王稷尧 Jiyao Wang

www.wangjiyao.top | jiyaowang130@gmail.com

Research Interest: Data Mining & Natural Language Processing

SUMMARY

Computer science student and data science enthusiast, dedicated to finding treasure from data. Experienced in data development, employing algorithms to achieve system optimization or resolving real-world issue, and delivering results on tight conditions. Zealous volunteer passionate about public benefit activities.

EDUCATION

Hong Kong University of Science and Technology 2022.09 - 2025.07

Ph.D in Intelligent Transportation

Accepted offer.

Hong Kong University of Science and Technology 2021.09 - 2022.06

M.S in Big Data Technology

Core Coursework: Big Data Computing, Data Visualization, Parallel Programming, Machine Learning, Data Mining and Knowledge Discovery, etc.

Sichuan University 2017.08 - 2021.05

B.E in Software Engineering

Core Coursework: Object-Oriented Programming, Data Structure and Algorithmn, Computer Network, etc.

PUBLICATIONS

[1] Preciser Comparison: Augmented Multi-layer Dynamic Contrastive Strategy for Question Classification. 2022 Conference on Neural Information Processing Systems. First author in progress

[2] Multi-Aspect co-Attentional Collaborative Filtering for Extreme Multi-label Text Classification. 2022 International Conference on Computational Linguistics. First author in progress

[3] An Multi-Aspect Attentional Model To Capture Multistratal Influence In Social Group, 2021 IEEE International Conference on Electronic Information Engineering and Computer Science. First author published

[4] Adversarial Social Recommendation With Capturing Multi-Modal Views Of Social Friends, 5th International Conference on Electronic Information Technology and Computer Engineering. Third author published

RESEARCH EXPERIENCE

Research Assistant in HKUST Transportation Behavioural Psychology and Safety Lab 2022.03 - Present

- Under the supervision of Prof. Dengbo He, I lead a project on modeling and evaluating drivers' trust in electric vehicle battery systems.
- The main areas and methods covered include ergonomics and applied statistics.

Research Assistant in HKUST Database Reseach Group 2021.10 - 2022.02

- This group is led by Prof.Lei Chen in HKUST. The main content is participating AI/DM competitions cooperating with a senior PhD. In both an AutoML competition held by Tencent and ATEC2021 which is related to identifying Internet fraud reports, my team's solution got top20% performance.
- Currently, I am participating in optimizing one of our lab's models in the OGB(Open Graph Benchmark).

PROFESSIONAL EXPERIENCE

Meituan 2021.06 - 2021.09

Autonomous Vehicle Algorithm Intern Beijing

- Utilized C++ to design and develop obstacle decision models, hoping to improve the decision jitter in the face of various types of obstacles
- Applied the clustering method to analyze and simulate the reasons for the jitter in the decision of autonomous vehicles when facing obstacles in the real environment
- Employed Hive SQL and PySpark to complete the evaluation index construction of the obstacle decision model, designed and built the required database and the optimal decision validation set

DiDi 2021.03 - 2021.06

Data Operations Intern Chengdu

- Using Hive SQL to complete the design and construction of related relational databases, and performing required data extraction; applying machine learning and statistical knowledge to clean up and process the sales data to provide data visualization and conclusions
- utilizing Python to analyze the user behavior data of the platform, and participating in the completion of the platform user hierarchy analysis report

PROJECT EXPERIENCE

ICM2021 Big Data Analysis of Music Genre Evolution 2021.10 - 2021.12

This project is based on the 2021 American College Students Mathematics Modeling Competition, using Hive+PySpark, deployed on 3 Linux servers, and completed the construction of data warehouse, data cleaning and basic analysis of style and genre evolution. I am mainly responsible for big data environment configuration, Hive database design, data extraction and analysis by tools in PySpark

KDD Cup 2021 Time Series Anomaly Detection 2021.10 - 2021.12

The project is provided on KDD Cup 2021, and the goal is to build a model that can automatically detect anomalies in time series through unsupervised learning methods. I mainly completed the Matrix Profile time series feature extraction, the Forecast-based LSTM time series anomaly detection sub-model, and the weighted ensemble model. The final Accuracy is around 71

Sichuan Provincial Statistics Bureau's Innovative Tech Platform 2020.09 - 2020.11

Took charge of platform's construction using PyQt and MySQL for the calculation and display of innovative economic indicators of Sichuan province. Deployed a downloadable Windows APP on cloud servers and obtained the software copyright

E-commerce Platform Portal Construction Based on JAVA EE 2019.07 - 2019.08

Utilized Java EE and database knowledge to complete the construction of an e-commerce online shopping platform website within two weeks. Completed design and construction of the database and some functional modules, and arranged and coordinated work within the group

HONORS & AWARDS

2020 Merit Student, Sichuan University 2020.09

First-class Scholarship, Sichuan University 2020.09

Successful Participant, 2020 MCM 2020.05

Second-class Scholarship, Sichuan University 2019.09