tools for R&D and production purposes, advising early stage startups (+3) in their AI strategy
- 2015-2016 **Data Scientist**, *AXA-IM Rosenberg equities*, San Francisco, 12 months.
Experimenting with Feed Forward and Recurrent Neural Networks on earning classification models. Analyzing textual and graph patent data for valuation purposes
- 2014-2015 **Machine Learning Research Project**, *ENSAE*, Paris, 6 months.
Clustering textual data and topic discovery using the Latent Dirichlet implemented from scratch in R for the Center for Research in Economics and Statistics supervised by J. Ridgway

Publications

- UDW **ELMoLex: Connecting ELMo and Lexicon features for Dependency Parsing**, *Oct. 2018*.
- EurNLP **Unsupervised Learning for Handling Code-Mixed Data: A Case Study on POS Tagging of North-African Arabizi Dialect**, *Aug. 2019*.
- Under review **Enhancing BERT for lexical Normalization**, *Aug. 2019*.

Education

- 2016-2017 **MSc in Data Science**, *Paris*, Ecole Polytechnique and the University of Paris-Saclay.
Studying Machine Learning from theory to practice: Statistical Learning Theory, Optimization for Data Science, Deep Learning, Compressed Sensing
- 2013-2017 **ENSAE, Institut Polytechnique de Paris**, *Paris*.
A leading French engineer school in Data Science mainly studying Theory of Probability, Statistics and Optimization
- 2011-2013 **Preparatory School in Maths and Physics**, *Nice*.
Two years undergraduate program leading to national competitive exam. Analysis, Algebra, Fundamental Physics

Languages

- Languages French (native), English (fluent), Italian (intermediate)
- Programming Professional experiences in Python, R and SQL. Familiar with Scala and Spark, MongoDB, Javascript,

Extracurricular Activities

- Reading Scientific culture, philosophy, philosophy of language
- Sport Competitive tennis, Scuba diving, hiking, running