Mhairi McNeill

mhairihmcneill@gmail.com   *http://stackoverflow.com/users/5098093/*
mhairihmcneill.com   *github.com/mhairi*

Experience

Freelance Data Scientist   2016-2020

I work with a variety of clients, analysing data and building data systems. I'm a generalist who can do most projects but I specialise in text mining, visualisation, Shiny development, data science training and web scraping.

Examples of projects:

- Using computer vision to automatically detect items likely to be antiques in an online store.
- Using web scraping to gather data from multiple sources to build a dataset of healthcare worker deaths.
- Worked as the key data scientist in a successful pilot project for a waste tracking system, still involved in stage two of the project.
- Attribution modelling for an online retailer, including interactive visualization of results.
- Creating visualizations of healthcare data, to help customers make better healthcare choices.
- Analysing the accuracy of continuous electricity monitoring and building an outlier warning system.
- Using Bayesian modelling to understand the drivers of election results.
- An automatic web scraping, text mining and sentiment analysis system in Python.

Lecturer at Edinburgh Napier University and Instructor at CodeClan   2016 - 2020

I teach and create content a MSc. level course at Edinburgh Napier. I also teach and create content at a Data Science Bootcamp.

Data Scientist at deltaDNA, Edinburgh   2014 to 2016

I did statistical research and development, along with some statistical consulting for clients, at a games analytics company.

Statistician at Lifescan, Inverness   2014

I provided statistical rationale for a project that involved making major changes to the manufacturing process at Lifescan. I worked mainly with Excel and Minitab.

Skills

- Advanced R programmer. Package specialties include: ggplot2, dplyr, knitr, shiny, tidytext.
- Strong knowledge of machine learning and computer vision systems and practical experience of applying them.
- Wide statistics knowledge, with a particular emphasis on Bayesian analysis.
- Interactive visualizations using R Shiny.
- Python programming including use of Keras, TensorFlow, NLTK, Pandas, matplotlib and scikit-learn.
- SQL and databases: HP Vertica and PostgreSQL.
- Linux systems and experience with AWS.

Education

MSc. Applied Statistics and Datamining at   **2013 to 2014**
St. Andrews University

18/20 - Distinction

A masters focused on applications of statistics and machine learning. My dissertation involved creating an R package which implemented Riemann Manifold Monte Carlo for Generalised Linear Models.

Statistics courses from The Open University   **2011 to 2013**

Practical Modern Statistics - coursework: 88% final exam: 98%
Analysing Data - coursework: 89% final exam: 97%

Mathematics BSc at the University of Edinburgh   **2008 to 2011**