Mathew W. McLean

mathew.w.mclean@gmail.com
https://mwmclean.github.io/

+61 4 5238 9372
Citizenship: Canadian

Education

Ph.D. in Operations Research, Cornell University, 2013

    Concentration: Applied Probability and Statistics

    Advisors: David Ruppert and Giles Hooker

M.S. in Operations Research, Cornell University, 2013

B.Sc. (Hons.) in Statistics, University of Manitoba, 2008

Professional Experience

Software Engineer - Implementation Team, Displayr, 2021-present (Sydney, Australia)

Developed and tested features (mostly written in R and JavaScript) for our market research software products, `Q` and `Displayr`. Since joining, the company has grown from 20 to 50+ employees and nearly ten million in ARR.

Maintained/co-authored a suite of ~25 R packages for statistical modeling, visualization, data cleaning and integrating with various web APIs.

Implemented new features for our apps, mostly using JavaScript and R, to simplify and automate data cleaning, analysis, and reporting for our users.

Experience with R, JavaScript, GitHub Actions, htmlwidgets, stan, Emacs, travis-CI, docker, Scrum, RStudio, tidyverse.

Team Lead - Data Science Team, Displayr, 2019-2021

Lead a small team of data scientists to write efficient, intuitive tools for performing text analysis, driver analysis, regression, and choice modeling in our software.

Our team wrote and maintained 1,000s of integration and unit tests for all data science features in our products.

Computational Statistician - Data Science Team, Displayr, 2017-2019

Designed user interface as well as interactive visualization and HTML widget output for data science and machine learning features in our apps.

Developed models implemented in *stan* and *R* for Hierarchical Bayes choice modeling, simulation, and experimental design.

Wrote blog posts detailing usage of new features and replied to customer questions.

Postdoctoral Research Fellow, University of Technology Sydney, 2015-2017

Supervisor: Matt P. Wand

Developed novel, fast, approximate algorithms for fitting Bayesian semiparametric and nonparametric models; applied methods to streaming data from robotics/computer vision

Maintained website for the online fitting of semiparametric regression models using data scraped in real-time from various resources

Research Assistant Professor, Texas A&M University, 2013-2015

Supervisor: Raymond J. Carroll

Developed new semiparametric methods to account for measurement error when analyzing dietary intake data; estimation performed using Markov Chain Monte Carlo implemented in R and C

Used generalized linear models to analyze count data from a large RNA-Seq experiment to study the effects of diet on colon cancer

Invited talks on Bayesian modelling for longitudinal data, hypothesis testing with mixed models, reproducible research, and authoring `R` software packages

Four submitted or published peer-reviewed journal articles

Sole author and maintainer for `RefManageR`, an R package available on CRAN. The package allows for importing bibliographic references via web APIs, PDFs, and `.bib` files into `R` and provides tools for easily manipulating, printing, and searching the stored references

Produced reproducible research manuscripts and presentations in `Markdown`, LATEX, and `HTML5` formats with `knitr`/`Sweave` and interactive web applications using RStudio's Shiny framework

Reviewed original applied statistics research submitted for publication for ten statistics journals

Research Assistant and Teaching Assistant, Cornell University, 2008-2013

Research in functional data analysis, semiparametric regression, variable selection for high dimensional models, Bayesian hierarchical modelling, mixed linear models, and time series analysis

Three papers published in statistics journals with code contributed to `R` package `refund`

Five invited talks and poster presentations on regression models for functional data

Lectured to classes of approximately 30 students on introductory statistics for engineers, stochastic processes, and data mining; as well as ran labs demonstrating use of `Excel` and `R` for statistics

Computing Skills

R, Linux, Emacs/ESS, Git, stan, RStudio, MATLAB, SAS, C, Java, MS Office, HTML

Publications, talks, and software are available at http://mwmclean.github.io/