Vincent Duchauffour
Machine Learning Engineer

⌂ Paris, France

☎ 06.09.78.66.27

✉ vincent.duchauffour@proton.me

in Linkedin

○ Github

Profile

My interest in computer science and mathematics has led my career path towards data science. After some experiences that allowed me to focus on business-related, modeling and data engineering tasks, I am nowadays working closely on software engineering aspects of machine learning projects.

This includes maintaining ML librairies for production use, developing and monitoring ML pipelines and improving CI/CD guidelines. I'm really looking towards contributing to open-source software and learning low-level languages (Rust).

Skills

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

Languages

- French (Native)
- English

Experience

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

- **Maintain our proprietary ML librairies** in collaboration with the R&D team, this include:
  - Rewrite large part of the ML codebase with **strong OOP principes** (SOLID principes / Design patterns / Refactoring)
  - Set up the **production environment** in accordance with DevOps team (AWS, code quality, shell scripting, CI/CD, Docker, unit tests, packaging, deployment)
  - Write from scratch a multiple object tracking library using pre-trained models
  - Split the ML monolith into multiple independent repositories
- **Reduce by 50% the inference time** on the main ML pipeline (**PyTorch**, profilers, model serving optimization)

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**), validate these additions by ensuring that they don't have side effects on other features
- Maintain and deploy **data pipelines** and internal scripting

Sep. 2019 – Sep. 2021 **Data Analyst/Engineer** - **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 at Yanport*
Master thesis: Improving an existing NLP system with ordinality, a deep dive into explainability

2019 – 2021

Msc Digital Humanities - CNAM - *Apprenticeship at Qare*
Master thesis: Data analysis for efficient telemedecine customer service
Math evening classes - CNAM
Analysis, Linear algebra, Graph theory, Optimization, Probabilities, Statistics

2012-2017

Licence & Master Philosophy - Paris Nanterre