📱 +33 6 59 29 14 50

✉ arthur.bernard.92@gmail.com

in /in/arthur-bernard-789955152

⌗ ArthurBernard

Arthur Bernard

Experienced Data Scientist and Python developer, specialized in the application and customization of Large Language Model (LLM).

Key skills

- Retraining LLM on specifics datasets.
- Expertise in Retrieval-Augment Generation (RAG) techniques: contextual enhancement.
- Proficiency in API integration for complex LLM applications.
- Development of user-friendly graphical interfaces for LLM applications.
- Advanced data collection and cleaning methodologies for LLM training.

Programming

- Highly advanced: **Python** (expertise in NumPy, Pandas, Cython, PyTorch, Sickit-Learn, Flask, StreamLit, Transformers, HuggingFace, etc).
- Advanced: **Shell**, VBA, LaTeX.
- Moderate: **C++**, HTML&CSS and JS.

MOOCs

- 2023 - **Sequence Models** (Natural Language Processing, Recurent Neural Network, Attention Models and Transformers), on Coursera
- 2018 - **Machine Learning**, Stanford University course, on Coursera.
- And other diverse courses about **Python**, **Linux**, **C++**, **HTML&CSS**, **JavaScript**, **Git**, etc.

Interests

- Artificial intelligence: Natural Language Processing (NLP) and time-series.
- Crypto-currencies/Blockchains.
- Various Open Source projects.

Experiences

Dec. 2022 – Present — **Freelance Data Scientist** *at LLM Solutions*
Operating as a freelance data scientist under the name LLM Solutions, offering AI customization services to businesses. Developed innovative solutions enabling companies to **internally integrate customized LLMs**, including **model retraining** and the use of **Retrieval-Augmented Generation (RAG)** techniques on their private data. **Technologies Used:** Python, Flask, Streamlit, GPU Computing, RAG Techniques, LLM Retraining, etc.

Mar. 2020 – Sep. 2021 — **Data Scientist Consultant** *at Coperneec*, Paris
For **Natixis** in Risk Management Department in the Market Stress Test team: automation of recurrent tasks, R&D on reverse stress test methodology, **fast pricing of products**, python package to manage data across different internal data sources (**SQL**, **MDX**, **Hadoop**), etc.

Jun. 2018 – Oct. 2019 — **Quant Researcher** *at Napoleon Group*, Paris
R&D of trading strategies, **portfolio allocation** algorithms, **multivariate prediction** with neural networks, **execution order algorithms**, development of **backtesting** and **financial analysis** tools, and web-scraping data.

Personal projects

2023 – 2024 — **Customized LLM: *HackerNewsGPT* project**
Objective: Developed solution to enhance technology monitoring by enabling users to ask real-time questions about the latest articles from *HackerNews*.
Role & Contributions: Retraining a LLM using RAG to incorporate the latest articles for real-time, context-aware question answering. Managed the backend on a GPU server with a Flask API, and developed a user-friendly web app interface using Streamlit.

2018 – 2019 — **Machine/deep learning tools adapted to finance**
Development of a Python and Cython package to create **neural networks**, **backtest strategies**, analysis with **econometric models** and **financial indicators**, etc. Published on PyPI as *fynance*.

2016 – 2020 — **Trading bot algorithms on crypto-currencies**
Development and maintenance of trading bots with Python and Bash scripts. Starting in 2016 with **arbitrage strategy**, and more recently create **strategies with neural network**. Partly available on my GitHub in the repository *Trading_Bot*.

Education

2017 – 2018 — Master's Degree **Econometrics of Banking and Financial markets** *at Aix-Marseille School of Economics*, Marseille
Courses: Stochastic finance, financial econometrics, financial engineering, econometrics of exchange rates, neural network, etc.

2013 – 2016 — Bachelor's degree **Economics and Management, specialized in Finance** *at Aix-Marseille University*, Marseille
Courses: Time series econometrics, financial markets, statistics, optimization, informatics (SQL and VBA), etc.

Miscellaneous

2019 — **Data-science competition** *at ENS Challenge Data*
6th place out of $100+$ competitors in a daily US stock market prediction challenge, organized by **Capital Fund Management**.

Present — **Hobbies**
Cooking, sports (swimming, cycling and running), and DIY projects.