ZHANLI LI

(+86) 156-2350-1965 ◇ lizhanli@stu.zuel.edu.cn ◇ HomePage ◇ GitHub

EDUCATION

Zhongnan University of Economics and Law *September 2023 - June 2027*
Wenlan School of Business, B.Sc. in Digital Economy
Weighted Avg: 93.4/100 Rank: 1/80 2025 National Scholarship
Core Courses: Introduction to Python Data Science (100), Mathematical Modeling (95), Mathematical Analysis (95.8), Probability & Mathematical Statistics (99)
Research Interests: Large Language Models (LLMs), Document QA Agents, Agentic RL, Deep Learning, Causal Inference

PUBLICATIONS

Zhanli Li, Huiwen Tian, Yixuan Cao, Ping Luo. *HDReAct: Multi-Turn Hierarchical Document Retrieval and Reasoning. KDD 2026* (CCF-A), **Under Review**, December 2025. [**First Author**]

Zhanli Li, Yixuan Cao, Lvzhou Luo, Ping Luo. *Navigating Large-Scale Document Collections: MuDABench for Multi-Document Analytical QA. ACL 2026* (CCF-A), **Under Review**, December 2025. [**First Author**]

Zhanli Li, Zichao Yang. *ESG Rating Disagreement and Corporate Total Factor Productivity: Inference and Prediction. Finance Research Letters* (JCR Q1, IF: 7.4), **Accepted**, October 2024. [**First Author**]

RESEARCH EXPERIENCE

HDReAct: Embedding-Free Long-Document QA via Hierarchical Driven Agents†
Role: Project Leader *September 2025 - Present*
Addressed the limitations of Retrieval-Augmented Generation (RAG) models, specifically the reliance on embedding-based recall and the inability of subsequent Rerankers to utilize document structural information.

- Proposed **HDReAct**, a method using document structure as a core prior. It guides the model to perform dynamic search and decision-making across chapters and paragraphs via an explicit **hierarchical indexing tree**.
- Achieved **88% answer accuracy** on **FinanceBench** without using any embedding models, relying solely on symbolic retrieval and structural navigation, significantly outperforming the concurrent work **FinSage** (CIKM 2025).
- ▷ **Preparing submission to** *KDD 2026 (CCF-A)* as First Author.

MuDABench: Multi-Document Analytical QA Benchmark and Multi-Agent Framework†
Role: Project Leader *May 2025 - December 2025*
Large-scale document collections contain rich knowledge, yet scalable multi-document analysis remains an unsolved challenge.

- Constructed the **largest** multi-document QA benchmark to date using **distant supervision** labeling to evaluate retrieval and reasoning systems.
- Experimental results revealed that OpenAI's **File Search** achieved a best accuracy of only **0.1368**.
- Developed a **Multi-Agent interactive workflow** that improved accuracy to **0.2651**, identifying that the primary bottleneck in current multi-document QA is the accurate extraction of information from single documents.
- ▷ **Submitted to** *ACL 2026 (CCF-A)* as First Author.

XGBoost-SHAP: Causal Inference from the Perspective of Interpretable ML*
Role: Project Leader *September 2024 - Januanry 2025*
Addressed the reliance on linear panel models in corporate finance research by expanding the interpretability of fixed-effect models.

- Introduced an **XGBoost** automatic hyperparameter optimization framework based on **Optuna**, achieving an R^2 of 0.76 in TFP prediction, outperforming all concurrent works.
- Integrated **SHAP** interpretability methods to reveal the non-linear relationship and **internal mechanisms** between ESG rating disagreement and corporate productivity.
- ▷ **Published in** *Finance Research Letters (JCR Q1, IF: 7.4).* Selected as **Outstanding Paper at Tsinghua University Causal Inference Seminar (Top 3%)**.

† *Conducted under the supervision of Associate Professors Yixuan Cao and Ping Luo, Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences*

* *Conducted under the supervision of Assistant Professor Zichao Yang, Wenlan School of Business, Zhongnan University of Economics and Law*

INTERNSHIP EXPERIENCE

Beijing Paoding Technology Co., Ltd. —Research Dept. *June 2025 - Present*
Role: AI Research Intern **Mentors:** Ping Luo, Yixuan Cao
Project: Commercial R&D and Implementation of Iterative RAG

- Leveraged the company's technical accumulation in **document parsing** by utilizing PDFlux and integrating results into the commercial RAG system ChatDOC.
- Implemented an **Iterative Retrieval-Reasoning (Iterative RAG)** framework. Utilizing multi-turn tool calling and context feedback loops, the model progressively filters, aggregates, and corrects retrieval results across different document levels.

OPEN SOURCE PROJECTS

Auto-Tutor: Intelligent Email Generation System for Research Applications *GitHub 80+ Stars*
Project Creator

- Developed a workflow to automatically parse user resumes and mentor information.
- Utilized LLMs to batch-generate personalized research application and communication emails, creating a one-click "Intelligent Application" workflow.

SELECTED AWARDS

- Hunan Province Outstanding Student (Top 0.1 % in Hunan), April 2023
- Central China Mathematical Modeling Competition (First Prize, Team Leader, Top 3%), May 2025
- Outstanding Paper of 2025 Workshop of Big Data and Causal Inference in Tsinghua University (First Author, Top 3%), July 2025
- China National Scholarship (Top 0.2% in china), Novemver 2025
- China Undergraduate Mathematical Contest in Modelling (First Prize, Team Leader, Top 3%), November 2025

TECHNICAL SKILLS & SERVICE

Language: Chinese (native) English (CET-6)
Programming Languages: Python, C/C++, LaTeX, Markdown
Tools: Docker, SSH, tmux, Ubuntu
Libraries & Frameworks: transformers, torch, sklearn, pandas, numpy, ms-swift, vllm, verl, llamaindex
Academic Service: Independent Reviewer for *Finance Research Letters*