Antoine SIMOULIN, PhD

+1 312-447-4198 antoine.simoulin@gmail.com simoulin.io

Technical Skills: **Python, Git, SQL, Linux/Unix/Shell, Bash, HTML, CSS, Excel, PowerPoint, LaTeX**

EXPERIENCES

Expert Data Scientist, NLP
Quantmetry

5 years April 2017 – August 2022 Paris

Optimized and automated processes for **real-world problems** by contributing to the ideation, framing, deployment, and monitoring phases of **end-to-end data science projects:**

- Reduced the email response time by 15% by **deploying an API** for classifying, summarizing, and automating email replies for one of the largest French insurance using **neural networks implemented in Tensorflow** and **deployed to production using a Jenkins CI/CD to package and host Docker containers on Red Hat OpenShift**;
- Supported R&D team by animating quarterly internal and external **presentations showcasing AI research trends** and its potential business applications, and **mentoring interns** in **implementing research papers** such as a dataset to analyze biases in large language models or a neural pointer network in **PyTorch**;
- Measured the impact of screening on breath cancer pathology in France by creating a dataset from unstructured medical records using **PySpark scripts** and **Natural Language Processing (NLP)** methods, computing **statistics and publishing the study in a national conference**.

Quantitative Analyst Intern, Securitization team
Crédit Agricole Corporate and Investment Banking

1 year September 2015 – August 2016 New York

- Contributed to the set-up of new worldwide securitization operations by conducting **data analysis** using **SQL** and VBA queries on large databases and reporting to notation agencies and internal teams;
- Improved the running time of the capital calculation of internal insurance by **implementing Monte-Carlo's algorithms in CUDA** for execution on graphic cards.

EDUCATION

PhD, Computer Science (NLP)
University Paris Cité

3 years February 2019 - July 2022 Paris

Conducted research advised by Professor Benoit Crabbé, member of LLF lab, focusing on NLP and deep learning methods to build sentence embeddings:

- Pushed forward the state-of-the-art of NLP by designing innovative neural networks following structured patterns inspired by linguistic insights, and **publishing 6 articles in top peer-reviewed NLP conferences with open-source contributions** such as a version of GPT-2 for French with over a billion parameters;
- Shared my knowledge by building and **teaching a graduate level NLP course** including 10 lectures and labs to around 25 students from the mathematics department at University Paris Cité between 2020 and 2022;
- **Contributed to open-source** by developing packages such as PyTree, implementing tree-structured networks, **awarded during the PyTorch Annual Hackathon 2021**, and releasing sentence embedding models pre-trained on 1B sentence pairs, **awarded during the Hugging Face Community week using JAX/Flax for NLP & CV 2021**.

Dual Master Program (MSc), Data Science
Ecole Polytechnique

1 year 2016 - 2017 Paris

The leading French research, academics, and innovation institution. Mathematical and numerical analysis. Statistical and machine learning. Large-scale calculations and distributed databases.

Master of Science (MSc), Simulation and Mathematical Engineering
ENSTA Paris

4 years 2013 – 2017 Paris

Top engineering school in France, accessible through extremely selective *classe préparatoire*. Computer science & mathematics degree with a major in optimization, practical implementations in **C++**. Last year advised by Professor Pierre Carpentier, director of the UMA lab.