Owais Siddiqi

omsiddiqi01@gmail.com **|** https://owaissiddiqi.co.uk/ **|** London, United Kingdom

EDUCATION

Imperial College London   **Oct. 2022 – Oct. 2023**
MSc in Biomedical Engineering - Merit

- **Reinforcement Learning:** Explored advanced algorithms including Deep Q Networks (DQN), Proximal Policy Optimisation (PPO), Soft Actor Critic (SAC), Monte Carlo methods, Temporal Differences, and Dynamic Programming.
- **Statistics And Data Analysis:** Proficient in Markov Decision Processes, Bayesian Probability, and both Linear and Non-Linear Regression techniques.
- **Image Processing:** Mastered techniques such as Image Classification, Object Detection, Image Segmentation, Fourier Transformations, and the application of Neural Networks in image data.
- **Brain Machine Interfaces:** Acquired knowledge in Neural Decoding, Information Theory, Time Series Analysis, K-Nearest Neighbours, Naive Bayes Theorem, SVM, Random Forests, and Multi-Layer Perceptron's.
- **Mathematical Methods:** Developed a strong foundation in Graph Theory, Time Series Analysis, Data Visualisation Techniques (PCA, ICA, SVD), and Linear Algebra.

Queen Mary University of London   **Sept. 2019 – Sept. 2022**
BEng (Hons) in Biomedical Engineering – 1st Class

- **Mathematics and Computing for Engineers I & II:** Expertise in Calculus, Differential Equations, Sequences, EigenValues, Vectors, and Gaussian Elimination methods.
- **Vector Calculus for Engineering:** Strengthened skills in Multivariable Calculus, Linear Algebra, and Second Order Differential Equations.

WORK EXPERIENCE

Genevation ltd.   **Apr. 2024 - Present**
Junior Genomic Data Scientist:

- **Cancer Biomarker Prediction**: Helped develop an AI model with **PyTorch** for the prediction and identification of neoantigen biomarkers, facilitating personalised vaccine production for cancer immunotherapy. Collaborated with cross-functional teams to refine model architectures and enhance development pipelines, enabling targeted cancer therapies.
- **Bespoke VCF Production**: Designed a specialised pipeline to convert raw sequence data (FASTQ) into customised VCF formats, focusing on efficient data processing tailored for downstream genetic analyses of tumour-specific biomarkers, including somatic/germline mutation annotation.
- **Bioinformatics and Computational Tools**: Applied bioinformatics tools for RNA-seq data processing and variant calling, leveraging **containerised environments** with **Docker** and managing **Linux-based virtual machines** on **Google Cloud Platform (GCP)** for high-performance, scalable data analysis.
- **Optimised Data Processing**: Reduced the genetic downstream analysis of neoantigen biomarkers **from a week to a few days**, significantly accelerating the development timeline for personalised vaccine production.
- **Continuous Learning**: Engaged in ongoing professional development to stay updated with advancements in genomics, AI, and bioinformatics. Participated in relevant workshops and online courses to sharpen skills in computational biology and AI-driven data analysis.

PDUK ltd.   **Aug. 2023 – Jan. 2024**
Data Scientist Intern:

- **Data Engineering:** Employed SQL queries to import, organise, and clean data in **MySQL**, resulting in a **25% reduction in data processing time**, ensuring comprehensive information retrieval on **3000+ customers, products and transactions**. Enhanced overall efficiency and improved data accessibility through streamlined processes.
- **Data Visualisation:** Employed Python tools such as **Pandas** for data visualisation, creating insightful charts and graphs with **Matplotlib** and **Seaborn**. Effectively communicated trends, patterns, and relationships within the data via reports, facilitating informed decision-making.
- **Predictive Modelling:** Leveraged advanced AI techniques, including transfer learning and pre-trained recurrent neural network (RNN) models within **PyTorch**, for predictive modelling. Adapted existing RNN architectures for

time series forecasting of customer purchase history, resulting in a **20% improvement in forecast accuracy**, leading to a **10% reduction in inventory costs**.
- **Customer Segmentation:** Utilised K-means clustering with the Elbow Method to achieve optimal cluster determination for segmenting customers based on purchasing behaviour. This approach led to a **15% increase in conversion rates** through targeted marketing strategies tailored to distinct customer groups.

Imperial College London   **Mar. 2023 – Sept. 2023**

MSc Project:
- **Bespoke Model Development**: Pioneered the creation of a cutting-edge, novel, end - to - end regression AI model using artificial microscopy data, revolutionising cell differentiation through deep learning. Resulted in an exceptional **80%+ accuracy** in seamless single stage cell quantification.
- **Data Management & Augmentation**: Optimised data preprocessing by implementing advanced methods, ensuring meticulous preparation of microscopy images for analysis. Implemented data augmentation techniques, which expanded the training dataset, resulting in a **30% increase in generalisation capabilities** and a **15% improvement in predictive accuracy**.
- **Tool Utilisation**: Led the seamless integration of a comprehensive toolset, featuring **Pandas**, **Matplotlib**, **OpenCV**, **TensorFlow**, **Keras**, and **NumPy**. This initiative optimised data preprocessing, parameter extraction, and model implementation, ensuring an efficient and effective workflow.

Imperial College London   **Oct. 2022 – Dec. 2022**

Neural Data Analysis:
- **Data Decoding:** Utilised advanced libraries such as **Matplotlib**, **NetworkX**, and **SciPy** to decode and analyse complex neural data from mice, extracting meaningful insights from intricate datasets.
- **Visualisation Expertise:** Orchestrated the development of a diverse array of visualisation tools, including autocorrelation plots, recurrence plots, and raster grams, to present a comprehensive and insightful perspective on neural activity patterns.

Queen Mary University of London   **Sept. 2021 – May. 2022**

Third Year Project:
- **Fuzzy Modelling:** Led the charge in researching and implementing **ANFIS** combined with a bagging ensemble approach, crafting a sophisticated multi-class classifier for brain-machine interfaces. Achieved a remarkable 64% accuracy in classifying/predicting new **SSVEP** data using **MATLAB**.
- **Recognition & Presentation:** Secured a prestigious position among the **top 20% of engineering projects**, showcasing an innovative approach. Presented compelling findings at the annual Industrial Liaison Forum, garnering accolades for groundbreaking research.

The Institute of Cancer Research   **Jun. 2021 – Aug. 2021**

Summer Intern:
- **Neural Network Development:** Led the development of a novel few-shot learning Siamese Network with a triplet loss function using **OpenCV**, **Keras**, and **TensorFlow**. Achieved a remarkable **70% accuracy score**, showcasing innovation and expertise.
- **Research Contribution:** Established the creation of an innovative imaging phantom, enabling independent experiments and generating valuable data and insights. Contributed significantly to the success of the research project, resulting in a **20% increase in experimental efficiency and data quality**.

Publications

Siddiqi, O., Winfield, J., Sormpas-Petridis, K., Harris, E., Ramkumar, A., Candito, A., Görner, S., Messiou, C., Blackledge, M., & Thrussell, I. (2022). Investigating the Repeatability of Multifrequency Magnetic Resonance Elastography applied to a Soft Gelatine Phantom. Presented as a digital poster at the ISMRM 31st Conference & Exhibition, Excel London, 7th – 12th May 2022.