

# HUGO MICHEL

## QUANT/DATA SCIENTIST



☎ 0620575205 🏠 26 years old 🇫🇷 French ✉ [hu.michel@orange.fr](mailto:hu.michel@orange.fr) 🌐 [Portfolio](#) 🐙 [GitHub](#)

### Education

---

**Sorbonne School of Economics — Rank: 1st in class (1/26) — Grades: 15.65/20** Sep. 2023 – September 2024  
*Master of Economics in Quantitative Finance* Paris, Ile de France

- **Electives (taught in English):** *Asset Pricing, Stochastic Processes, Financial Econometrics, Simulation and Numerical methods, Derivatives, Quantitative Risk Management, Machine Learning & Credit Scoring, Financial market sentiment analysis, Network Analysis & Data Mining for on-chain analysis, Financial Market Microstructure, Data Privacy & Data Anonymization techniques*
- **Master Thesis:** *Physics Informed Neural Network (PINN) for European Option Pricing*, (Supervisor: Dr. Eric Vansteenberghe Economist Researcher at ACPR Banque De France)
- Won **1st place** in Credit Risk Challenge - [The DRIM Game](#) (organized by Deloitte, SAS, SFIL, Crédit Agricole) [\[Slides\]](#)
- Won **1st place** in [Sorbonne Data Challenge](#) (organized by ENGIE, DataScientest) [\[Slides\]](#)

**Télécom Paris (Institut Polytechnique de Paris (IP Paris) — Grades: 16.90/20** Sep. 2021 – February 2023  
*Post Master degree in Big Data & Artificial Intelligence* Paris, Ile de France

- **Electives :** *Massive data processing, Distributed algorithmic, Statistical Learning, Distributed Big Data architecture, BI & Data Warehousing, Data visualization, Semantic Web, Machine Learning, Convex Analysis & Optimization theory, Deep Learning, Reinforcement Learning, Computer Vision & NLP*

**Aalto University — GPA: 4,62/5** Jan. 2020 – July 2020  
*Computer Science - Exchange semester in Finland* Helsinki, Finland

- **Electives:** *Information Security, Cybersecurity, Networking principles*

**University of Technology of Troyes - UTT — GPA: 3,99/5** Sep. 2016 – July 2021  
*Master of Engineering in Computer Science with major in Data Science* Troyes, Aube

- **Electives:** *Algebra & Analysis, Algorithmic, Bayesian Statistics & Probability, Graph theory, Information System security, Data Exploratory Analysis, Usage Centered Design, BI Architecture, Web Technologies, Information System Analysis and Modelling, Symbolic AI, Multi Agent Systems & Game Theory*

### Experiences

---

**Crédit Agricole - CIB — Front office (Global Markets Division)** September 2023 – September 2024  
*Quant/Data Scientist Intern* Paris, Ile de France

- Development of a tool for the rapid valuation of structured and vanilla products through the use of models that lie at the intersection between physics and deep learning. Work carried out in close collaboration with the Quant Researchers team

**Capgemini Insights & Data** July 2022 – September 2023  
*Junior Data Engineer* Paris, Ile de France

- Mission for an international financial institution as a consultant
- Implementing a BI architecture to calculate diverse business KPIs, enhancing data observability and provide insights:
  - Automating RWA (Risk-Weighted Assets) calculation process
  - MLOps: Deploying at scale of the credit scoring model in collaboration with data scientists
  - Building ETL pipelines combining third party data from APIs with the datalake in order to automate the reporting of several financial documents (financial audit trails, syndication,...) and KYC digitalization
- As part of a data migration project, creation of an internal framework using sqoop technology for the mass recovery of historical data from a MAPR cluster based on Hive technology to Oracle Cloud.
- DevOps: Led release management and deployment in production, maintaining the CI-CD pipeline

**Télécom Paris x Baalbek Management** November 2021 – July 2022  
*Data Scientist Intern* Paris, Ile de France

- Academic project in connection with a company from energy sector: Research work on NILM (Non Intrusive Load Monitoring) technologies. Building unsupervised end-to-end model based on an convolutional autoencoder 1D architecture for activity detection from the household's aggregated power signal [\[🔗 Project\]](#)

