

WORK EXPERIENCE

Senior Data Scientist Decision Technology

10/2014 – Present London

Consulting company specialising in understanding and predicting customer behaviours with experimental psychology, cognitive science and statistics

Achievements/Tasks

- Lifted customer conversion rate by 17% by building statistical models and mining clients' big data sets to identify enablers and blockers for product consideration
- Accurately forecast sales for clients by deploying ARIMA and neural network based time series models accounting for marketing and promotion activities
- Boosted customer lifetime value for clients by optimising product offerings and improving experience through A/B tests and randomised control trials
- Slashed data model deployment time from days to hours by implementing a DevOps framework leveraging Docker and container orchestration
- Designed and launched multiple research to predict the 'new normal' to help clients with post-lockdown strategic business planning
- Effectively managed a team of junior data scientists and worked closely with other business functions on multiple concurrent projects
- Translated data to actionable insights and recommendations and communicated results to (non)technical stakeholders across different functions

EDUCATION

MSc Advanced Computing Technologies (Machine Learning)

Birkbeck, University of London

09/2016 – 09/2018 Distinction

Selected Courses

- Big Data Analytics
- Machine Learning
- Parallel Computing and Distributed Systems
- Data Mining & Warehousing

MSc Sociology

University of Oxford

09/2013 – 09/2014 Distinction

Selected Courses

- Statistics
- Experimental Design

BA Linguistics and Sociology

University of Manchester

09/2011 – 09/2013 First

SKILLS

Machine Learning

Data Mining

Time Series Analytics & Forecasting

Neural Networks

Statistical Modelling

Predictive Modelling

Survival Analysis

Regression

Dimension Reduction

Clustering

A/B Testing

Multivariate Testing

Hypothesis Testing

Experimental Design

Natural Language Processing

Data Visualisation

R

R Shiny

Python

TensorFlow

SQL

Tableau

Hive

Spark

AWS

Docker

Kubernetes

PROJECTS

Data Science blog: databentobox.com

(10/2016 – Present) 📄

- databentobox.com is a blog where I showcase my personal data science projects and share thoughts on statistics and programming
- Blog articles are frequently re-posted by R-bloggers and KDnuggets and have gained traffic in the data science community
- Specifically, the recent project Simulating Coronavirus Outbreak with Origin-Destination Matrix and SEIR Model attracted 20k views

MSc Machine Learning Project: Tackling Toxic Comments Using Convolution and Recurrent Neural Networks (06/2018 – 09/2018)

- Going beyond standard sentiment analysis, the goal of this project is to build an automated tool to identify online toxic comments
- Neural network based classification models (bidirectional GRU-CNN) with pre-trained word embedding were built
- The final model achieved state-of-the-art classification results and the MSc project was awarded distinction

MSc Sociology Thesis: Gender Equity and Women's Fertility Intention in European Countries (06/2014 – 09/2014)

- This thesis provides an empirical test of gendered fertility theory by examining whether low gender equity at home lowers women's fertility intentions in 31 European countries
- Using multilevel logistic regressions, together with the data from the second and the fifth round of European Social Survey, the MSc thesis generated unique insights and was awarded distinction

LANGUAGES

English

Native or Bilingual Proficiency

Mandarin Chinese

Native or Bilingual Proficiency