Joel Smith

SUMMARY

Data Science & Artificial Intelligence MSc graduate with a strong focus on building reliable, well-structured machine learning systems. Experienced in translating mathematical models into clean, reproducibile code, and in designing end-to-end pipelines spanning data, training, evaluation, and deployment. Demonstrated depth across generative modelling and applied ML systems, with hands-on experience using PyTorch and modern ML tooling.

Education

University of Liverpool Sept 2024 - Sept 2025
MSc in Data Science & Artificial Intelligence, Distinction

University of Chester Sept 2020 - July 2023
BSc in Economics, First Class

- Awarded the University Prize for Economics

Projects

LLM-Based Financial Advisor Web App GitHub Repo

- Designed and implemented a full-stack LLM-based financial Q&A system to reduce information asymmetry in retail finance, integrating retrieval-augmented generation with real-time market APIs and vector search (ChromaDB/pgvector) to ground responses in up-to-date, verifiable data.
- Engineered a production-oriented inference pipeline with a FastAPI backend and React frontend, containerised via Docker and deployed on a local Kubernetes (kind) cluster to support modular services and reproducible serving.

Manipulation Lab GitHub Repo

- Designed and implemented a modular benchmarking framework for robotic imitation learning in NVIDIA Isaac Sim, supporting end-to-end pipelines for data collection, training, and evaluation.
- Reduced experimentation overhead by shifting task and training configuration from code to Hydra-based configs, enabling rapid iteration and reproducible comparisons across benchmarks.
- Improved maintainability by decoupling environment, dataset, and model components, reducing boilerplate and simplifying the addition of new tasks.

Generative Modelling from Scratch GitHub Repo

- Authored a blog-style series of deep technical studies on generative models (e.g., mixture models, VAEs, diffusion models) developing detailed mathematical derivations and explaining core probabilistic concepts from first principles.
- Produced supporting PyTorch/NumPy implementations and visualisations to illustrate the theory and compare learned generative distributions against real data.

Reinforcement Learning from Scratch GitHub Repo

- Authored a notebook-based series on reinforcement learning fundamentals, deriving and implementing core algorithms (bandits, dynamic programming, Monte Carlo methods) from first principles.

Machine Learning and Deep Learning Projects University Coursework

- Implemented a range of applied machine learning and deep learning projects across multiple problem domains using classical and neural approaches.

Experience

Commercial Finance Analyst June 2024 – Sept 2024
Matalan Liverpool, UK

- Queried, cleaned, and validated large-scale SQL datasets (10M+ rows) to support sales and promotional analysis, delivering time-sensitive insights to commercial finance stakeholders.

Team Supervisor Sept 2020 – January 2024

Costa Coffee   Liverpool, UK

- Led a team in a high-throughput retail environment while completing a full-time undergraduate degree, developing management and communication skills.

Skills

Machine Learning: PyTorch, scikit-learn, Torchvision, Hydra
Data & Experimentation: PostgreSQL, Google BigQuery, Microsoft SQL Server, Weights & Biases
LLM & Retrieval: OpenAI API, ChromaDB, pgvector
Systems & Deployment: FastAPI, Docker, Kubernetes (kind), Git, Linux, AWS (EC2)
Languages: Python, SQL