Yuxuan (Wayne) Wang

929-583-8556 • yw5954@nyu.edu • https://www.linkedin.com/in/wayne-wang-364174231/

EDUCATION SUMMARY:

New York University, Tandon School of Engineering, Brooklyn, NY   Expected Graduation Date: May 2025
Bachelor of Science, **Computer Science**, Math. **GPA: 3.935**

TECHNICAL SKILLS:

Python || C++/C || Linux || Github || Search Engine Optimization (query processing, dense/sparse re-ranking) || **Machine Learning** (Deep Learning, Reinforcement Learning) || **DevOps Tools** (Docker, Circle-CI, Clang) | **Database Management** (MySQL, NoSQL) || **Web Development** (Flask, SwiftUI)

RELEVANT COURSEWORK:

Computer Science: Algorithmic Machine Learning and Data Science || Network Security || Computer Networking || Data Structure || Algorithm || Computer Architecture and Organization || Database || Operating System

EXPERIENCES:

Researcher – Evaluation of Graph-Based Vocabulary Mismatch Solution in Information Retrieval   June 2024 – Present
Supervised by Professor Torsten Suel @ New York University, Tandon School of Engineering
(Information Retrieval, Search Engine, HPC, Query Processing, Database, Python)

- Investigated the effectiveness of graph-based neighbor search and expansion to address vocabulary mismatch in information retrieval, with a primary focus on LADR (Lexically-Accelerated Dense Retrieval).
- Designed an experimental setup to evaluate LADR's efficiency in reranking by using various initial sets of 200–1000 passages obtained through different retrieval methods, including BM25, DeepImpact, DeeperImpact, SPLADEv2, TILDE, and docT5query. This setup allowed for efficient reranking within a pool of 3,000 passages rather than the entire 8.8 million document corpus.
- Testing with multiple dense models (TasB, ColBERT, DeBERTa, and MiniLM) Demonstrated that the pool-based ranking within 3,000 passages achieves results comparable to full reranking of the entire corpus. Additionally, experiments using both KNN and HNSW graph-building methods confirmed that graph-based expansion provides an efficient intermediate solution for handling vocabulary mismatch, allowing cheaper initial retrieval methods to achieve robust results in reranking.

Internship - Industrial Investment Data Analytics   June 2023 - August 2023
Sinosure (Tianjin) Equity Investment Fund Management Co., Ltd., Guangzhou, China
(Deep Learning, Web Crawler, Database, Python, MySQL)

- Developed and implemented an automated data pipeline to efficiently extract, analyze, and present insights on Series B funding rounds in the Electronic Design Automation (EDA) software sector, improving the timeliness and quality of information delivered to decision-makers by 50%.
- Consolidated and standardized disparate data sources into a unified forecasting model, increasing data accuracy by 15% and enabling a 20% faster turnaround on investment analysis reports, which improved the precision of profitability assessments and strengthened decision-making.

PROJECTS:

DTCC Lab-The Future of Coding with AI (Vertically Integrated Project at Tandon)   September 2023 - May 2024
(Prompt Engineering, Automated Code Generator, Python)

- Designed and optimized AI-driven prompts leveraging the ChatGPT 3.5/4 API to elevate Python code generation, with a primary focus on complex data manipulation and algorithmic problem-solving for LeetCode-style challenges.
- Achieved a benchmark with over 80% of structured prompts under 75 lines executing flawlessly enhancing LLM precision and efficiency in automated coding workflows.

ContextWIN: Whittle Index Based Mixture-of-Experts Neural Model For Restless Bandits With Contextual Information Via Deep RL (https://arxiv.org/abs/2410.09781)   September 2023 - December 2023
(Theoretical Computer Science, Probability, Recommendation system, Reinforcement Learning, Python)

- Developed ContextWIN, an advanced architecture enhancing Neural Whittle Index Network for Restless Multi-Armed Bandit problems. Utilized reinforcement learning and a mixture of experts to improve decision-making in dynamic environments, especially recommendation systems.
- Proved the theoretical robustness of both NeurWIN and ContextWIN models, laying the groundwork for future applications in complex decision-making with contextual data.