Sassan Bhanji

sb2700@cam.ac.uk | +447484210939 | linkedin.com/in/sassan-b-92b772215

Education

University of Cambridge, *MEng in Computer Engineering and Bioengineering* Oct. 2023 – Jun. 2027

- Grade: 1st class – awarded the Townsend scholarship and Wright Prize.
- President of the St John's College Engineering Society (Parsons Society).
- Music scholarship and Principal Violist of the Cambridge University Orchestra.
- Cambridge AI Safety Hub (CAISH) committee.
- **Relevant Modules:** Deep Learning, Reinforcement Learning, Computer Vision, Mathematical Methods, Control, Signal Processing, Information Theory, Inference, Bioengineering, Neuroscience.

King's College School Wimbledon Sep. 2012 – Jul. 2023

- 4A* in A level Maths, Further Maths, Physics and Chemistry. 13 Grade 9s or equivalent at GCSE.

Experience

Geodesic Research – AI Safety Research Intern *Python, Hugging Face Transformers* Oct. 2025 – Present

- Researching obfuscation generalisation under chain-of-thought optimisation pressure; working towards ICML submission.
- Developed experimental pipelines for reinforcement learning on Qwen-4B, including dataset construction and augmentation.

National University of Singapore – Machine Learning Research Intern *Python, JAX, PyTorch* Jul. 2025 – Sep. 2025

- Investigating ML-augmented subgrid-scale modelling to accelerate CFD simulations - supervised by Dr Qianxiao Li.
- Implemented and benchmarked Transformer and UNet architectures for turbulence prediction.
- Analysed the effect of different filtering methods on a priori and a posteriori model error under different boundary conditions.

BWT Alpine Formula 1 team – Engineering Intern *Python, JavaScript, SQL, Docker, Kubernetes* Jul. 2024 – Oct. 2024

- Developed internal tools to optimise the stress analysis pipeline, improving processing efficiency by $\sim 20\%$.
- Conducted FEA using Hypermesh on components including the rocker, drybrake, and nose cone; supported FIA certification and crash test analysis.
- Analysed live telemetry and cable data using ATLAS during practice and qualifying sessions in the Operations Room.

Projects

Alignment Research Engineer Accelerator (ARBOx3)

- Selective training program in AI safety research; built GPT-2 from scratch, applied mechanisic interpretability to analyse reasoning circuits, fine tuned models using PPO for RLHF.

Constitutional AI with Direct Preference Optimisation (DPO)

- Reimplemented Anthropic's Constitutional AI pipeline on open weights (Mistral-7B) using LoRA and DPO; completed full stack: critique/revision, SL fine-tuning, preference gen, and DPO training.
- Replaced RLAIF with DPO for efficiency, creating a modular pipeline (base –> SL –> DPO).

Time Series Forecasting and Data Science

- Modelled real-world time series in Python using regression pipelines with cross-validation and residual diagnostics.
- Forecasted atmospheric CO_2 trends with Fourier terms and modelled early COVID-19 growth via log-linear regression.
- Analysed London Underground ridership to detect seasonality using periodograms, hybrid models, and overfitting prevention techniques.

Mars Lander Simulation and Autopilot

- Developed a Mars lander simulation in C++ incorporating gravity, thrust, drag, and aerodynamic effects.
- Designed and implemented an autopilot system using control theory to ensure stable landings.
- Used OpenGL for real-time 2D visualisation; implemented procedural terrain generation with Perlin noise.

Languages

- English – native, Persian – native, French – C1 qualification, Spanish – B2 qualification, Italian – B2 level