

HAOYANG LING

✉ hyfrankl@umich.edu 🔄 frankling2020 🌐 Haoyang Ling 📄 frankling2020.github.io

EDUCATION

- Master of Science in Information (Big Data Analytics Track)** **May 2024**
University of Michigan, Ann Arbor *Ann Arbor, MI*
- Excelled in database, data mining, information retrieval, and natural language processing courses with a **GPA of 4.0/4.0**.
- Bachelor of Science in Electrical and Computer Engineering** **Aug 2023**
Shanghai Jiao Tong University *Shanghai*
- Honored as **Outstanding Graduate with a GPA of 3.9/4.0** showcasing excellence in computer science and data science.

SKILLS

Programming and Tools	Python, C/C++, Verilog, Docker, MongoDB, Linux, Git, Bash
Web Development	HTML, CSS, JavaScript, Django, React
Data Analysis	Hadoop, Spark, SQL, R, Tableau
Linear Programming	Gurobipy

WORK EXPERIENCE

- Research Intern / Mcity AI Assistant and Safety Validation of Autonomous Vehicles** **Apr 2023 – Aug 2023**
Mcity at the University of Michigan, Ann Arbor *Ann Arbor, MI*
- Adapted the LangChain framework and constructed a React-based AI chatbot for Mcity in the Docker deployment, boosting understanding and public exposure.
 - Leveraged Git for version control and issue tracking with a record of over 100 commits with modifications spanning thousands of lines of code.
 - Coupled code simulation of urban and roundabout driving behaviors with linear programming methods to refine traffic accident distributions, achieving a precise correlation with real-world data, with less than 3% deviation.

PROJECT EXPERIENCE

- Research Assistant / Reasoning Ability Benchmark of LLMs via Complexity Classes** **Nov 2023 – Jan 2024**
University of Michigan, Ann Arbor *Ann Arbor, MI*
- Introduced a monthly updated, complexity-based benchmark with nine algorithmic tasks for Large Language Models (LLMs), effectively reducing susceptibility to fine-tuning approaches and mitigating the risk of model over-fitting.
 - Pinpointed reasoning challenges in high-complexity classes and showcased the enhanced learning proficiency of closed models compared to open models.
- Data Scientist / Movie Search Engine** **Oct 2023 – Dec 2023**
University of Michigan, Ann Arbor *Ann Arbor, MI*
- Pioneered a novel dataset with 47 queries from specific genres to movie scenarios with 2760 manually labeled data points and identified limitations in genre drift and thematic interpretation.
 - Led the development of a movie retrieval pipeline with LGBMRanker, surpassing the BM25 baseline by a 0.11 increase in NDCG and achieving better alignment with user intent.
- Software Engineer / Magnetic Digital Microfluidic Control System** **May 2023 – Aug 2023**
Shanghai Jiao Tong University *Shanghai*
- Engineered a real-time droplet-control feedback loop for high-risk chemistry experiments, integrating the YOLOv5s model and robotic control on an Nvidia Jetson Nano board, achieving over 90% precision.
- Data Scientist / Machine-Generated Text Detection for ChatGPT** **Feb 2023 – Apr 2023**
University of Michigan, Ann Arbor *Ann Arbor, MI*
- Innovated a machine-generated text classification technique using TF-IDF logistic regression, attaining 90% accuracy on par with BERT models across four datasets.
 - Unveiled model robustness through data augmentation and underscored dataset design's role in model resilience after revealing a 5% performance drop due to sentence truncation.
- Big Data Developer / Music Recommendation with Spark** **Jul 2022 – Aug 2022**
Shanghai Jiao Tong University *Shanghai*
- Authored a reusable Dockerfile customized for Hadoop and Spark ensuring efficient setup in 20 minutes.
 - Enhanced parallel breadth-first search for music recommendation with Spark and data compression via Apache Avro achieving a 10x speed boost over MapReduce.