



Vincent Duchauffour

Machine Learning Engineer

Paris, France

☎ 06.09.78.66.27

✉ vincent.duchauffour@proton.me

🌐 [Linkedin](#)

🌐 [Github](#)

Profile

Former social science student. My interest in computer science and mathematics led me to reorient my professional project towards data science.

After some experiences that allowed me to focus on business related tasks.

Since then I am focus on more software engineering tasks of ML librairies. I'm really looking towards contributing to open-source software.

Skills

- Python, Shell scripts
- Scikit-Learn, PyTorch
- Linux, Git, Docker
- AWS, GCP, DVC, W&B
- SQL, Elasticsearch, MongoDB

Languages

- French (Native)
- English

Experience

Since Oct. 2022 Machine Learning Engineer - Kinetix

- Principal maintainer of our proprietary ML librairies in collaboration across the R&D team, this include:
 - Rewrite large part of the ML codebase with strong OOP principles (SOLID principles / Design patterns)
 - Split the ML monolith into multiple independent repositories
 - Setup and responsible of our MLOps backend (AWS) and **production environment** (CD/CI, unit tests, packaging, deployment)
- Reduce by **50%** the inference time on the main ML pipeline (PyTorch, profilers)

Sep. 2021 – Sep. 2022 Data Scientist - Yanport

- Design, build and deploy an improvement on existing NLP system with ordinal classification with focus on explainability (NLTK, Spacy, PyTorch, HuggingFace)
- Design, build and deploy **new features** on the home value estimate engine (CatBoost)
- Maintain and deploy **data pipelines** and internal scripting

Sep. 2019 – Sep. 2021 Data analyst - Qare

- Design, build and deploy a monitoring tool for data and ETLs quality
- Analytics engineering (DBT)
- Contributing to ETLs and RETLs (MongoDB, AWS Redshift, Salesforce)
- EDA and ad hoc analyses, **interact with** concerned **corporate departments** and make **public presentation** of results
- Build **Tableau** dashboards

Education

2021 – 2022

Msc Data Science (IASD) - Paris Dauphine - *Apprenticeship*

Master thesis: Improving an existing NLP system with ordinality, a deep dive into explainability

2019 – 2021

Msc Digital Humanities - CNAM - *Apprenticeship*

Master thesis: Data analysis for efficient telemedicine customer service

- Programming, Statistics, ML concepts, Linear models, Clustering

Math evening classes - CNAM

- Analysis, Linear algebra, Graph theory, Linear optimisation, Probabilities, Statistics

2012-2017

Bachelor & Master Philosophy - Paris Nanterre

- Ethic, epistemology