Yuxuan (Wayne) Wang

https://waynewangyuxuan.github.io/personal_site/ || yw5954@nyu.edu

EDUCATION SUMMARY:

New York University, Tandon School of Engineering, Brooklyn, NY Expected Graduation Date: May 2025
Bachelor of Science **Computer Science**, Math. **GPA: 3.944**

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, Kubernetes, Prometheus) || **Database Management** (MySQL, NoSQL) || **Web Development** (Flask, SwiftUI, React) || **PyTorch** || **TensorFlow**

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:

Research - New York University June 2024 – Present

Evaluation of Graph-Based Vocabulary Mismatch Solution in Information Retrieval

Supervised by Professor Torsten Suel @ New York University, Tandon School of Engineering
(Information Retrieval, Search Engine, Query Processing, Database, Python)

- Conducted a systematic study on how seed quality impacts graph-based expansion in LADR (Lexically-Accelerated Dense Retrieval); designed a plug-and-play retrieval pipeline supporting multiple sparse retrieval models (BM25, DeepImpact, SPLADE, etc.) as initial seed sets.
- Implemented a graph expansion module for candidate documents using KNN/HNSW; incorporated a vector dimension masking mechanism (PRFDIME) that constructs semantic masks via pseudo relevance feedback to enhance query representation and improve reranking effectiveness.
- Through multiple experiments, demonstrated that reranking using only the top-3000 graph-expanded passages via Bi-Encoder/Cross-Encoder models can achieve comparable performance to full-corpus reranking, while improving recall and reducing computational cost.

Data Analyst Intern - Sinosure (Guangzhou) Equity Investment Fund Management

June 2023 - August 2023
(Machine Learning, Web Scraping, Data Engineering, Python, AWS, NoSQL, NLP)

- Led the development of a structured investment news processing system by integrating Wind data and web crawlers; extracted key fields such as company names, funding rounds, amounts, and investors; enriched company profiles using financial indicators (P/E ratio, revenue, valuation) to improve data quality.
- Designed semantic analysis and event extraction modules using a Chinese BERT model for sentiment classification, keyword aggregation, summary generation, and news type labeling; output results in a unified JSON format and pushed them to Kafka for daily reports and frontend display.
- Prototyped a visualization component using Streamlit, supporting word clouds, investor rankings, label distribution charts, and news card previews; contributed to the standardization of data formats and field definitions across frontend and backend systems.

PROJECTS:

TripPlanner September 2024 - Present
(Python, AWS Serverless, API Integration)

- Designed and implemented a route optimization module leveraging Amazon Location Service and AWS Lambda, supporting multi-point route planning and POI sequence optimization (lightweight TSP); cached static route maps in S3 to improve loading efficiency.
- Built a configurable real-time notification system enabling users to customize frequency and level of alerts; combined OpenWeatherMap API with SNS to push alerts for weather or itinerary changes; introduced a message aggregation mechanism to reduce redundant notifications.
- Adopted a fully Serverless architecture (Lambda, DynamoDB, Step Functions, ElastiCache, SNS) to ensure high availability, low maintenance cost, and scalability.

ShadowDash September 2024 - Jan 2025
(Build Optimization, C, CI/CD Deployment)

- Redefined build configuration by developing a C-based declarative language that replaced .ninja files, enabling direct in-memory graph construction and eliminating inefficiencies in Ninja's character-by-character parsing process.

- Engineered advanced optimizations, including scope encapsulation and parallel compilation, achieving up to a 2x speedup for LLVM manifests and a 1.73x improvement for zlib parsing, making ShadowDash a high-performance alternative for large-scale projects.
- Streamlined CI/CD pipelines by integrating CircleCI, enabling automated testing, benchmarking, and deployment workflows to ensure reliability and maintainability across complex configurations.

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)

- Proposed ContextWIN, a Whittle Index-based Mixture-of-Experts neural model, and provided rigorous mathematical proof for its robustness and optimality in Restless Multi-Armed Bandit (RMAB) problems.
- Designed a novel algorithmic framework combining reinforcement learning with Mixture-of-Experts models to effectively handle complex contextual information, offering theoretical support for recommendation systems and other dynamic decision-making tasks.