

Behzad Karkaria

+447884092948, London, UK

behzad.karkaria@gmail.com <https://behzadk.github.io/> <https://github.com/behzadk>

EXPERIENCE

Data Scientist II

Apr. 2023 – Present

Data Scientist I

Oct. 2022 – Apr. 2023

Bactobio | London, UK

Lead data science projects in bacterial generation in a multifaceted role:

- Work collaboratively with lab scientists to design experiments that are informative to ML and statistical approaches.
- Develop and apply ML and statistical approaches, and deliver meaningful and actionable results to lab scientists.
- Build and maintain our ML infrastructure for standardised and reproducible experimentation.
- Exploratory data science: developing new techniques for getting the most out of our data, drawing on developments in the public domain and fine-tuning them for our requirements, integrating in-house and public data to improve the power of our models.

My experience of both computational and wet lab biology is essential for the successful partnership between myself and the wet lab team. I chair the data science research group, contribute daily to code reviews and am active in shaping our software engineering stack and maintaining our software standards.

Skill Highlights: Probabilistic models, PyTorch, Scikit-Learn, SHAP, Interpret-ML, Genomics, MySQL, Git, MLflow, Optuna, Facebook-Hydra, Pandas, Plotty, Autoencoders, SQL-alchemy, Docker

Postdoctoral Research Associate

Sep. 2021 – Oct. 2022

University of Cambridge | Cambridge, UK

Developing machine learning pipelines to analyse microbial community data and predict experiments across several projects. In all projects I worked collaboratively with experimentalists

- Parameterising genome-scale metabolic models with multiomic data.
- Predicting growth requirements of ‘unculturable’ microbes.
- Developed ML approaches to predict the effect artificial sweeteners on microbial growth.

Skill Highlights: genome-scale metabolic models, Autoencoders, PyTorch, Git, MLflow, Optuna, Facebook-Hydra, Plotty, SLURM, Scikit-Learn, Pandas

Data Scientist / Bioinformatics Intern

Mar. 2021 – Jun. 2021

Hummingbird Diagnostics | Heidelberg, Germany

Developed machine learning pipelines for early-stage lung cancer diagnosis using bulk RNAseq data from blood samples.

- Used PyTorch extensively to build, train and test ML architectures.
- Implemented variational-autoencoders to improve disease classification metrics.
- Compared batch effect removal techniques and developed methods to assess how our models compensate for batch effects.
- Gained valuable experience of working in a start-up environment and working on collaborative code being used in production.

Skill Highlights: PyTorch, Git, MLflow, Scikit-Learn, Pandas, RNAseq, R, DEseq2, Anndata, Optuna, Facebook-Hydra, SLURM (HPC)

PhD Student – Computational Biology

Sep. 2016 – Mar. 2021

University College London | London, UK

- Implemented approximate Bayesian computation methods to perform model selection, and parameter estimation.
- Developed software to design genetic circuits for bacteria, intended to stabilise microbial communities.
- Developed a methodology to design engineering strategies that can be used to produce different community population dynamics still being used to inform new work in Barnes lab.

Publications and Presented Talks:

- [Chaos in small microbial communities](#) – PLOS Computational Biology
- [Automated Design of Synthetic Microbial Communities](#) - Nature Communications, 2021
- [Single strain control of microbial consortia](#) - Nature Communications, 2021
- [From Microbial Communities to Distributed Computing Systems](#) - Frontiers in Bioengineering and Biotechnology 2020
- [Computational Design of Synthetic Microbial Communities Incorporating Competitive and Cooperative Interactions](#) - 2nd International Conference on Microbiome Engineering (ICME 2019)

Skill Highlights: Scipy, CUDA, C++, Boost C++, Scikit-Learn, Numpy, Pandas, Bayesian statistics, model selection, parameter estimation, matplotlib, ggplot, seaborn, Sun Grid Engine (HPC)

Vision and AI Summer Placement

Mar. 2019 - Jun. 2019

Imagination Technologies | London, UK

Developed model systems and algorithms to improve the compilation time and resource efficiency of neural networks being run on dedicated hardware designed for smartphones and autonomous vehicles.

- Learned about architectures used in leading convolution neural networks used for image classification.
- Learned to build and train deep learning models using Tensorflow and Keras frameworks.
- Expanded and applied my knowledge of numerical optimisation algorithms, improved performance of company hardware.

Skill Highlights: PyTorch, Keras, Caffe, Tensorflow, CNNs, Computer Vision, Agile, Perforce (version control), parameter optimization

Innovation Consultant

Sep. 2015 – Jul. 2016

Agilisys | London, UK

- Performed research to support technology board decision making regarding new product development and company acquisitions.
- Worked with the bid team to compete for new contracts and assess industry competition.
- Developed visualisations and data reporting dashboards from SQL databases, providing clients with customer insights and working with clients to improve customer engagement.

Skill Highlights: SQL, competitor analysis, acquisition analysis, bid development, stake-holder development, partner management

Undergraduate Research Experience

July. 2013 – July. 2015

King's College London | London, UK

- Conducted miRNA knockout screen in *Drosophila*.
- UCL iGEM 2014 – Engineered new metabolic pathways into *E.coli*.
- Awarded *Gowland Hopkins Biochemistry Prize* for knock-out screen in fission yeast, confocal microscopy, and automated image analysis.

Skill Highlights: Image segmentation, confocal microscopy, *Drosophila* husbandry, molecular biology, cloning

TEACHING

SYSMIC | Teaching maths and data science to biology PhD students

Sep. 2016 – Mar. 2021

iGEM | Teaching mathematical modelling and mentoring projects of summer students

Sep. 2016 – July. 2020

EDUCATION

University College London | PhD

Sep. 2016 – Mar. 2021

King's College London | Bachelor of Science in Biomedical Science

Sep. 2012 – July. 2015

- Upper second-class Hons.