Sam Robson

Bioinformatics Lead at Centre for Enzyme Innovation

Portsmouth, United Kingdom

(+44) 07863 328347

linkedin.com/in/samrobson/

irunfasterthanmycode.github.io

scb.robson@gmail.com

About me

Computational biologist with a strong mathematics and statistics background. Extensive experience of maintaining, processing, and analysing Big Data from next generation DNA sequencing. Expertise in a wide variety of data mining, data visualisation, deep learning, and machine learning methods to identify statistically significant trends in high-dimensional data. Bioinformatics Lead at the Centre for Enzyme Innovation and Faculty Bioinformatics Lead at the University of Portsmouth.

Skills

R

perl

MySQL

python

TensorFlow

C++

RStudio⋆5 Git⋆4 LaTeX⋆4

(*)[The skill scale is from 0 (Fundamental Awareness) to 6 (Expert).]

Education

2004-2008 PhD in *Mathematical Biology and Biophysical Chemistry*
University of Warwick

2003-2004 MSc in *Mathematical Biology and Biophysical Chemistry* Class: 1^{st}
University of Warwick

1861-1863 BSc in *Mathematics (Hons)* Class: 2:1
University of Warwick

Experience

Since 2017 Centre for Enzyme Innovation, University of Portsmouth Bioinformatics Lead
Responsible for a small team of bioinformatics researchers to use machine learning algorithms to identify potential novel plastic-degrading enzymes from microbiological genome data

2014-2017 Sam Robson Consulting (Self-Employed) Statistics Consultant
Identified key factors influencing doctor burnout through the use of multivariate mixed-effects regression analysis in the largest study of doctor burnout yet conducted

2010-2017 The Gurdon Institute, University of Cambridge Bioinformatician
Developed and maintained pipeline and database for in-depth processing, mining and analysis of high-dimensional genome wide DNA sequencing data

2008-2010 Wellcome Trust Sanger Institute Statistical/Mathematical Biologist
Responsible for maintaining, processing and normalizing genome-scale Big Data, including sample QC, concordance analysis with previously published SNP data, data optimization and case-control association testing

Analysis Skills

Big Data wrangling, maintenance and analysis of extremely large data sets, pipeline development for high-throughput DNA sequencing data, normalization of complex data sets, data visualisation, machine learning, regression analysis (linear and generalised linear models), classification models (*unsupervised*: K-means, hierarchical clustering, mixture models; *supervised*: random forest, K-nearest neighbour, SVM), PCA dimensional reduction

Leadership Skills

Bioinformatics Lead and Board Member at the Centre for Enzyme Innovation in charge of a small group of bioinformatics researchers, supervisor for a number of PhD students, Faculty Bioinformatics Lead working with and advising researchers throughout the University on a number of distinct projects, explaining complex technical outputs to non-experts and management

Communication Skills

Excellent communication skills at the interface between Life Sciences, able to effectively explain complex analysis concepts to non-specialists, able to maintain extremely high standards when working across a large number of disparate projects, ability to work effectively across disciplines, experience collaborating with industry, excellent ability to identify and solve problems, ability to work on own initiative or as a keen team player, highly motivated

Awards

2019 Award of £6 million from Research England E3 Fund
2019 Award of £5,000 Google Cloud Platform research credits
2017 Awarded CStat and CSci Membership of the Royal Statistical Society