



## Vincent Duchauffour

### Data Scientist

Paris, France

☎ 06.09.78.66.27

✉ [vincent.duchauffour@proton.me](mailto: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.

My first experience as a data analyst allowed me to focus on business oriented data analysis and analytics engineering.

My current experience allowed me to confirm this career path by contributing to data science projects in production-ready context.

## Skills

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

## Languages

- French (Native)
- English (B2)

## Experience

Since Oct. 2022 **Machine Learning Engineer - Kinetix**

- Rewrite large part of the codebase with strong OOP principles (**SOLID principles** / Anti-code smells policy)
- 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
- **EDA** and ad hoc analyses for external client

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