

EDUCATION

University of Connecticut

Ph. D. in Statistics GPA: 3.91/4.0

08/2021 - Present

Storrs | CT | USA

Courses: *Applied Statistics, Design of Experiments, Mathematical Statistics, Linear Models, Advanced Probability, Statistical Inference*

University College London, UCL

09/2019 - 09/2020

MSc in Data Science WES GPA: 3.75/4.0 (Distinction)

London | UK

Courses: *Information Retrieval and Data Mining, Machine Learning, Applied Machine Learning, Statistical Data Science, R for Statistics*

Heriot-Watt University, 2-year Exchange Program with SWUFE

09/2017 - 05/2019

BSc in Actuarial Science WES GPA: 4.0/4.0 (First-Class Honors)

Edinburgh | UK

Awards: Deputy Principal's Award (07/2018 and 07/2019)

Courses: *Discrete-time Financial, Continuous-time Finance, Time Series, Stochastic Processes, Statistical Machine Learning, Statistical Models*

Southwestern University of Finance and Economics, SWUFE

09/2015 - 07/2017

Bachelor of Economics in Insurance (Actuarial Science - Bilingual) GPA: 84/100

Chengdu | China

Honor: Third-Class Scholarship (Top16%, 12/2017)

Courses: *Risk Management, Financial Mathematics, Probability Theory, Excel in Economic Management*

RESEARCH

On Joint Modeling of Mean-Variance-Correlation of Clustered Data: Fallacies and Corrections, UConn 05/2022 - Present

- Identified the correct one among two different “sandwich” variance estimators
- Conducted several simulation studies to investigate under what situation the two estimators are similar and under what situation they are not
- Developed a new criterion for joint model selection

NIR Spectroscopic Data Classification by Ensemble Methods, Dissertation, UCL

06/2020 - 09/2020

- Used PCA to reduce dimensions to 7 to maintain 95% information of high-dimensional spectral data
- Tried two ensembles (random forest and XGBoosting) to classify the data
- Used a double cross-validation scheme to evaluate each model's performance

Information Retrieval Model Project, Information Retrieval and Data Mining, UCL

01/2020 - 04/2020

- Built a passage re-ranking system of a given candidate list of passage to a query using retrieval models
- Built an inverted-index for the collection for efficient passage retrieval, tried retrieval models such as BM25 model, logistic regression, LambdaMART Model and neural network, with Word2Vec embedding

COMPETITIONS

Otto Group Product Classification, Kaggle Competition

01/2020 - 03/2020

- Conducted Otto Group product data visualization and preprocessing
- Trained classification model such as random forests, neural networks, XGBoosting, to predict product classification

INTERNSHIPS

DataGrasp

10/2019 - 12/2019

Strategic Analyst, Medicspot Team, 10hrs/week

London | UK

- Processed Medicspot's data (pharmacy addresses and installed device) to construct a linear regression model
- Calculated and ranked the expected revenue of installing Medicspot device in pharmacies across the UK

China Life Insurance (Group) Company

07/2018 - 08/2018

Data Analysis Assistant, Health Insurance, 40hrs/week

Chengdu | China

- Extracted long-term care insurance compensation data of Chengdu of the past five years by SQL
- Summarized and analyzed the claims data by R and estimated the future compensation amount
- Understood Markov model and fitting application regression statistical model in the life insurance industry

POSTER PRESENTATIONS

- Variance estimation for generalized estimation equations of mean-variance-correlation for clustered data, **Excellence in Statistical Science**, Storrs, CT, 2022
- Disparity in county-level low-income job loss rate during the Covid-19 pandemic, **The Joint Statistical Meetings**, Washington, DC, 2022

ADDITIONAL INFO

Skills: Master of R, Python, SAS; Proficient in MATLAB, SQL,