

Yihui Fan

Data Scientist | Consumer Behaviour & Insights

Postgraduate training in both machine learning and social science, together with consulting experience in public and private sectors equipped me with a unique skill set for solving business problems, generating data-driven insights, and developing data products.

pressto.fanyihui@gmail.com ✉

+44(0)7738 630073 📞

London, United Kingdom 📍

databentobox.com 🌐

linkedin.com/in/yihui-fan-datascience in

WORK EXPERIENCE

Senior Data Scientist Decision Technology

10/2014 – Present

London

Consulting company specialising in experimental psychology, cognitive science and statistics

Achievements/Tasks

- Apply advanced statistical models and conduct online experiments in diverse domains such as marketing and media strategies, proposition and price optimisation, customer intelligence, sales forecasting, etc.
- Lead the firm's R&D team to develop analytical models, automated data pipeline and visualisation services to facilitate data-driven decision making.
- Work with client-side product, marketing and commercial teams to evaluate and improve propositions and marketing effectiveness.
- Experienced in translating data to recommendations and communicating results to (non)technical stakeholders across different functions.

EDUCATION

MSc Advanced Computing Technologies (Machine Learning)

Birkbeck, University of London

09/2016 – 09/2018

Distinction

Courses

- Big Data Analytics using R
- Data Warehousing and Data Mining
- Cloud Computing
- Information Security
- Intelligent Technologies (Machine Learning)
- Advanced Database Management

MSc Sociology University of Oxford

09/2013 – 09/2014

Distinction

Courses

- Social Statistics
- Social Network Analysis
- Quantitative Method and Survey Research
- Experimental Social Science
- Research Design
- Qualitative Research Methods

BA Linguistics and Sociology University of Manchester

09/2011 – 09/2013

First

SKILLS

Statistical Modelling

Survival Analysis

Regression

Random Forests

Dimension Reduction

Clustering

Time Series Analytics & Forecasting

A/B Testing

Multivariate Testing

Experiment Design

Survey Design

Lifetime Value Modelling

R

R Shiny

Python

TensorFlow

SQL

MongoDB

Hive

Spark

AWS

Docker

PROJECTS

Data Science blog: databentobox.com

(10/2016 – Present) 🌐

- databentobox.com is a blog where I showcase my personal data science projects and share my thoughts on statistics and programming.
- Blog articles are frequently re-posted by R-bloggers and KDnuggets and 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 a series of 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