Yunjie He （Roya）

+49 152 05313060 | yunjie.he@ipvs.uni-stuttgart.de
https://royahe.github.io

EDUCATION

Ph.D. University of Stuttgart IMPRS-IS & Bosch Center for Artificial Intelligence (BCAI) Jun 2022
Department of Computer Science Stuttgart
Research topic: neural-symbolic methods for knowledge graph logical query answering, knowledge representation, and reasoning.

MSc Computational Statistics and Machine Learning, University College London Sep 2020 - Sep 2021
Department of Computer Science London

- Grade: Distinction (82.0/100.0)
- 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, MSc Project thesis.

BSc Economics and Statistics, University College London Sep 2017 - Jun 2020
Department of Statistical Science London

- Grade: First Class Honours (78.0/100.0)
- Core Modules: Advanced Linear Algebra, Calculus, R, Python, Stochastic System, Statistical Inference, Probability and Statistics

PUBLICATIONS

Modeling Relational Patterns for Logical Query Answering over Knowledge Graphs
Authors: Yunjie He, Mojtaba Nayyeri, Bo Xiong, Evgeny Kharlamov, Steffen Staab

Graph Attention With Hierarchies for Multi-hop Question Answering
Authors: Yunjie He, Philip John Gorinski, Ieva Staliunaite, Pontus Stenetorp

Distr6 Package in R
An R6 object-oriented distributions package. Unified interface for probability distributions and kernels. Link

RESEARCH EXPERIENCE

HUAWEI Noah's Ark Lab - NLP group Jun 2021 - Oct 2022
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 I 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

SKILLS LIST

Python, Java, R, Pytorch, Github, Linux, SQL