Jordan Browne-Moore

Machine learning engineer, London, UK

jordan@jordanbrowne-moore.com github.com/ElSnacko linkedin.com/in/jordan-j-browne-moore

Summary

I build machine learning systems for regulated finance, and I study how language models work on the inside. Credit scoring, KYC automation, fraud detection and fine-tuned LLMs, taken from problem definition through production deployment at a neobank and a market-intelligence platform. Alongside that, independent research into LLM internals: where refusal lives in a model, and what agents do when their tools lie to them.

Experience

Senior Data Scientist, then Consultant, Kuda Technologies

Hired to build the credit risk and fraud detection functions from the ground up. Retained as external consultant to continue leading these workstreams after transitioning out of the full-time role. Owned model development, validation methodology, and production deployment across credit scoring, KYC automation, and fraud detection. Mentored 4 data scientists and participated in cross-seniority hiring.

- **Credit scoring system.** 50% reduction in first-payment defaults. 75% reduction in total portfolio defaults. Model serves as the production scoring engine for the entire lending portfolio.
- **KYC automation pipeline.** Verification time reduced from 7 days to 4 hours. System now handles the majority of KYC decisions autonomously, with human review reserved for edge cases only.
- **Graph-based fraud ring detection.** Identified and disrupted 20+ fraud rings (3+ connected accounts each). Output fed directly into risk decisioning, triggering account restrictions and recovery actions across affected networks.
- **Salary and recurring inflow detection.** 35% coverage uplift over the third-party supplier. 45% uplift on accounts where the supplier returned no data. 9% false positive rate. Directly improved credit eligibility accuracy for underbanked customers.

Data Scientist III, CB Insights

LLM fine tuning, NLP, production ML engineering.

- **LLM fine-tuning for revenue estimation.** 80% reduction in API costs vs third-party LLMs. Production inference API serving revenue estimates across the platform.

Senior Data Science Consultant, Capgemini

Financial services consulting.

Data Science Consultant, Beyond Analysis

Independent research

- **The refusal axis is layer-local.** The refusal axis rotates ~90° between consecutive transformer blocks, so single-vector steering hits a geometric ceiling. Extends the refusal-steering method of García-Ferrero, Montero & Orus (arXiv:2512.16602). My own results, cross-checked against the committed data: 12 of 12 claims reproduce.
- **Self bootstrap exfiltration in open weights agents.** Zero autonomous self bootstrap in 1,332 mundane control trials; tool response poisoning produces about 36% pooled compliance.
- **LLM activation steering toolkit** (github.com/ElSnacko/llm-steering).
- **Feature bootstrapping toolkit** (github.com/ElSnacko/feature-bootstrapping-toolkit).

Writing

- Self Bootstrap Exfiltration in Open Weights Agents: Absent Without Instruction, Induced by Injection (July 2026)
- I Built a System to Map My Own Beliefs. It Showed Me the Easy Half of a Harder Problem. (June 2026)
- The Refusal Axis Is Layer Local: Why Single Vector Steering Hits a Geometric Ceiling (June 2026)
- We Cut KYC From 7 Days to 30 Min Using VLMs. The Hard Part Wasn't the Model. (June 2026)
- Your LLM Extraction Pipeline Is Out of MRM Scope. Your Credit Model Isn’t. Now What? (May 2026)
- We Finetuned Qwen 3 32B to Replace an LLM API. The Evaluation Was Harder Than the Training. (May 2026)
- Your Best Feature Might Be Your Biggest Liability (April 2026)

Education

- **MSc Financial Economics**, Birkbeck, University of London
- **BSc Economics & Finance**, Southern Oregon University

Certifications

- GCP Associate Cloud Engineer
- Azure AZ-900

Tools

LightGBM, Python, SciPy, SHAP, SQL, AWS, Qwen 3 32B, LoRA, FastAPI, HuggingFace Transformers, Gemini Flash, Qwen3 8B VL, NetworkX.