r12323011@ntu.edu.tw | githubjacky | blog | Hsiu-Hsuan Yeh

Hsiu-Hsuan Yeh

A second-year master's student researching social networks and LLMs' economic applications.

I. EXPERIENCE

Department of Economics, National Taiwan University . . . . . January 2024 – Present
Independent Researcher

- Currently collaborating with Prof. Ziho Park on improving the construction of macroeconomic textual indicators using LLMs as black-box optimizers, served in a containerized environment with Ollama and Docker Compose.
- Previously worked with Dr. Chung-Chi Chen, exploring the potential of deploying LLMs to automate and enhance textual index curation workflows. Our work has been published in CIKM'24, a top conference of data mining.
- Developing a Python package with disciplined Git workflow, llm-research, built on top of the LangChain framework to provide an easy-to-use few-shot prompting API with MLflow integration for enhanced inspection of LLM outputs.

Behavioral and Data Science Research Center (BDSRC) . . . . . July 2023 – Present
Research Assistant

- Worked with Prof. Chih-Sheng Hsieh, assisting him with data collection and wrangling, with the main focus on cleaning a year's worth of call detail records with over a billion rows per month by utilizing PySpark and cuDF-Polars.
- Streamlined the data wrangling process for over 600,000 individuals by utilizing Joblib and Numba's guvectorize, reducing processing time by up to 90% and accelerating the computation of user-defined functions (UDFs).
- Collected and processed over 50 years of data on NFL and NBA players, teams, and coaches, enhancing the scraping process by leveraging httpx for sending HTTP requests and selectolax for HTML parsing.
- Leveraged GoogleTrendsAnchorBank, to collect monthly discussion volume data for both victims and perpetrators involved in the #MeToo movement, incorporating rate limiting mechanism to prevent 429 Error.
- Deployed rootless Docker with GPU support on the BDSRC server to safeguard access to system resources and maintain the privacy of containers by preventing users from running others' containers.

Institute of Economics, Academia Sinica . . . . . November 2022 – May 2024
Research Assistant

- Delivered a comprehensive presentation to a study group on fundamental machine learning concepts, including overfitting prevention techniques, regularization methods, and adaptive boosting algorithms.
- Worked with Prof. Tzu-Ting Yang on econometric research, applying XGBoost to classify treatment and control groups in minimum wage and estate tax policy studies, while optimizing model performance through Optuna.

Department of Economics, National Taiwan University . . . . . July 2022 – June 2023
Research Assistant, Teaching Assistant

- Redesigned and optimized the Julia package, SFrontiers.jl (originally developed by Prof. Hung-Jen Wang), through modular code architecture and enhanced type stability via multiple dispatch.
- As the primary teaching assistant of the course, Introduction to Numerical Methods in Economics and Econometrics, I taught the basics of the Julia programming language, wrote homework solutions, and graded all homework assignments.

II. PUBLICATIONS

Hsiu-Hsuan Yeh, Yu-Lieh Huang, Ziho Park, and Chung-Chi Chen. 2024. Automation of Text-Based Economic Indicator Construction: A Pilot Exploration on Economic Policy Uncertainty Index. In Proceedings of The 33rd ACM International Conference on Information and Knowledge Management (CIKM'24) (Acceptance Rate: 26.76%, 141/527)

III. EDUCATION

National Taiwan University . . . . . September 2019 – Present
M.A, B.A. in Economics