Yuanyuan Zhang

zhang.huanzhiyuan@gmail.com · (+86) 136-933-89103 · https://hengqujushi.github.io/ ·

Summary

Yuanyuan Zhang received a B.S. degree from Jilin University in 2012 and has done pioneering work in the industry in many leading companies in the track, such as Baidu, Ledongli, Alibaba, and People's Car Service. He has accumulated more than 10 years of work in data science and data platforms in mobility data, trajectory data, and other fields.

At Ledongli:

- Independently responsible for the data science work of the first annual selected app in the Apple Appstore in China.

At Alibaba:

- Developed the AI sports product, the industry's first application to count fitness movements in real-time on a mobile phone.
- During the epidemic, the product helped hundreds of thousands of college teachers and students successfully carry out physical education teaching and was promoted to tens of millions of primary and secondary students.
- Responsible for the commercial algorithm work of walking and running, with an annual income of nearly 100 million yuan.

During the People's Car Service Period:

- Led the team in developing an industry-leading SDK and UBI solutions for dangerous driving behavior identification.
- Applied for 7 patents and published 10 top-tier conference papers, including 8 CCF-A papers.
- The project became the first in the insurance industry to be selected for the supervision sandbox of the People's Bank of China for financial technology innovation applications.
- Led the development of the industry's first scalable, high-performance, cloud-native, one-stop geospatial data platform, aimed at addressing the issue of missing data/AI infrastructure in the field of geospatial data.
- Currently leading the team in developing a full-process agent system for vehicle insurance based on Large Language Models (LLMs).

I am deeply passionate about language models (LMs) that are designed to meet a broad spectrum of human needs and center around human-centric concerns. My focus revolves around several critical questions:

- **How can we develop robust and scalable methods to evaluate LMs, particularly for tasks where human performance is challenging?**

- **What approaches can we employ to create LMs that encapsulate the full spectrum of human desires, including ethical values, personal preferences, and specific competencies?**

- **Is it possible to tackle complex issues by enhancing the interactive capabilities of LMs? This enhancement could include augmenting tools and memory, facilitating multi-agent communication, and fostering effective human-AI collaboration.**

I am convinced that these areas of evaluation, alignment, and interaction are interconnected, forming a synergistic loop where each element influences and bolsters the others. My aspiration is to delve into each of these realms not only as isolated topics but also in terms of their interplay, with the ultimate goal of crafting general-purpose language models that are truly human-centric.

Campus Experience

Renmin University of China, Beijing 2020 – 2025

Master Tutor (Industry) Department of Applied Statistics, School of Statistics

- Give undergraduate students the course "Data Science Practice" and teach the part of graph neural network, which was highly praised.
- Tutor many students in scientific research activities, and has published 8 CCF-A papers, and three papers are being submitted.
- Guided students in participating in the NeurIPS Large Language Model Efficiency Challenge, achieving a top ten placement.

Jilin University, Changchun,Jilin 2008 – 2012

Bachelor Information and Computing Science

- In the first three academic years, the professional course score is in the top 10% of the major
- Main Courses: Calculus, Linear algebra, Probability and Statistics, Real Analysis, Numerical analysis, Partial Differential Equations, Information Theory, Data Structures and Algorithms.

Work Experience

Beijing Baixingkefu Network Technology Co., Ltd. Director of Data Science and Data Platform Department 09.2020-Present

- The industry-leading dangerous driving behavior recognition system keeps ahead of competitors in many key links such as event recognition, scene recognition, driver and passenger identification
- The industry's first scalable, high-performance, cloud-native, one-stop geospatial data platform
- Developing a full-process agent system for vehicle insurance based on Large Language Models (LLMs).

