

# Maël FABIEN

## Education



### Data Scientist

(22 years old)

Nationality: France, Luxembourg

+ (00) 33 7 87 20 40 33

mael.fabien@gmail.com

10 rue du Roi de Sicile, 75004  
Paris

### Bachelor of Science

Economics  
HEC Lausanne



Economics, Statistics,  
Financial Mathematics

2013 - 2016

### Mater of Science

Actuarial Science  
HEC Lausanne



Statistics, Mathematics,  
Stochastic Processes

2016 - 2018

### Specialized Master

Big Data & Data Science  
Telecom Paris



Big Data, Machine  
Learning, Deep Learning,  
NLP, Computer Vision

2018 - 2020

## Social



<https://maelfabien.github.io/>



[https://github.com/  
maelfabien](https://github.com/maelfabien)



[https://www.linkedin.com/  
in/mael-fabien/](https://www.linkedin.com/in/mael-fabien/)

Medium

[https://medium.com/  
@mael.fabien](https://medium.com/@mael.fabien)

## Language

- French (Mother Tongue)
- English (C2) (Bilingual Scientific Baccalaureate)
- German (B2)

## Interests

- ML and Data Science
- Sports (Fitness, running, hiking)
- Travel (South East Asia, New Zealand, South America)

## Experience



07.2019 - Now : Data Scientist Intern, Paris

Created innovative algorithms for automated text transformation understanding using tree-based structures and genetic algorithms. Worked with back-end and front-end engineers, R&D team and collected end-users feedbacks.

*Python, Algorithm optimisation, Natural Language Processing*



03.2019 - 07.2019 : NLP Data Scientist Intern, Remote

Designed and implemented a Semantic Similarity Search Engine in Python for a messaging service that groups all user's messages in a single feed.

*Python, Search Engine, Hierarchical Embedding, Sentence Embedding*



09.2018 - 06.2019 : Deep Learning Research Project, Paris

Research project for the French employment center on multimodal emotion recognition for Text, Sound and Video. Built a web app on top with Flask.

*Python, NLTK, Keras, Tensorflow, Flask, HTML, CSS*



02.2018 - 08.2018 : Non-Life Actuary Intern, Vaudoise, Lausanne

Redesign of a private Accident Insurance product. Statistical analysis, anomaly detection, product pricing (MLE, robustness tests, profit testing, simulations).

*R, VBA, Tableau*

## Entrepreneurship

MedikHub



2019: Hackathon Winner

Winner of the 2019 « Hubs as a service » Hackathon. Developed a mobile medical service for rural areas, a booking platform and a pre-diagnosis tool using NLP. Won a 5'000 euros prize.



Wanago

2016 - 2018 : Founder of Wanago

Crowdfunding platform for van acquisition. Jury Prize winner of the Entrepreneurial Contest Start Lausanne after a 6 months competition among 65 startups, with 5'000 euros prize.

## Teaching

Stat4decision

Data science company

09.2019 - Now : Data Science Instructor

Machine Learning Instructor for Stat4Decision. I provide 3 days intensive trainings « Python for Data Science » in the banking industry.

VIVADATA

02.2019 - Now : Machine Learning Instructor

Machine Learning Instructor for VIVADATA Bootcamp. Created course supports and gave over 35 lessons on Stats, ML, DL and Data engineering in Paris and Dakar.



09.2016 - 03.2018 : Student Teaching Assistant

Student Teaching Assistant for the 3rd year Bachelor course of Corporate Finance and Public Finance, 40% position on the side of my studies.

# Writing

## **Stat4decision** 09.2019 - Now : AI Writer

Data science company

Contributing to Stat4Decision's blog in French.

## **Towards Data Science**

### 08.2019 - Now : Editorial Associate

Editorial associate role at TDS. Working part-time on AI articles review and content creation.



## **EXPLORIUM** 03.2019 - Now : AI Writer

Contributing to Explorium's blog, a Data Science platform startup in Israel. Examples :

<https://www.explorium.ai/blog/whos-the-painter/>,

<https://www.explorium.ai/blog/interpretability-and-explainability-part-1/>



## 09.2018 - Now : Blog Writer

I started writing a data science blog in which I share articles (over 100 so far) and tutorials on Statistics, Machine Learning, Deep Learning, Reinforcement Learning, Data Engineering... and detailed projects from scratch. <https://maelfabien.github.io/>. I share some of my articles on Medium and earned earlier this year the « Top AI writer » badge.

# Technical Stack



### Daily user :

- Python (Pycharm, Jupyter, Sublime Text...)
- Machine Learning libraries: Sk-learn, NLTK, Spacy, OpenCV
- Deep Learning frameworks: Keras, Tensorflow, PyTorch

### Frequent user :

- R (RStudio), Scala (Zeppelin), Matlab
- SQL (Postgre, MySQL), HTML, CSS
- Cloud : AWS (EMR, EC2...), Google Cloud Platform
- Flask
- Unit tests

### Used in several projects :

- Hadoop, Spark,
- Cassandra, MongoDB, Neo4j
- Hive
- ElasticSearch, Kibana

# Recent Projects

## **fsText : Few Shot Text Classification**

Created fsText, a few-shot learning text classification library on GitHub, using pre-trained embeddings and siamese networks.

<https://github.com/fsText-suite/fsText>

## **Character-level LSTM for genre detection**

Created and trained character-level LSTM networks for genre detection from first names with over 90% accuracy.

## **Language Generation Model**

Created a language generation model using Tensorflow. I trained the model on my own articles.

<https://towardsdatascience.com/i-trained-a-network-to-speak-like-me-9552c16e2396>

## **What language is this? Language Recognition**

Built a Language Recognition web app using Markov Chains and a web scraper for Wikipedia.

[http://maelfabien.github.io/machinelearning/HMM\\_3](http://maelfabien.github.io/machinelearning/HMM_3)

## **GANs for face generation**

Built a GAN for face generation using Facial Emotion Recognition 2013 database (fer2013 on Kaggle)

## **Face Detection using HoG and CNNs**

Implemented face detection algorithms and compared several approaches relying on DL or not

<https://towardsdatascience.com/a-guide-to-face-detection-in-python-3eab0f6b9fc1>

## **Data Engineering Project**

Processing large amount of data (GDEL database) using AWS EC2, Cassandra and Spark.

<https://github.com/maelfabien/Cassandra-GDEL-Queries>

## **Fraud Detection**

Xente Fraud Detection competition (banking transactions) from Zindi platform.

## **Open Machine Learning Tutorials**

Published Machine learning tutorials on GitHub (NLP, Computer Vision, Reinforcement learning)

[https://github.com/maelfabien/Machine\\_Learning\\_Tutorials](https://github.com/maelfabien/Machine_Learning_Tutorials)

