Jui-Chien (Ryan) Lin

Troy, NY ⋄ 518-816-5091

ryanjclin@gmail.com ⋄ https://www.linkedin.com/in/ryanjclin

EDUCATION

Rensselaer Polytechnic Institute (RPI) | Troy, NY | GPA: 3.93   Sept. 2023 –
Master of Computer Science, Interest: AI, DL, NLP

Tsinghua University | Beijing, China   Sept. 2019 – Jan. 2023
Master of Engineering in Automation

National Central University | Taoyuan, Taiwan   Sept. 2015 – June 2019
Bachelor of Science in Mechanical Engineering

WORK EXPERIENCE

AI Intern | GliaCloud Co., Ltd. | Taipei, Taiwan   May 2024 – Aug. 2024

- Developed **RAG** system, reducing 'type: ignore' annotations by 25% and improving code readability
- Implemented unsupervised video classification using **LLMs** for market insights on video preferences
- Created Intelligent Flowchart Conversion System to enhance workflow efficiency and data reusability

AI Intern | Didi Chuxing Technology Co. Ltd. | Beijing, China   Sept. 2021 – Nov. 2021

- Implemented machine learning techniques to determine optimal taxi subsidy rates, boosting customer demand
- Identified key subsidy-related features from a pool of 1000+ variables using **correlation analysis**
- Developed predictive models using **XGBoost**, employing label optimization to enhance model performance and stability, resulting in reduced subsidy variations by 7%

PROJECTS EXPERIENCE

Language Models (LM) for Textual Data Evaluation | RPI   May 2024 –

- Generated synthetic data using **LLama3** and **GPT-4o** to enhance language model learning on text variations
- Developed hybrid approach combining **LM** and **deterministic functions** to assess text data quality
- Leveraged deterministic methods to measure the similarity and diversity between datasets to address the inherent opacity of neural networks

Preference Learning Using NLP techniques | RPI   Sept. 2023 – Mar. 2024

- Skillfully condensed extensive text using **LLama2**, emphasizing preference identification for improved detection
- Employed **BERT**'s output for precise preference classification, guided by instructive sentences to enhance performance. Implemented masking and segment embedding to facilitate entity comparison

Industrial Multivariate Time Series Generation | Tsinghua University   Feb. 2021 – Nov. 2021

- Utilized **AutoEncoder** to merge strongly correlated variables, enhancing inter-variable analysis
- Applied **wavelet decomposition** for efficient information extraction, shifting data from time to frequency domain
- Employed **GAN-TEST** and Maximum Mean Discrepancy (**MMD**) metrics for evaluation, resulting in **MMD** differences below 0.1 and **GAN-TEST** differences below 10% when comparing original and generated data

PUBLICATION

- Kang, Inwon, Sikai Ruan, Tyler Ho, **Lin, J.**, Farhad Mohsin, Oshani Seneviratne, and Lirong Xia. "LLM-augmented Preference Learning from Natural Language." arXiv preprint arXiv:2310.08523 (2023)
- **Lin, J.**, Mohsin, F., Bhamidipati, S., Xia, Li. "Generating Election Data Using Deep Generative Models." Submitted to The Workshop on Artificial Intelligence for Social Good(AI4SG-23) 2023
- **Lin, J.**, Yang, F. (2022) "Data Augmentation for Industrial Multivariate Time Series via a Spatial and Frequency Domain Knowledge GAN." Submitted to IEEE International Symposium on Advanced Control of Industrial Processes (AdCONIP) 2022