Tong Xiao

Phone +44 7756536490 | **E-mail** jackxiao031104@gmail.com | LinkedIn | Github

EDUCATION

Imperial College London London

MSc Statistics & Data Science *10. 2024 – 10. 2025*

- Applied Statistics with focus on deep learning and Bayesian statistical theory

BSc Mathematics with Statistics *10. 2021 - 06. 2024*

- Graduated with First Class Honours
- Relevant Modules: Bayesian Probability, Stats Modeling, Time Series, Monte Carlo Simulation, Optimisation, ML

BASIS International School Park Lane Harbor Shenzhen

High School Diploma *09. 2018 - 06. 2021*

- GPA: 4.0/4.0 Awarded Full Merit-based Scholarship (USD 150k) from BASIS Global (Acceptance Rate: <3%)

AWARDS

- Awarded Distinction for ranking Top5% globally in the 2021 American Maths Competition 12B, qualifying for AIME
- Awarded High Distinction for ranking Global Top3% in the 2019 Australian Maths Competition (Intermediate Division)
- Team & Individual Second Prize Winner at the Duke Math Meet China finals 2019
- Perfect Score in Waterloo Maths Contest 2019 (Fryer level)

WORK EXPERIENCE

Data Science Intern Manchester

Natwest Group *06 - 08. 2024*

- Improved ML classifier for Suspicious Activity Reports (SAR) by developing an LLM-based model trained on text datasets from S3 database in AWS Sagemaker. Integrates Retrieval-Augmented Generation (RAG) features
- Applied clustering analysis and machine learning to enhance data quality, **boosting 30% accuracy** to existing model
- Rotated as Scrum Master within our agile team to ensure efficient delivery of our solution to Natwest's FinCrime team

AI Research Intern London

Imperial EEE Department, Deepwok Lab *04 - 09. 2024*

- Designed a novel LLM architecture for DNA tasks by incorporating GPT4's tokenizer into a Transformer head
- Developed a modular evaluation framework that integrates our DNA LLM with downstream tasks (e.g. gene prediction, Protein-DNA interaction prediction) and benchmarks, enabling automated performance evaluations and comparisons
- Engineered GPU acceleration for our model, resulting in a 10x speedup in performance on downstream tasks
- Evaluated our model using my framework, showing its ability to **outperform** DNABERT **(state-of-the-art)** on 6 of 8 Genomic Benchmarks (a collection of 8 datasets covering downstream tasks on DNA of humans and different species)

Optiver Trading Academy London

Optiver *11 - 12. 2023*

- Developed a delta-hedging trading algorithm that autonomously traded options in Optiver's simulated exchange, Optibook, leveraging the Black-Scholes model for option pricing
- Trained a SVM model to dynamically adjust the credit for bid/ask prices based on liquidity, volatility, time to expiry, and historical price movements. Utilized GARCH to predict market volatility, achieving an R-squared of 0.7
- Achieved positive PnL of $500,000, **ranked in top 20%** of participants in the contest with over 50 student traders

Undergraduate Research (UROP) London

Department of Mathematics, Imperial College London *07 – 09. 2023*

- Developed 3 math proofs around the 'Continent of Stability' concept within relativistic quantum mechanics
- Developed a novel algorithm in Julia for numerically solving Schrodinger's equation in 2D quantum harmonic oscillator
- Designed a Python algorithm for computing and visualizing Dirac's equation for a free particle using Finite Difference
- Further explored quantum algorithms like Variational Quantum Eigensolver (VQE) for complex optimization

TECHNICAL SKILLS

- **Languages**: Python (proficient), MySQL/SQLite (proficient), Julia, C, C++ , R, MATLAB
- **Technologies**: Git, Bash/Shell, AWS Sagemaker, Azure, Snowflake, S3 data storage
- **ML Packages**: Pandas, PyTorch, Tensorflow/Keras, Llama3, AzureOpenAI API, Langchain, Pytorch Lightning