Alexis Vannson

French and Austrian nationality | vannson.alexis@gmail.com | GitHub: https://github.com/alexisvannson

EDUCATION

CentraleSupélec - Essec business school **Paris, France**

BSc in Artificial Intelligence, Data and Management Sciences *September 2023- Present*

Related Coursework: Advanced Linear Algebra, Integrals & Differential Calculus, Foundations of Statistics & Data Modeling, Algorithms & Software concepts, Optimization, Programming for data science, machine learning, NLP, Computer Vision

EXPERIENCE

Software Engineering Intern- *Polyfact* *May - August 2025*

- Reverse-engineered French government APIs and built an **automated web scraping pipeline** using **Puppeteer** to capture and integrate real-time political data—agendas, press releases, and articles into existing infrastructure across 10+ ministries.
- Built a search agent with **TypeScript**, **LangGraph, OpenAI**, **Perplexity**, **Tavily**, **SerpApi**, and **Pappers** APIs to generate detailed reports on lesser-known French officials, expanding the profile base by 40%, enabling real-time search and storage via **Supabase/PostgreSQL**

Research Intern- *CentraleSupélec* *March - June 2024*

- Conducted a comparative performance analysis of 4+ machine learning architectures (**MLPs, CNNs, U-Nets and GNNs**) for **computer vision** tasks, benchmarking accuracy, training time, and model efficiency. Applied statistical methods to evaluate model performance and validate experimental results
- Optimized graph-based data preprocessing by using **KNN clustering** to connect superpixels, achieving a 35% reduction in processing time and boosting model accuracy by up to 7%.

Data Science Intern - *Emeria Technologies* *April 2024*

- Performed **feature engineering** to isolate high-confidence data subsets, achieving 95% accuracy on a subset affecting 6% of the 5 million unprocessed invoices per year.
- Developed a **Random Forest** model using **Scikit-learn**, analyzing 138,000+ records to identify key features, boosting contract number prediction accuracy to 85% through hyperparameter optimization and cross-validation, reducing misclassification by 4%.

AI Practitioner- *Automatants* *September 2023 - June 2024*

- Explored and implemented various AI models, including **CNNs, ResNets and Autoencoders,** to deepen my understanding of artificial intelligence through hands-on application. Worked on pneumonia detection, deepfake identification, image denoising, audio classification processing as part of lab sessions and hackathons using **PyTorch**.

PROJECTS

Human Face Generation with DCGAN – PyTorch, PyTorch Lightning, torchvision

- Architected and trained a **Deep Convolutional Generative Adversarial Network** from scratch, implementing custom generator and discriminator networks with batch normalization, leaky ReLU activations, aiming at stable training convergence.

Drug-Perturbation Analysis Pipeline | Python, AWS, ChemBERTa

- Built end-to-end bioinformatics pipeline integrating gene expression signatures from LINCS L1000 Level 5 data (20GB) with molecular structure embeddings using an AWS S3 bucket
- Implemented Random Forest regression with k-fold cross-validation on high-dimensional genomic data, establishing predictive models for drug-gene interactions (achieving moderate predictive power R^2 of 0.19 on 978 landmark genes)

SKILLS

Programming Languages: Python, SQL, TypeScript, C, JavaScript
Mathematics & Statistics: Linear Algebra, Multivariate Calculus, Statistical Modeling, Optimization, Probability Theory
Data Science & ML: PyTorch, TensorFlow, Scikit-learn, NumPy, Pandas, PySpark, OpenCV
Cloud & Tools: AWS (S3, EC2), Git, Docker, UNIX/Linux, Shell scripting
Languages: French (Native), German (Native), English (Fluent), Spanish (Basic), Dutch (Basic)

INTERESTS AND ENGAGEMENTS

Hockey - French National team *2018 – 2022*

Represented France as a member of the national Field Hockey Team for international games