Alibaba Staff Engineer 09.2017-09.2020

- Commercialization of mass sports data such as walking and running
- The electronic coupon allocation based on purchase intention estimation and MCKP, which also inspired a work of 2020 CIKM
- The industry's first mobile fitness action real-time counting and timing system

Ledongli Co. LTD Data scientist 05.2014-09.2017

- The industry's first mobile pedometer that can run continuously in the background using IMU and Magnetometer
- Automated human activity recognition system based on motion sensor
- Mobile automatic wake-up system with power saving and timely
- Forecast of sales lead conversion possibility

Baidu Internship software engineer/Software Engineer 08.2011-05.2014

- Web Traffic Forecasting using ARIMA
- Session-aware document recommendation
- Session-aware Personalized music recommendation

Publication

- Zhang, Y., Du, Y., Zhang, Y., and Zhang, Q. "Moral Hazard and Transparency in Peer-to-Peer Auto Insurance with Telematics." To appear in ICIS 2023 Proceedings.
- Yuanyuan Zhang; Bo Peng; Yalei Du; Jinhua Su, GeoLake: Bringing Geospatial Support to Lakehouses, IEEE Access 2023
- Zehong Zeng; Yueyang Liu; Xiaoshi Lu; Yuanyuan Zhang; Xiaoling Lu, An Ensemble Framework Based on Fine Multi-Window Feature Engineering and Overfitting Prevention for Transportation Mode Recognition, UbiComp '23
- Shiyao Huang; Junliang Lyu; Sinian Zhang; Ruiying Tang; Huan Xiao; Yuanyuan Zhang; Xiaoling Lu, A Post-processing Machine Learning for Activity Recognition Challenge with OpenStreetMap Data, UbiComp '23
- Jiebi Deng; Jingqiu Xu; Zicheng Sun; Danning Li; Hongxuan Guo; Yuanyuan Zhang; Xiaoling Lu, Enhancing Locomotion Recognition with Specialized Features and Map Information via XGBoost, UbiComp '23

- Mengyuan Li; Jun Zhu; Yuanyuan Zhang; Xiaoling Lu, Enhanced SHL Recognition Using Machine Learning and Deep Learning Models with Multi-source Data, UbiComp '23
- Yaya Zhao; Lin Song; Cheng Ni; Yuanyuan Zhang; Xiaoling Lu, Road Network Enhanced Transportation Mode Recognition with an Ensemble Machine Learning Model, UbiComp '23
- Hanchao Yan; Xinran Huang; Yiling Ma; Ruizhe Yao; Zhiyu Zhu; Yuanyuan Zhang; Xiaoling Lu, AttenDenseNet for the Sussex-Huawei Locomotion-Transportation (SHL) Recognition Challenge, UbiComp '23
- J Su, Y Zhang, Triple-O for SHL Recognition Challenge: An Ensemble Framework for Multi-class Imbalance and Training-testing Distribution Inconsistency by OvO Binarization with Confidence Weight of One-class Classification, UbiComp '21, September 2021, Pages 401–407
- Y Duan, Y Zhang, C Gao, M Tong, Y Zhang, K Bian, W Yan, Trajectory-matching prediction for friend recommendation in anonymous social networks, GLOBECOM 2017-2017 IEEE Global Communications Conference, 1-6

♡ Award

National Encouragement Scholarship	08-09
First-Class Scholarship, Jilin University	09-10
National Encouragement Scholarship	10-11
Meritorious Winner in the Mathematical Contest In Modeling	10.02
Top 10 in Computational Mathematics in Peking University's Direct Doctoral Mathematics Examination	11.05
Both teams I coached ranked in the top 10 at the Sussex-Huawei Locomotion Challenge	21.10
All six teams qualified for the Sussex-Huawei Locomotion Challenge finals, with four placing in the top 10	23.10
Top 10 in the NeurIPS Large Language Model Efficiency Challenge	23.12

i Social Experience

- Permanent member of the Council of the Capital of Statistics, the largest data science community in China.
- Invited to Guanghua School of Management, Peking University, Xi'an Jiaotong University, Jilin University, and Renmin University many times to share the practice of data science in the industry.
- Participated in the translation and proofreading of many classic books, including the classic book Deep Learning in the DL field.