DR JOEL MILLER

Principal Data Scientist

joelsamiller@gmail.com | +447787808981 | London, UK
joel-miller-0630a21aa | JoelM935 | 0000-0001-7775-8803

EXPERIENCE

Principal Data Scientist
UK Health Security Agency

July 2023 – Ongoing | London

- Lead developer on a Python package for creating automated slide decks. Data is pulled from S3, SQL databases or local files, with pandas used for data manipulation and Ploty for data visualisations. This package has enabled hours of recurring work to be automated.
- Created a Python package to simplify authentication when connecting to AWS from our analytics platform. Reduces terminal commands and many lines of code to a single line. This has been well-received by other data scientists and has increased efficiency.
- Took an existing internal report that was manually created each month using Excel spreadsheets, with data received by email. Designed and built a new data pipeline and automated generation using Python. This saves three days of work each month.
- Automated two GOV.UK reports using R markdown, saving 10 days of analyst time per production cycle and improving quality.
- Leading a review of wastewater analysis conducted during the COVID-19 pandemic. Investigating the discrepancy between the infection rate predicted from wastewater by linear regression models and the COVID-19 Infection Survey.
- Conducted a data quality review of a SQL database following the identification of an ingestion error. Measured impacts and made recommendations to key stakeholders.
- Created and delivered training sessions on Python and Git to interns. These were rated in the top three sessions they received.
- Managed and trained three employees who were promoted shortly afterwards, with two going from non-coding to coding roles.

Senior Data Scientist
UK Health Security Agency

April 2022 – July 2023 | London

- Improved a high-profile COVID-19 product built using a large object-oriented Python codebase and AWS. Reduced 60k lines of code to 10k, significantly increased speed and reduced operating costs by 85%. I also improved the visualisations to improve clarity and provide a consistent look and feel.
- Supported incident responses, delivering data analysis and visualisations on high-priority public health situations at pace.
- Worked to a tight deadline on a greenfield project using Python to extract data from PDF forms into a tabular form. This provided a solution where the existing manual process could not process forms fast enough during an outbreak.
- Lead a team of 5 data scientists on a project to automate the production of a regular GOV.UK report using R-markdown. This reduced the production time from 2 days to ~30 seconds.
- Trained an LLM using Pytorch and HPC, producing a binary classification model with >90% accuracy. This model was then applied to free text data to improve reporting on Avian Influenza.

ABOUT ME

I have over seven years of experience applying data science techniques across academia and public health, and nearly 15 years of experience in computer programming over eight different languages. I am a self-driven individual with a scientific mindset, who loves building tools, solving problems and learning. I am also an international athlete who has represented Great Britain at the senior level, competing in many high-pressure scenarios.

SKILLS

Python | R | SQL | C | Git | GitHub | GitLab | Linux | Windows | MacOS

Regression | Calculus | Simulations | Statistics | Mechanics | EDA | Machine Learning | Linear Algebra

Communication | Problem Solving | Leadership | Teamwork

EDUCATION

Ph.D. Physics
University of Nottingham

Oct 2017 – August 2021

Thesis title: Searching for the Shadows of Giants: The Absorption Signatures of Proto-clusters in the Lyman-alpha Forest

Publications:

- Miller, J. S. A., Bolton, J. S., & Hatch, N. A. (2021). *Monthly Notices of the Royal Astronomical Society*, *506*(4), 6001–6013. doi:10.1093/mnras/stab2083
- Miller, J. S. A., Bolton, J. S., & Hatch, N. (2019). *Monthly Notices of the Royal Astronomical Society*, *489*(4), 5381–5397. doi:10.1093/mnras/stz2504

M.Sci. Physics with Astronomy
University of Nottingham

Sept 2013 – June 2017

First Class

SKILLS PROFILE

Self Learning

One of the many skills I developed during my PhD was teaching myself new concepts and topics, in a few days, with minimal or no supervision.

- At the beginning of my PhD, I taught myself to write code in C to run complex calculations on data from state-of-the-art hydrodynamical cosmological simulations.
  - This code allowed me to determine column densities of neutral hydrogen gas along specified sightlines and then generate mock spectra from the simulations.
- At the same time, I learnt how to use the IDL and Python languages to analyse the data extracted from the simulations and produce visualisations.
- I applied these skills in my non-academic career. When I was placed on R and SQL projects, I learnt these languages on the job and quickly produced results.

Critical Thinking and Problem Solving

For my PhD, I spent 3 and a half years conducting original research in my field of astrophysics. I continuously analysed my results using statistical methods such as Welch's t-tests, bootstrapping and Spearman's rank coefficients, before placing them within the broader context of the literature of my area of research. By doing this, I drew new conclusions and decided on the next questions to answer with my research.

Communication and Presentation

- Throughout my PhD, I gave several intra-departmental talks and a talk at an academic conference on my research.
- I delivered many outreach talks on my research, and astronomy in general, to a range of audiences from primary school pupils to prospective undergraduates. I was even interviewed by the BBC about astrophotography [Link].
- I wrote two published papers and three annual reports on my research, in addition to my thesis.
- My communication skills have been further bolstered through multiple sports coaching roles.

As a data scientist, I have applied these skills when liaising with customers from varying technical backgrounds—to capture their requirements and provide a product to suit them—and delivering presentations to technical and non-technical audiences at directorate and divisional meetings.

Teamwork and Leadership

- I work on projects in cross-disciplinary teams using an Agile framework.
- I use GitHub/GitLab to work collaboratively on codebases and Jira boards to track progress and capacity.
- I delivered training on R and Git to non-data scientists so that they could contribute to shared projects, and to facilitate smooth product handover.
- I have taken on the line management of multiple graduate and senior data scientists (one of whom was recently promoted).

Outside of work, playing competitive ultimate from university to international levels enabled me to develop robust teamwork skills. Furthermore, I succeeded in leadership roles both as an on-pitch captain and through several committee roles, including being twice elected as the university club president.

CERTIFICATIONS

Machine Learning, Modeling, and Simulation Principles

MIT xPro

April 2024

PERSONAL PROJECTS

Disc Tracker

Computer Vision

- Takes stereo input of video files where a flying disc is thrown.
- Implements a Gaussian mixture model for background subtraction.
- Blob detection is used to identify foreground objects.
- Jonker-Volgenant assignment algorithm with cost matrix given by distances between previous and new locations used to track objects over time.
- Tracks corresponding to the flying disc are de-projected to give the path in three dimensions.

Virtual Stall Count

Android App

- Simple Android app designed to support COVID-19 measures for the sport of ultimate [Rules of Ultimate].
- Plays a voice counting numbers from the selected start number up to 10 or until the user releases the button.

ACHIEVEMENTS

Competitive team sports

I am a keen ultimate player, competing for my university's 1st teams, high-level club teams and Great Britain. Some highlights of my playing career include:

- Winning gold at the 2019 European Championships with Great Britain.
- Winning 5 BUCS gold medals and 2 regional titles with university teams.
- Captaining my university mixed indoor 1s through an undefeated season to become national champions.

INTERESTS

Ultimate Frisbee | Astronomy | Gym | Photography | Playing Piano | Coding | Computer Games | Science-Fiction