

Jonathan Phelps

Phone: 07999056012

Email: jphelps130591@gmail.com

PERSONAL STATEMENT

Data Scientist with experience managing and developing in collaborative and independent projects, delivering actionable insights for some of the UK largest retailers in E-commerce. Comfortable building and maintaining production R code in a Linux environment across a wide team. Looking for a role to improve my statistical modelling and machine learning skills, whilst solving challenging real-life problems. Believe in using Data Visualisations to support all key business and R&D decisions and would love a chance to develop these skills further.

WORK EXPERIENCE

Summit

January 2014 – March 2017: Statistician in Analytic Digital Commerce

April 2017 – Present: Data Scientist

Summit are an E-commerce company, with a focus on managing Ad Campaigns on Google and Bing. Work with Argos, GAME, Boohoo, Carpetright amongst other large UK retailers. My role is to maintain and develop Summit's Forecaster technology, a statistical tool that models the profitability of each Product to optimise the bids set in Paid Search and Shopping Comparison. Responsibilities detailed below:

- Leading role in managing the product specifications and time planning for Summit's Budget Explorer tool. Worked with a Junior Data Scientist to develop a product that optimises Budgets across multiple Product Categories, Search Engines and Devices. The team works cross location and required me to lead sessions in the Prague office with our front and back end developers. Saw an increase in return on investment of 4-8% in alpha testing for Argos.
- Researched and implemented a Gradient Boosted regression tree model for predicting Impressions for Products shown in Google Shopping Comparison. Built a cross-validation training suite to optimise the hyperparameter set and validate against the default GLM solution. In testing, we observed an improvement in RMSE between 8-19% across three marketing categories.
- Worked on an A/B Geo-Location test with Argos to assess the incremental benefit of showing Website Banner Ads, in conjunction with Google search results. Presented back to Argos the findings of the test; a valuable experience in communicating actionable insights.
- A large part of the role has been to develop and maintain R code. This included implementing logging and error testing; using the data.table package to improve run time and memory management and speeding up legacy codes with vectorisation, new packages and parallelisation.
- Work with strict deadlines before Christmas Peak, requiring efficient team practices. Work to release cycles with GIT management, and Agile workflow. Best practices include feature & release tests, peer review, code style guidelines and packaging code.
- Represented the company in attending the EARL and Predictive Analytics Innovation Summit conferences in 2017.
- Knowledge sharing sessions plus code and visualisation challenges have encouraged engaging discussion and self-improvement. For my recent personal development, I completed *Python Programming: A Concise Introduction* on Coursera in 2017, and am currently working through *Machine Learning Engineer* Nanodegree on Udacity.
- Promoted internally within the team. Given more responsibility in supporting the Junior members and more trust in designing solutions and planning for new projects.

EDUCATION

- 2012 – 2013 **MSc in Statistics (Distinction).** University of Warwick
Dissertation title: *Bayesian modelling of Cortical Excitability and Sleep in Parkinson's Disease*. Project co-supervised by a medical doctor, and modelling results were well received.
- 2009 – 2012 **BSc (Hons.) Mathematics degree (1st class).** University of York
Project title: *Analysis of house prices in Boston, USA*.
- 2007 – 2009 **GCE Advanced Levels.** Jack Hunt School, Peterborough
Mathematics - A, Geography - A, History - B, General Studies - B, Further Mathematics (AS) - A, Religious Studies (AS) - A

SKILLS

- Advanced R programming, in Windows and Linux environments: using both Docker and Azure Clusters. Use regularly for data wrangling (data.table); database interaction (RODBC); visualisation (shiny, ggplot2, plotly); feature tests & packaging (testthat, devtools, Roxygen2); modelling and ML (caret, xgboost) and reporting (markdown, xlsx). Use Sublime and Rstudio for development
- Personal experience with Python
- Daily use of SQL with Microsoft SQL Server
- GIT management with Sourcetree and Bitbucket
- Agile ways of working with Jira, Trello and Slack
- Work areas: Gradient Boosting, Random Forests, GLM's, Sigmoid & Gompertz growth curves, constrained optimisation routines, Beta-Binomial Bayesian inference

INTERESTS

- Sport – Table Tennis and Badminton recreationally. Play football with work during the Summer. Keep fit by running and with aerobic exercises.
- In my spare time I enjoy reading fiction, board games, hiking, snowboarding, city breaks, cooking and playing the keyboard.
- Enjoy practicing data visualisations, and code challenges e.g. Euler and Regex. Participate in Stack Overflow community, answering R programming questions.