Dr. Scott Vinay

Data Scientist | Quantum Physicist

Derby, UK

+44 7877 078 569 | scott.vinay.93@gmail.com | scottvinay.com | ScottVinay | ScottVinay

A data scientist with significant experience designing and developing novel machine-learning solutions, particularly regarding time-series and language data. This is supported by a background in publishing well-received research in the statistical dynamics of quantum systems, and excellent skills at writing high-performance code and designing clear and extensible software architectures. With strong mathematical skills and demonstrated aptitude guiding teams in both the UK and Australia, I am a motivated engineer with ambitions to lead data science teams to solve some of the biggest challenges in society.

Employment

Transformative AI

United Kingdom

Senior Machine Learning Researcher — *Oct. 2021 – Present*

- Working on the development of deep classification models for the prediction of arrhythmias from electrocardiogram data. Developed intricate multi-stage training pipelines and novel deep-learning algorithms for characterising the shapes of time-series data.
- Lead author on report to FDA detailing our product, as well as lead author on publication under preparation for Nature Medicine.
- Lead the move from Keras and traditional RNNs to a fully Pytorch-native Transformer architecture, which simplified code while allowing for explainability of model decisions and introducing pretraining into the pipeline.

Machine Learning Researcher — *July 2020 – Oct. 2021*

- Took responsibility for the data pre-processing system; drawing from a large number of messy, disparate datasets and consolidating them to a single data standard, and refactored the codebase from the ground up into a unified pipeline.
- Introduced novel Bayesian-based algorithms to adapt models to an individual patient. This improved test-set F_1 score from 40.1% to 77.7%, and area-under-ROC from 64.8% to 92.7%.
- Produced high-quality reports and papers, which contributed significantly to the winning of research grants worth over a million USD.
- Supervised a teams of external contributors and interns, including producing resources and guides for moving to a cloud-based environment.
- Adapted cutting-edge results from research in bias-reduction in AI to improve generalisability to new databases.

isgood.ai

Melbourne, Australia

Data Scientist — *Aug. 2019 – April 2020*

- Co-designed, developed, and tested a novel machine learning pipeline for rapid understanding of a client's text, and semantically matched it to a database of existing text.
- Developed a highly modular code structure to allow for ease of delegation of tasks across a team and deployed iteratively via Gitlab.
- Took on a leadership role during the absence of the team leader, which involved tutoring new members on complex methods and assigning and reviewing progress on subtasks.

Black.ai

Melbourne, Australia

Data Technician — *Jan. 2020 – April 2020*

- Helped create and verify the dataset for a massive-scale image recognition system.
- Worked closely with development and machine-learning teams to help optimise efficiency of classification models.

University of Warwick

Coventry, United Kingdom

Undergraduate Research Internship - Fluid Dynamics — *Summer 2014*

- Modelled swimmers in the Stokes regime as sets of singularities, and used a method of images similar to that used in electromagnetism to calculate the resultant flow and behaviour when brought near to a wall. Used to investigate the behaviour of so-called "squirmer"-type creatures.

Undergraduate Research Internship - Computer Vision — *Summer 2013*

- Designed and constructed a computer vision system that used OpenCV to capture an image of a set of counters at a skewed angle, load the positions into an atomic energy simulator, and find the local-minimum energy. This allowed for researchers to easily test hypothesised atomic configurations in a tactile way, which I demonstrated at a university open day.

Education

PhD in Quantum Cryptography
Sheffield, United Kingdom

The University of Sheffield
Sept. 2015 - March 2019

- Thesis on *The Statistics and Security of Quantum Key Distribution.*
- Research included development of both practical proposals for technological protocols and novel mathematical frameworks for efficient statistical analysis of arbitrary networks.
- Supervised by Pieter Kok and Stephano Pirandola.

MPhys in Physics
Coventry, United Kingdom

The University of Warwick
Oct. 2011 - Jun. 2015

- 1^{st} class honours with an 84% average grade.
- Dissertation on *Energetics of Knotted Defects in Nematic Liquid Crystals.*
- Supervised by Gareth Alexander.

Machine Learning MOOC
Via Coursera

Stanford University
2018

Publications

Total citations: 62

Statistical analysis of quantum entangled network generation
Physical Review A

Scott E. Vinay and Pieter Kok
April 2019

- We developed novel numerical and analytic techniques for the analysis of probabilistic network generation, which we apply to the specific case of quantum communication to prove new security bounds.
- Improved secure communication rates by three orders of magnitude over more simplistic methods.
- Techniques used: Markov chains, Cauchy's residue theorem, Fourier transforms, complex analysis.

Extended analysis of the Trojan-horse attack in quantum key distribution
Physical Review A

Scott E. Vinay and Pieter Kok
April 2018

- We analysed the effect of a side-channel attack on quantum communication, proving an increased bound on security while relaxing the assumptions previously used.
- Techniques used: Covariance matrices, Krauss-decomposition of channels.

Practical repeaters for ultralong-distance quantum communication
Physical Review A

Scott E. Vinay and Pieter Kok
May 2017

- We presented a complete protocol for a quantum repeater, and an analysis of its efficiency that takes probabilistic effects into careful consideration. We show that it produces high secret key rates relative to comparable protocols.
- Techniques used: Order statistics, entanglement distillation.

Skills and Personal

Software
- Highly proficient in the use of Python for research applications, including object-oriented and parallel performance-focused approaches. Highly proficient with Pandas, Numpy, Matplotlib, Git, and machine learning packages including Sklearn, Keras, and Pytorch.
- Significant experience using Mathematica, Unix, LaTeX, R, and SQL.
- Comfortable and confident in the Google Cloud Platform (GCP) architecture, including the Cloud Compute, Storage, and AI Platform systems, as well as Amazon Web Services (AWS), using EC2 machines and S3 storage.
- Experience with C and MATLAB.

Presentational
- Confident at public speaking on complex topics.
- Delivered talks at multiple department symposia, as well as at meetings of the multi-university White Rose group.
- Presented research at conferences at the University of Cambridge (UK), Technische Universität Wien (Vienna), and the University of Nottingham (UK).

Personal
- In my personal time I like to keep active through rock climbing and snowboarding.
- I am a keen chess and go player.