

Eshwen Bhal

DATA SCIENTIST

22 Rownham Court, Rownham Mead, Bristol, United Kingdom. BS8 4YE

☎ +44 (0) 78414 09961 | ✉ milanman.eshb@gmail.com | 🗣 eshwen | 🌐 eshwen-bhal-phd-714557195

Profile

I am a data scientist developing a web scraping pipeline for the Consumer Price Index alternative data sources transformation. I have recently completed a PhD in particle physics, searching for dark matter using big data from the Large Hadron Collider (LHC) at CERN. This involved statistical analysis of large datasets with expertise in Python, and solving wide ranging problems from physics standpoints to programming challenges. Additional responsibilities included providing detector-related expertise and software, teaching undergraduates, and public outreach. Aside from work, I partake in many activities such as taekwondo, weight lifting, hiking, and skiing.

Experience

Office for National Statistics

Newport, United Kingdom

DATA SCIENTIST

2021 – Present

- **Web scraping** products and prices from various retailers to assist in creating a more accurate Consumer Price Index.
- Building spiders with Scrapy in Python to reliably scrape websites. Data cleaning and validation with **NumPy** and **pandas**.
- Spurred good coding practices like **unit testing**, efficiency, and automated documentation.
- Maintaining and improving deployment on **Google Cloud Platform** for scheduled, automated execution. Starting to learn Docker, Terraform, and Kubernetes for additional responsibilities.

University of Bristol

Bristol, United Kingdom

COMPUTATIONAL PHYSICS TUTOR

2019 – 2020

- Taught programming in Python to third year undergraduate students, with applications for **numerically solving physics problems** with NumPy and visualisation with matplotlib. Emphasis on **vectorisation** over looping, clear code, and presentation of results. Formally marked assignments and provided feedback.

Education

University of Bristol

Bristol, United Kingdom

DOCTOR OF PHILOSOPHY IN PHYSICS

2016 – 2021

- Thesis: **Searches for dark matter with a focus on invisibly decaying Higgs bosons using the full Run-2 dataset of the CMS experiment at the LHC** — *Under supervision of H. Flücher*
 - Explored various physics models in search of dark matter by analysing data from LHC's CMS experiment. Set **world leading limits** on Higgs boson decay to invisible states.
 - Executed robust, comprehensive statistical analysis with hypothesis testing on real and simulated data, meticulously documenting concepts, results and code.
 - Composed robust, efficient code to simulate data and perform each step of analysis. Written predominantly in Python, leveraging modern data science tools, vectorisation, and distributed computing to process multiple terabytes of data.
- **Postgraduate student representative** for particle physics, 2019–20 — *Role in Student-Staff Liaison Committee*
- Long term attachment at world's largest particle physics laboratory **CERN** — *18 month placement abroad in Switzerland*
- **Calorimeter Layer-2 on call expert** and **Level-1 Trigger shifter** — *Additional responsibilities with CERN*
 - Assisted in detector operations and monitoring so CMS experiment could take data efficiently and operate smoothly.
 - Developed and deployed software for subsystem of Level-1 Trigger to apply corrections and calibrations to data on the fly.

- Grade: **First Class** — 77% overall mark (4.0 GPA equivalent)
- Dissertation: **Simulations of Exoplanet Light Curves** — Under supervision of T. Harries
 - Developed software in C to simulate photons interacting simple planetary atmospheres, producing light curves akin to data from telescopes. Visualised model planets with maps of density and composition.
 - Able to model more complex atmospheres for comparison to real exoplanets to infer their composition.
 - Utilised Monte Carlo random sampling for scattering of photons, and parallelisation to efficiently run the code over millions of them.

Skills

Software and computing expertise

- Developed analysis software for dark matter searches and detector-effect calibrations in Python and C++. Optimised to run complex set of modules over billions of events on timescale of 1 hour.
- High proficiency in Python and use of data processing libraries like numpy and pandas. Using fully-vectorised approach often to reduce inefficiency.
- Visualisation with matplotlib and ROOT. Diverse range of graphs, charts, tables, and figures used to showcase results. Formal presentations of results with LaTeX and Microsoft PowerPoint.
- Utilisation of high performance computing and parallel programming with HTCondor and various computing grids.
- Expertise in git, bash, command line interface on Unix systems (Scientific Linux, CentOS, macOS).
- Meticulously and clearly documented code and running instructions for users.

Collaboration

- Belonging to, and working alongside, several groups of around a dozen people as well as a global research collaboration of over 4,000 people.

Problem solving

- Principal component of any physics degree. As an undergraduate, conducted more traditional pen-and-paper problem solving in many topics. As a PhD student, wrote code to solve physics problems numerically and perform data analysis for high energy particle physics.

Communication

- Presented formally in my PhD at all levels: regular working group updates, research group and collaboration-wide talks, and several conference posters and talks.
- Participated in outreach to the public at the CERN Open Days 2019 and at @Bristol.

Awards

2016	Dean's Commendation , in recognition of outstanding achievement in my fourth year	University of Exeter
2015	Physics Award , for being one of the three students with the highest marks in my third year	University of Exeter
2014	Dean's Commendation , in recognition of outstanding achievement in my second year	University of Exeter
2013	Physics Award , for being one of the three students with the highest marks in my first year	University of Exeter
2013	Dean's Commendation , in recognition of outstanding achievement in my first year	University of Exeter