

Yiyang Bian

☎ (+1) 216-354-5035 | ✉ yiyangsteven@gmail.com | 🏠 www.bianyiyang.me

Education

University of California, Riverside

PH.D. IN COMPUTER SCIENCE

- Research: spatial data management, information retrieval, and RAG/LLMs.

Riverside, CA, USA

Sep. 2024 – Present

Case Western Reserve University

M.S. IN COMPUTER SCIENCE

- Research: recommendation systems, information retrieval, and database systems.

Cleveland, OH, USA

Aug. 2021 – May. 2024

Central China Normal University

B.S. IN COMPUTER SCIENCE

Wuhan, China

Sep. 2017 – Jun. 2021

Projects

Multi-Agent RAG System for Spatial Reasoning

RESEARCH PROJECT

Riverside, CA, USA

Apr. 2025 – Present

- Designed and implemented a multi-agent Retrieval-Augmented Generation (RAG) system to answer natural language queries with geographic constraints and user preferences.
- Built modular agents for spatial parsing, semantic retrieval, and answer generation, coordinated via LangChain's Agent-executor for structured task delegation.
- Formulated ranking as a learning-to-rank problem, integrating transformer-based semantic embeddings with geospatial features (distance, route/area containment, travel time).
- Implemented a dual-stage retrieval pipeline: FAISS dense vector search for candidate recall, followed by a cross-encoder ranking model trained with pairwise/listwise objectives (BPR, nDCG).
- Achieved higher Precision@10 and NDCG@10 than Pareto rule-based baselines on TourismQA, demonstrating the benefits of ML-driven ranking within a RAG framework.

Graph-based Top-k ML Model Recommendations

RESEARCH PROJECT

Cleveland, OH, USA

Feb. 2023 – Nov. 2023

- Developed ModsNet, a knowledge-graph-enhanced framework for selecting top-k data science models given an example dataset, addressing cold-start scenarios without retraining or fine-tuning.
- Formulated the task as a GNN-based regression problem, learning from a model-data interaction graph enriched with meta-features from models, datasets, and historical performance.
- Designed a probe-and-select strategy to adaptively identify promising candidate models under a cost budget, and applied clustering-based sparsification to prune redundant interactions, reducing training cost by 22.85%.
- Achieved state-of-the-art Precision@k and NDCG@k on image classification, XRD peak-finding, and text classification benchmarks, outperforming 10+ baseline methods in recommender systems and information retrieval tasks.

Work Experience

BioInVision Inc.

SOFTWARE ENGINEER INTERN

Cleveland, OH, USA

Mar. 2022 – Aug. 2024

- Designed and implemented a scalable cloud-based platform for large-scale biomedical imaging data retrieval and visualization, enabling multi-gigabyte 2D/3D datasets to be processed and displayed in real-time.
- Built RESTful APIs in Flask to support high-throughput queries and metadata search, integrating Azure Cosmos DB for distributed storage and low-latency retrieval.
- Applied WebGL-based rendering pipelines to accelerate interactive exploration of volumetric data, supporting downstream analytical and machine learning workflows.

- Developed back-end services for a Smart City data integration system, automating ingestion, transformation, and indexing of 100,000+ municipal records to support real-time analytics and predictive modeling.
- Designed and optimized RESTful APIs with Java/Spring Boot and MySQL, reducing query latency by 30% through indexing strategies and query execution plan tuning.
- Collaborated with cross-functional teams to align data pipelines with AI-driven decision support modules, enabling large-scale retrieval and analysis tasks.

Research Publications

M. Wang, H. Ma, **Y. Bian**, et al.. *Generating Skyline Datasets for Data Science Models* — EDBT 2025

Y. Bian. *Extending Collaborative Filtering for Machine Learning Model Recommendation* — OHIOLINK 2024

M. Wang, S. Guan, H. Ma, **Y. Bian**, et al.. *ModesNet: Performance-Aware Top-k Model Search Using Exemplar Datasets* — VLDB 2024

M. Wang, S. Guan, H. Ma, **Y. Bian**, et al.. *Selecting Top-k Data Science Models by Example Dataset* — CIKM 2023

M. Wang, H. Ma, A. Daundkar, S. Guan, **Y. Bian**, et al.. *Crux: Crowdsourced Materials Science Resource and Workflow Exploration* — CIKM 2022

Skills

ML & IR	Recommender Systems, Learning-to-Rank, Knowledge Graphs, Graph Neural Networks
NLP & LLM	RAG, LangChain, HuggingFace Transformers, Semantic Retrieval, Vector Search (FAISS)
Frameworks	PyTorch, TensorFlow, scikit-learn, LightGBM, XGBoost
Data & Systems	Data Pipelines, Apache Spark, RESTful APIs, Microservices, Performance Optimization
Databases	PostgreSQL, PostGIS, MySQL, Azure Cosmos DB
Programming	Python, Java, C++, SQL
DevOps & Tools	Git, Docker, Kubernetes, Linux, VS Code