## Aubay

Machine Learning Engineer Intern

January 2020 – August 2020

Paris, Ile de France

- Designing end-to-end application oriented summarization and automatic generation of French texts.
- Applying a transfert learning approach in order to fine-tune state-of-the-art models based on a transformer neural network architecture for the extractive synthesis task and the abstractive synthesis task
- Integration of fine-tuned models in a web application. Back-end (Flask) and front-end (SvelteJS) ([Project](#))

## AkaBI

Software Engineer Intern

July 2019 – December 2019

Luxembourg

- Development of an expense report management module for consultants within the company's ERP intranet

## Languages

---

French (Mother tongue), English (TOEIC C1, BULATS C1), German (Level B1)

## Projects

---

### Derivatives Pricing | *Python, SciPy* | ([Project](#))

- Estimate the fair price of derivative based on an underlying asset:
  - Pricing in continuous time using Monte Carlo simulations, Euler Mayurama method, Black-Scholes, Analytical solution of the Geometric Brownian Motion and Finite Difference Methods. Pricing in discrete time using Binomial Tree
  - Derive implied volatility with Black-Scholes and Newton-Raphson algorithm
  - Simulating stochastic volatility with Heston model
  - Modeling the movement and evolution of interest rates with a single-factor short-rate model (Vasicek)

### BI Architecture | *Bash, PySpark, Cassandra, Zeppelin, Spark* ([Project](#))

- Designing resilient Big Data architecture for distributed storage and distributed query execution based on the GDELTA database

### Lienar and Non linear Econometrics | *Python, Statsmodel, SciPy* | ([Project](#))

- Stationary test, forecasting with SARIMA model, Cointegration test, Variational Autoregressive model, Impulse response function to analyze shocks, Granger causality tests, Markov Switching Model and Smooth Transition Regression to capture non-linearity, Difference and Difference method, Logit-Probit model

### Data Challenge IDEMIA: Face Recognition | *Python, Tensorflow, Scikit-Learn* | ([Project](#))

- Designing model to predict whether or not two photos correspond to the same identity. Special attention is given to the fairness of the prediction model according to the gender of the person

### Quantitative Risk Management | *Python, SciPy* | ([Project](#))

- Risk measures: Computing parametric VaR, non parametric VaR, stressed VaR and Expected shortfall (CVaR)
- Valuation and Hedging Bonds & Yields (Computing YTM, DV01, duration and convexity)
- Hedging: Computing Greeks (Delta hedging, Gamma hedging, Vega Hedging)

### Sentiment Analysis on ECB's statement | *Python, NLTK, SpaCy, Huggingface* | ([Project](#))

- Web Scrapping on ECB website to collect ECB press releases
- Sentiment analysis on ECB statements and their impact on the abnormal cumulative returns of European equities market using BERT transformers models

### Financial Tweets Sentiment Analysis | *Python, NLTK, Huggingface, MongoDB Cloud* | ([Project 1](#), [Project 2](#))

- Apply dictionaries based model (VADER), Conventional classifier algorithms (Naïve Bayes, SVM, Random Forest,...) and more advanced model based on transformer architecture (FinTweetBERT) to predict sentiment's link with stock returns

### Explainable AI for credit risk management | *Python, Scikit-Learn* | ([Project](#))

- Apply statistical learning algorithms (Logit, Ensemble learning and Boosting algorithms) to predict the probability of default
- Applying portfolio optimization strategy by finding the best acceptance rate (threshold) which minimize the most the total expected loss
- Building a XAI dashboard (shapley values, partial dependence plots, features contributions,...) and applying DICE technique

## Technical skills

---

**Programming languages:** Python, R, Bash, SQL, LaTeX, pySpark

**Data Viz:** Power BI, Tableau, Kibana, Seaborn, Matplotlib, Plotly, D3.js, Gephi

**Data Engineering:** Hadoop Ecosystem, Elasticsearch (ELK stack), pySpark, ODI, Talend ETL, MongoDB Cloud, PostgreSQL

**CI-CD DevOps:** Jenkins, Git, Jira

**Data Science/Framework:** NumPy, Pandas, bs4, Scikit-Learn, Statsmodel, SciPy, PyTorch, NLTK, SpaCy, HuggingFace, Flask