Henry Wilde

PhD thesis (ongoing)

Title **New methods for algorithm evaluation and cluster initialisation with applications to healthcare**

Supervisors Dr Jonathan Gillard, Dr Vincent Knight, Mr Kendal Smith (NHS Wales)

Description The purpose of this project is to investigate and model variation in the service and cost of treating hospital patients in South Wales; this is achieved by identifying structures in routinely gathered, administrative healthcare datasets via clustering. In addition, this thesis considers some of the issues surrounding the process of evaluating an algorithm's quality, and presents novel methods for enriching the picture painted by traditional processes.

Education

2017–present **PhD Applied Statistics, OR and Data Analytics (expected: Spring 2021)**, *Cardiff University*.

- *Research interests:* healthcare modelling, unsupervised learning, algorithm evaluation, game theory, research software development.
- *Key skills:* confidence in the principles of data science and mathematical programming, ability to develop industrial relationships, proficiency in LaTeX and Python.

2014–2017 **BSc Mathematics (First Class Honours)**, *Cardiff University*.

- *Key areas of study:* mathematical methods for data mining, game theory, algorithms and heuristics.
- *Projects included:* building a simulation of a hospital emergency department, and a principal analysis of two game-theoretic strategies within an Iterated Prisoner's Dilemma tournament.

Relevant experience

2019–present **Final year project allocation**, *Cardiff University*.
I have installed a new framework for allocating dissertations to final year students in the School of Biosciences, Cardiff University. By making use of the Matching library I maintain, I have been able to reduce the bulk of their workflow down to a matter of seconds. In doing so, the allocation is both mathematically fair and student-optimal, with almost all students being allocated to their first or second choice of project.

2018–2019 **MMORS dissertation co-supervisor**, *Cardiff University*.
I assisted in the supervision of a MMORS final year project with Dr Vince Knight conducting an empirical study of Folk Theorems in repeated games. My primary role was to consult on how best to develop the supporting research software, and in the writing process.

2018–2019 **Assessment advisor**, *Cardiff University*.
I have acted as part of the assessment team for a MSc module on computational methods. The assessment consists of a two-day hackathon involving several dozen students and my role has been to gauge and rank their abilities to work in teams to develop a piece of software for a certain mathematical task.

2018–present **Maths Support assistant**, *Cardiff University*.
I provide mathematical aid to students from all schools within the university as part of the Maths Support Service. This drop-in service affords little scope for preparation and requires me to make use of my nature as a mathematician — to be proactive, analytical and logical — often covering far-reaching branches of mathematics in a session.

2017–present **Module tutor**, *Cardiff University*.
Throughout my time as a postgraduate student, I have supported a number of modules as a tutor, leading regular sessions with students, and assessing their progress. This role has given me countless opportunities to teach in group and one-on-one settings, and to play a key role in several active learning schemes.

henrydavidwilde@gmail.com • *daffidwilde* • *daffidwilde*

Publications and pre-prints

[KPW20] Vince Knight, Michalis Panayides, and Henry Wilde. *Python for Mathematics*. 2020. DOI: 10.5281/zenodo.4074114.

[Wil+20] Henry Wilde et al. *Segmentation analysis and the recovery of queuing parameters via the Wasserstein distance: a study of administrative data for patients with chronic obstructive pulmonary disease*. 2020. arXiv: 2008.04295 [stat.AP].

[WKG19] Henry Wilde, Vincent Knight, and Jonathan Gillard. "Evolutionary dataset optimisation: learning algorithm quality through evolution". In: *Applied Intelligence* 50.4 (2019), pp. 1172–1191. DOI: 10.1007/s10489-019-01592-4.

[WKG20a] Henry Wilde, Vincent Knight, and Jonathan Gillard. *A novel initialisation based on hospital-resident assignment for the k-modes algorithm*. 2020. arXiv: 2002.02701 [cs.LG].

[WKG20b] Henry Wilde, Vincent Knight, and Jonathan Gillard. "Matching: A Python library for solving matching games". In: *Journal of Open Source Software* 5.48 (2020), p. 2169. DOI: 10.21105/joss.02169.

Software projects

[Kni+19] Vince Knight et al. *Blackbook: Black for Jupyter notebooks*. 2019–present. DOI: 10.5281/zenodo.2553362.

[WK18a] Henry Wilde and Vince Knight. *EDO: A library for generating artificial datasets through evolution*. 2018–present. DOI: 10.5281/zenodo.2552890.

[WK18b] Henry Wilde and Vince Knight. *Matching: A package for solving matching games*. Software. 2018–present. DOI: 10.5281/zenodo.2553125.

Additional activities

Jan 2020 **SIAM UKIE Annual Meeting 2020**, *The University of Edinburgh*.
Received a travel award to present a poster on evolutionary dataset optimisation — including a case study comparing k-means and DBSCAN clustering.

May 2019 **Welsh Mathematics Colloquium**, *Gregynog Hall*, Powys.
Gave an in-depth talk on the mathematical principles of evolutionary dataset optimisation and some of the issues surrounding algorithm evaluation.

Mar 2019 **Data Science Campus Seminar Series**, *Office for National Statistics*.
Spoke on my work into evolutionary dataset optimisation. In particular, its applications to the field of data simulation and synthesis.

Feb 2019 **NHS Wales Modelling Collaborative**, *South Wales*.
Invited to speak about how my data-driven approach to my research would impact the Cwm Taf Morgannwg Health Board.

2018–2020 **Advanced Python Workshop**, *Cardiff University*.
Founded a group for postgraduate students to engage in monthly, tutorial-based sessions about more advanced aspects of Python (such as parallelisation and automated testing) followed by a code clinic.

2018–2019 **STEMLive**, *Cardiff University*.
Volunteering, designing and running a stall at the annual outreach event for school children in the South Wales region to engage with STEM subjects.

2018 **EURO Support for NATCOR Bursary**.
Financial support to attend NATCOR courses in Approximation Algorithms & Heuristics and Predictive Analysis & Forecasting.

References

Academic

Dr Vince Knight
knightva@cardiff.ac.uk

Dr Jonathan Gillard
gillardjw@cardiff.ac.uk

Personal

Mr Simon Naylor
snaylo@hotmail.com

Ms Jessica Lockwood
jess.lockwood95@gmail.com