

# Yihui Fan

## Data Scientist | Business Analytics & Consumer Behaviours

Postgraduate training in both Machine Learning and Social Science, together with 6 years of work experience in data science, behavioural science and management consulting equipped me with a unique skill set for delivering data-driven business insights.

pressto.fanyihui@gmail.com

+44(0)7738 630073

London, United Kingdom

databentobox.com

linkedin.com/in/yihui-fan-datascience

## 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 mining clients' big data sets and building statistical models 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
- Automating the process and classification of textual data by training and deploying deep learning models, reducing operational cost for clients
- Slashed data model deployment time from days to hours by implementing a DevOps framework leveraging Docker and container orchestration
- Designed and built dashboards that track clients' key metrics during COVID-19 and facilitate 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 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

## LANGUAGES

English

Native or Bilingual Proficiency

Mandarin Chinese

Native or Bilingual Proficiency