

WANG ZIYUAN

Address: 777 Guoding Road, Yangpu District, Shanghai 200433

Tel: (+86) 188-1787-6918; **E-Mail:** wangziyuan@163.sufe.edu.cn

Date of Birth: April 6th, 1996; **Date of Graduation:** June, 2024; **Gender:** Male; **Nationality:** China

Research Area: Natural Language Processing, Deep learning, Quantitative Analysis/, Design Science

My Website: [HomePage](#); [inLinkedIn](#); [GoogleScholar\(100+Citation\)](#); [Github\(1K+Star\)](#); [CSDN](#)

Education

- **Tippie College of Business, University of Iowa, Iowa**
2022.4-2023.4, Visiting Scholar, majoring in Business Analytics.
Supervisor: [Professor Weiguo Patrick Fan](#)
Research Area: Deep Learning, Semantic Textual Similarity, Design Science.
- **School of Information Management and Engineering, Shanghai University of Finance and Economics, Shanghai**
2018.9-now, PhD Candidate, majoring in Management Science and Engineering.
GPA: 3.6/4.0 **Rank:** Top10%
Supervisor: [Professor Hailiang Huang](#), [Associate Professor Songqiao Han](#)
Research Area: Text Classification, Semantic Textual Similarity, Quantitative Trading, Design Science.
Awards: First Prize National Scholarship for Graduates, Second Prize National Scholarship for Graduates, Outstanding Class leaders of School, Outstanding Students of School
- **College of Finance and Statistics, Hunan University, Changsha**
2014.9-2018.6, Bachelor of Science in Statistics.
GPA: 3.8/4.5 **Rank:** Top5%
Awards: National Scholarship, Outstanding Graduate of Hunan Province, Citibank "Future Elite" Scholarship, First-Class Scholarship for undergraduates, School of Merit student, Outstanding Class Leader, First Prize in the "Agricultural Development Bank Cup" competition, H Award in "Certification Cup" International Mathematical Modeling Competition.

Publications & Research Project

How Close is ChatGPT to Human Experts? Comparison Corpus, Evaluation, and Detection 2022.12-2023.07
Under Review.

- Co-first author. We initiated the ChatGPT Comparison and Detection Project, which collected human-ChatGPT comparison data and conducted statistical and linguistic analysis. Based on deep learning and machine learning, we developed a series of ChatGPT content detectors.
- We were the earliest team in academia and industry to open-source comparison datasets and detection models. The detector demo has been accessed over 200,000 times worldwide, and the open-source model has an average monthly download volume of over 30,000, with the dataset having an average monthly download volume of over 30,000. The project has received over 1000 Github stars, with paper citation over 100. Recently, our work has been accepted by [LLM@IJCAI](#)(CCF-A, CORE-A*). Please feel free to visit our [Demo](#) and [Paper](#).

Investigating Effectiveness of Whitening Post-processing on Modifying LLMs 2022.09-2023.06
Accepted by 35th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2023, CCF-C, CORE-B)

- First author. We explored different whitening methods for matrices to normalize the representation of the BERT model to reduce correlation, aiming to solve the assumption that traditional Bert semantic representation ignores orthogonality of cosine similarity basis, and evaluated performance on various NLP tasks.
- Supported by the National Natural Science Foundation of China (Project No. 72271151) and the Graduate Innovation Fund of Shanghai University of Finance and Economics (Project No. CXJJ-2021-052).

New Hybrid Model of Crowdfunding Project: A Perspective of Prospect Backers 2022.08-2023.02
Under Review.

- Corresponding author. This work captures the heterogeneity of investor support behavior in crowdfunding by introducing the prospect theory (PT), evaluates the prospect utility of the project, and makes support decisions based on the evaluated prospect utility.

IDEA: Interactive Double Attentions from Label Embedding for Text Classification 2022.06-2022.09

Published by 34th IEEE ICTAI, CCF-C, CORE-B, acceptance rate = 15.7%

- First author. We proposed a simple method to improve the performance of BERT in text classification by utilizing label embedding technology in supervised learning. Based on the original BERT model, we proposed a new model structure that integrates text and label information, which has achieved significant improvement effect on the open datasets. Please feel free to visit our [Paper](#).
- Supported by the National Natural Science Foundation of China (Project No. 72271151).

Web Tool: SUFE-CS-CONF-DDL 2022.01-2022.03

- Project initiator/project leader. We built a countdown system tool for the computer conference of Shanghai University of Finance and Economics based on Vue-CLI, providing dual retrieval of tenure track tier/CCF level.
- Please feel free to visit our visualization website [SUFE-CS-CONF-DDL](#), and the open-source code of the project is available on [Github](#).

Analysis of Shanghai's Biopharmaceutical Industry Chain Based on Big Data and Comparative Study of the Yangtze River Delta (with Suzhou Bank) 2021.02-2022.01

- Algorithm Engineer. Combining big data intelligent industrial research technology with traditional industrial economics, we constructed a knowledge graph of the biopharmaceutical industry chain and conducted an analysis of the biopharmaceutical industry chain in the Yangtze River Delta.
- In May 2021, the leaders of the Shanghai Municipal Party Committee visited an economic regulatory platform in a certain district and fully recognized the achievements of the platform construction, pointing out that further efforts should be made to build a "city brain" upgraded version in Shanghai.

Policy and Business Classification in Technology Public Opinion Recommendation Systems (with Shanghai Flaginfo Information Technology & PingAn Technology) 2020.08-2021.01

- Project leader of text classification. We crawled and cleaned policy texts released by the government, and conducted multi-label classification based on the business types of PingAn Technology's business group, with traditional and SOTA ML/DL text classification methods. Partial code: [Github](#).

The Dilemma of Charity—How to Save Crowdfunding Goodwill under Social Media. 2017.02-2017.05

- Project leader. I participated in the questionnaire writing, data collection and modeling, and paper writing throughout the 2017 National College Student Market Survey and Analysis Competition, and obtained the National Market Survey and Analysis Skills Certificate (CRA).
- Our team won the national second prize and the provincial first prize, and was the only team from Hunan Province to advance to the finals.

Working Experience

Millennium Management, CRTIC 2023.7-now

- **Junior Quantitative Researcher Internship**
As a Quantitative Analyst, my responsibilities included collecting data and constructing a framework for market sentiment analysis. I utilized data mining techniques to research, analyze, and optimize model outcomes. I also conceptualized, examined, and predicted the performance of alternative alpha signals. And leveraged SOTA NLP models to develop deep learning time-series models. Ultimately, my role involved establishing investment portfolios, informed by an extensive review of the latest academic and industry research findings.

On-campus employment

- **Academic Reviewer:** Applied Intelligence, International Journal of Computational Science and Engineering, ACM International Conference on Information and Knowledge Management.
- **Teach Assistant:** Advanced Machine Learning (Fall 2019), Application of Big Data in Economics (Fall 2021).
- **Research Assistant:** Working as RA in Fall 2020, Fall 2021, and Spring 2022 semesters.

Skills

- **IT:** Python (Pytorch, Tensorflow, Keras), R, SPSS, MATLAB, MySQL, Vue-CLI; LaTeX; Microsoft Office
- **ACCA:** Association of Chartered Certified Accountants, F1-3
- **Language:** IELTS: 7.0(R/L/S/W:7.5/7.5/6.5/6.0); Mandarin Certificate (Level 1 B)
- **Others:** National Invention Patent, C-1 Driver License, Saxophone Level7