Yunjie He （Roya）

Phone: +86 15228188888 Email: zcakebx@ucl.ac.uk

EDUCATION

MSc Computational Statistics and Machine Learning, University College London Sep 2020 - Sep 2021

Department of Computer Science London

- Grade: Distinction
- Core Modules: Supervised Learning, Probabilistic and Unsupervised Learning, Approximate Inference, Statistical Models and Data Analysis, Statistical NLP, Applied Machine Learning, Reinforcement Learning, Machine Learning Seminar.
- In-course NLP project: Discovering the Effectiveness of Pre-trained Masked Language Model for English Spelling Correction
- MSc project supervised Dr. Philip Gorinski: Improving Hierarchical Graph Neural Networks for Multi-hop Question Answering Via Hierarchical Graph Attention Mechanism and Graph Completion

BSc Economics and Statistics, University College London Sep 2017 - Jun 2020

Department of Statistical Science London

- Grade: First Class Honours
- Related modules: Advanced Linear Algebra, Calculus, R, Python, Stochastic System, Statistical Inference, Probability and Statistics

RESEARCH EXPERIENCE

HUAWEI Noah's Ark Lab - NLP group Jun 2021 - Present

Research project supervised by Dr. Philip Gorinski and Dr. Pontus Stenetorp London

- Researched into the Multi-hop machine reading comprehension(QA) task, Graph Neural Networks and their related applications in downstream tasks, especially the question answering problem.
- Presented two novel extensions on the existing HGN model to improve its performance on Multi-hop QA task. Firstly, we reconstructed the graph structure in the HGN model by introducing question2sentence edges and leveraged the influence of graph structure on the model performance. Secondly, we proposed a hierarchical graph node update mechanism for the graph attention network (GAT) in the Multi-hop Question Answering task based on the HGN (Hierarchical Graph Network) model.

In-course NLP research project on spell correction Feb 2021 - May 2021

Topic: Discovering the Effectiveness of Pre-trained Masked Language Model for English Spelling Correction

- Examined BERT’s suitability in spell correction, then conducted different empirical experiments to improve BERT’s performance.
- Based on experiment results, we propose a novel model CLMBER (Char-CNN-LSTM-Multilingual-BERT model) which achieves the best performance in the spell correction experiments.

Alan Turing Institute Jun 2019 - Sep 2019

Research project supervised by Dr. Franz Kiraly London

- Engaged in the Machine Learning Toolbox Design project aimed to build Object-Oriented Programming based interface framework for statistical distributions and machine learning algorithms applications in R.
- Contributed to the design of the mlrpro and Distr6 toolbox based on knowledge of mlr3, skpro, and mathematical theory including Bayesian modelling in mlrpro and MCMC in mlrpro, BUGS, JAGS, Stan

Kaggle Competition in ASHRAE Energy Predict | Ranking:top 1%

- Applied machine learning and deep learning to fraud detection & energy cost problems in Python
- Completed the individual part of EDA; polished feature engineering part; employed machine learning models for data prediction; and used Bayesian optimization for parameter tuning

PROFESSIONAL EXPERIENCE

HUATAI Technology co., ltd Jun 2020 - Aug 2020

Big Data Analyst Intern

- HUATAI, a fast-growing data tech startup, commits to provide data-driven professional business suggestions to clients. I worked as part of the big data team to perform exploratory and statistical analysis to reveal trends, understand user behaviors and draw insightful conclusions on the performance of products with the help of big data analytical tools, such as Hadoop, Hive, and Python

Rshiny app UCL, London

developer&maintainer

- Optimized space utilization of the Statistical Science Department at UCL by creating a web-interactive Rshiny app for office use information. Realized a connection between online SQL database and Rshiny app and made use of the collected information for statistical analysis using R.

SKILLS LIST

Python, R, Pytorch, Git, Linux, SQL, Latex