

James O'Reilly

Bongenotenlaan 137, Leuven 3000, Belgium

☎ (+353) 861210196 | ✉ jamesdanielloreilly@gmail.com | 📺 [jamesdanielloreilly](#) | 📄 [jamesdanielloreilly](#)

Mathematics & Computer Science graduate currently in final year of a BioInformatics MSc. Seeking machine learning and genomics research roles in industry. Significant experience in implementing pipelines for statistical analysis of NGS data, machine learning applications in genomics, and commercial software engineering. Available for employment June 2021. As an Irish citizen I have full employment and residency rights in the UK.

Education

KU Leuven

Leuven, Belgium

M.Sc. IN BIOINFORMATICS

Sept 2019 - Present

- Expected grade: H1.1 (First Class Honours)
- Thesis: Using machine learning to disentangle deterministic and stochastic epigenetic sources of cellular heterogeneity in cancer. Supervisor: Prof. Bernard Thienpont at the VIB Lab for Functional Epigenetics, Leuven.
- Bioinformatics modules: genome analysis, regulatory and comparative genomics, quantitative genetics, statistical methods for bioinformatics, management of large-scale omics data.
- Mathematics modules: Neural networks and deep learning, Bayesian analysis, applied multivariate statistical analysis, machine learning, dynamical systems, support vector machines.

University of Bristol

Bristol, England

B.Sc. IN MATHEMATICS AND COMPUTER SCIENCE (JOINT HONOURS)

Sept 2016 - June 2019

- Final grade: H2.1 (67%)
- Maths: foundational courses in calculus, probability, analysis, and linear algebra. Specialised in discrete mathematics, graph theory and information theory; focusing on codes, communication, and cryptographic schemes.
- Computer Science: strong understanding of machine learning (77%) and computational neuroscience (84%). Extensive programming and development work in Python.

Work Experience

DataCamp

Leuven, Belgium

SOFTWARE ENGINEER INTERN

Jun 2020 - Aug 2020

- Contributed to a variety of projects as part of the engineering team. Implemented automated testing of code correctness in Python, built and optimised Docker images used across all of DataCamp's online courses, and improved the python backend and feedback system.
- Gained experience with DevOps infrastructure: shipping production-level code in an agile environment using Git, Jira, and CI.

University of Bristol - Dept. of Computer Science

Bristol

COMPUTATIONAL NEUROSCIENCE RESEARCH INTERN

Jun 2019 - Aug 2019

- Investigated the application of deep neural networks in determining the statistical structure of language.

Significant Projects

- **Thesis KU Leuven:** Investigating epigenetic sources of cellular heterogeneity in cancer. Using single-cell latent variable models to disentangle deterministic and stochastic sources of heterogeneity in scRNA-seq data of CRISPR knockout-screens in mouse lung adenocarcinomas.
- **Integration of single-cell multi-omics data:** Used Multi-Omics Factor Analysis to integrate PBMC data from the novel 10X-Genomics single-cell multiome assay. Identified the combined contributions of the transcriptome and epigenome to cellular variation. Performed cell-type annotation using the Human Blood Atlas. In collaboration with Prof. Stein Aerts at the VIB Laboratory for Computational Biology. [GitHub](#).
- **Prediction of single-cell expression data:** Imputed gene expression from single-cell chromatin accessibility data using ArchR. Evaluated prediction accuracy with single-cell multiome data and trained machine learning models to predict expression for specific gene ontologies. [GitHub](#).
- **Factorial single cell latent variable modelling of PBMC data:** Used the Slalom package to perform f-sclVM on PBMC data, using predefined genesets to identify specific biological factors contributing to transcriptional variability. Identified and regressed out technical and biological factors for downstream analyses.
- **Chromatin accessibility analysis in NSCLC:** Implemented a full ATAC-seq pipeline for analysis of chromatin accessibility in non-small lung cancer cells. Investigated epigenetic state shifts between active and dormant states across treatments. Identified TF binding motifs and upregulated promoters of EMT. [GitHub](#).
- **Image classification with deep neural networks:** Implemented deep feature learning with a deep convolutional NN for image classification in Matlab. Worked to optimise network topology and hyperparameters. Critically evaluated different learning algorithms for NNs. [GitHub](#).
- **Generative neural networks for image generation:** Implemented different generative models, including deep Boltzmann machines and generative adversarial networks (GANs). Investigated convergence and stability of DCGANs, with a focus on Wasserstein GANs. [GitHub](#).
- **Bristol Final Year Project:** Conceptualised, designed, and evaluated a virtual reality learning environment (VRLE) for calculus education. Demonstrated that student understanding of abstract mathematical concepts could be improved using VR technology, while facilitating both distance learning and learning for disabled students. [GitHub](#).

Skills

Programming	Advanced: Python, R, Matlab, and Bash. Experience with: Julia, C, C#, Java, JAVAScript, haskell, LaTeX.
DevOps	Git, Docker, CircleCI, Jira
Bioinformatics	Seurat, Signac, Slalom, ScanPy, Cell Ranger, DeepTools, MACS2, Bowtie2, MOFA, ArchR, DeSeq, and more