Xinmeng Li

www.linkedin.com/in/xinmengli2026
(+1) 339-674-8882
xinmengmass@gmail.com
https://xinmengli.github.io

Professional Summary

Accomplished **Sr. Applied Scientist** (Ph.D. Computer Science) with 3+ years of experience designing and productionizing state-of-the-art **Machine Learning** and **Large-scale Modeling** solutions. Expert in **Graph Neural Networks (GNNs)**, high-dimensional data ranking, and retrieval systems. Proven track record of scientific leadership, including a co-authored patent and a 30% quarterly performance uplift on complex prediction tasks. Proficient in bridging the gap between cutting-edge research and enterprise-grade products through scalable, science-driven AI architectures.

Professional Experience

2022 – present **Computational Modeling Scientist**, *Montai Therapeutics*

- Led the scientific design and **product integration** of a proprietary forecasting framework, utilizing multi-step reasoning and automated data-retrieval to process complex biological datasets.
- Architected a production-ready prototype within 3 months, securing **intellectual property protection** (co-authored patent) for novel algorithmic approaches to data-driven decision making.
- Developed robust **evaluation frameworks** for model reliability, ensuring auditable outcomes for regulated enterprise environments and teaching research best practices to engineering peers.
- Advanced deep neural network models using **GNN foundations** and ensemble structures, driving two consecutive quarterly OKR goals with a **30% performance uplift** per quarter.
- Recognized with the **Great Exhibition Behavior (GEB) Award**, the company's highest performance honor, for technical leadership and business impact.

2015 – 2022 **Research Assistant & Teaching Assistant**, *Tufts University*

- Defined complex research problems in **Natural Language Processing** and Graph ML, developing scalable solutions for high-dimensional data ranking and retrieval.
- Guided technical approaches for grant proposals and **mentored junior researchers** in implementing production-quality ML code and scientific workflows.

Summer 2019 **Summer Intern**, *Food and Drug Administration (FDA)*

- Developed **deep-learning models** for time-series forecasting, focusing on data grounding and feature engineering over large-scale, unstructured demographic datasets.

Technical Skills

Applied ML **Graph Neural Networks (GNNs)**, Ranking & Retrieval Systems, Recommender Systems, Causal Inference, Time-Series Forecasting, Statistical Analysis.

Modeling Deep Learning, Ensemble Methods, Feature Engineering, Grounding & Retrieval-Augmented Generation (RAG) foundations.

Programming Python (Expert: PyTorch, PyG, Scikit-learn, NumPy, Pandas), C++, SQL, MATLAB.

MLOps & Tools Docker, Kubernetes, Git, CI/CD for ML, Experiment Tracking (MLflow/WandB), Spark, Hadoop.

Education

2022 **Ph.D. in Computer Science**, Tufts University

2017 **M.S. in Computer Science**, Tufts University, GPA 3.9/4.0

2015 **B.S. in Computer Science**, Sichuan University, Top 2%

Selected Research Projects

Retrieval & Ranking **Large-Scale Deep Learning Ranking System (ESP):** Developed a state-of-the-art ranking model to identify optimal matches from large candidate corpora. Achieved **41% improvement** over industry baselines by optimizing retrieval and grounding processes.

Statistical Analysis **Sequence Analysis Pipeline (ASAP-SML):** Built a novel pipeline combining statistical testing and ML for high-dimensional data, focusing on **model stability** and feature importance in low-data regimes.

Publications

Main Conference & Journal Papers

2024 **Li X.**, et al. "An Ensemble Spectral Prediction (ESP) model for high-throughput discovery of small molecules." *Bioinformatics*.

2022 **Li X.**, et al. "Boost-RS: Boosted Embeddings for Recommender Systems in biological data discovery." *Bioinformatics*.

Workshops & Other Venues

2021 **Li X.**, et al. "Sequence Analysis Pipeline (ASAP-SML): A hybrid statistical and machine learning approach." *Journal of Computational Biology*.