(+44) 7450050241
York, UK
hpetchsmith98@gmail.com

Harry Petch-Smith

Msc Bioinformatics

GitHub: hpetch-smith
LinkedIn: Harry-Petch-Smith

Education

MSc Bioinformatics - Distinction, *University of Birmingham* 09 2022
Modules: *Essentials of Mathematics and Statistics, Genomics and Next Generation Sequencing, Data Analytics and Statistical Machine Learning, Metabolomics and Advanced omics Technologies, Computational Biology for Complex Systems*

BSc Biomedical Science With a Year In Industry - 2:1, *The University of Sheffield* 07 2021

A levels, *South Hunsley School and Sixth Form College* 07 2017
Psychology (A*), Chemistry (B), Biology (B)

Experience

Volunteer Instructor/Introduction to Python Programming **Oct 2021 — Dec 2021**
Code First Girls *Remote*

- Taught beginners python course with the aim of introducing more women to the field of tech.
- Collaborated in a group of three instructors to deliver weekly lessons.
- Managed lesson planning alongside masters studies.
- Led interactive sessions in an engaging manner, with a mix of activities and lecture format.
- Clear communication of programming principles to women from many backgrounds.

Student Biomedical Scientist **06 2019 — 07 2020**
Hammersmith Medicines Research *Wembley, London*

- Year in Industry Placement at UK leading clinical trials company.
- Responsibilities: Sample processing, Monitoring patient data, Validation of new assays, Quality control of documentation, Stock/supplies management
- Gained experience working in a fast paced professional environment that valued both team work and independence

Projects

A Baysian non Parametric Approach to Time Series Modelling
Machine Learning

- **Research Question:** Can multiple output Gaussian processes accurately model and predict time series data.
- **Objective:** Test multiple output GPs on two datasets; Simulated dependant time series data, and blood glucose monitoring data.
- Implemented single output gaussian process using **R.**
- **Simulated** time series datasets in **R and Python.**
- **D**eveloped a Multiple Output Gaussian Process from scratch in R.
- Results reported in **nature article format** using **Latex**

Identification of Cyp21a2-dependent, glucocorticoid regulated target genes in zebrafish larvae
Data Science/Bioinformatics

- **Research question:** What are the effects of glucocorticoid deficiency on gene expression in Zebrfish larvae?
- **Objective:** Given a **large dataset** of zebrafish gene expression counts; **combine and analyse** the data to identify differentially expressed genes.
- constructed a Bioinformatics pipeline to conduct **data preprocessing** and **analysis and visualisation** in **R**.
- Used **R** to combine multple datasets for easy comparison
- Performed **quality filtering and normalisation**
- Implemented **principle component analysis, MA plotting, Volcano plots** and **downstream gene ontology analysis**
- Maintained a lab book recording each step generated as an R markdown file.
- Work was awarded a **first class grade**

Skills

Languages	Bash, Python, R, HTML/CSS, Javascript
Tools	Linux command line, Rstudio, Jupyter Notebook, Spyder, LateX, Markdown, Visual Studio Code.