JEFFREY YUAN

Glenview, IL | mailjeffreyyuan@gmail.com | 847-905-6888 | Personal Portfolio | Research Repository | Github

EDUCATION

Northwestern University Evanston, IL

MS in Statistics and Data Science June 2026 (expected)

BS in Data Science and Computational and Systems Biology, GPA: 3.95/4.00

- Honors: CME Group Excellence Award, CME Trading Challenge Top 5%, CME AI/ML Team Challenge Winner, MD+ Datathon Runner Up, CTC Options Theory Intensive, Northwestern Summer URG, Northwestern Conference Travel Grant
- Relevant Courses: Machine Learning and Sensing, Machine Learning on Graphs, Regression Analysis, Advanced Machine Learning, Information Management, Data Structures and Algorithms, Advanced Statistical Theory and Methods
- Teaching Assistant: Advanced ML, Data Science with Python, Data Science Project

SKILLS

- Programming: Python(Pytorch, Tensorflow, Keras, LangChain, Scikit-Learn, Statsmodel, Pandas, Numpy), R, SQL, Java, Git
- Technical: Deep Learning, Time-Series, Probability, Optimization, Statistics, LLMs, RAG, Reinforcement Learning, NLP, Derivatives, Recommendation Systems, Gen AI, Agentic AI, MLOps, Docker, Linear Algebra, Calculus, Data Visualization
- Tools: Cursor, Claude Code, Jupyter, GCP, Vertex AI, Azure ML, AWS, BigQuery, Snowflake, Tableau, MongoDB, Hive

EXPERIENCES

CME Group, *Year Round Data Science Intern*, Chicago, IL June 2024 - Present

- Engineered full life-cycle **ML pipeline** on Vertex AI for a **self-supervised GNN recommendation system** forecasting trading behavior with trading and digital activity data, resulting in **$254,050 of revenue** through strategic sales positioning
- Developed deep learning **time series financial model** to **forecast weekly returns** of 10Y Treasury Note futures using large financial data sets resulting in the sale of CBOT dashboards on CME derivatives and equity products
- Designed automated **Agentic AI** system for **Natural Language to SQL** data pipelines using Google's Gemini foundational model to retrieve data across multiple CME databases, reducing development time by **99%**
- Authored **BigQuery** code repository for institutional clients to calculate **TWAP/VWAP** benchmarks, enhancing CME Data Services products and generating **$360K in ARR** through improved tools for targeted trading strategies

Reliable and Efficient AI Lab, *Graduate Research Assistant*, Evanston, IL May 2025 - Present

- Building a **hierarchical reinforcement learning (HRL) AI agent** on a **fine tuned LLM** (Deepseek R1**),** integrating **dual-policy architecture** and **multi-component reward system** to optimize ICD-10 code generation and reasoning
- Designing a custom **multi-modal framework** leveraging **Large Language Models (LLMs**) and **Graph Neural Networks** (GNN) for novel **compound prioritization** with SMILE strings and molecular graphs

Royal Liverpool University Hospital, *Machine Learning Engineer,* Remote March 2025 - June 2025

- Developed **dynamic patching algorithm** optimizing spatial coverage and minimizing overlap by a **contour-based sliding window**, generating 150K+ high-quality patches for medical image analysis across 3 stain modalities
- Benchmarked **deep convolutional neural networks (CNNs)** architectures on per-stain and multi-modal fusion pipelines for CMIL (Conjunctival melanocytic intraepithelial lesion) classification, achieving **87.5% AUC** and **84.2% F1**-score

Significance Lab, Harvard Medical School, *Research Assistant*, Boston, MA Dec 2022 - June 2025

- Fine tuned 144 **supervised learning** models on the MIMIC IV dataset with **Azure ML**, optimized each model through cross-validation and Bayesian optimization, using SHAP values for enhanced ER resource allocation
- Performed retrospective cohort study on shock index trajectory, categorized patients into 5 groups via **unsupervised clustering** algorithms, analyzed with the Bayesian Information Criterion - validated with ANOVA and chi-squared tests
- Leading team undergrads in a **web-analysis** of Clinical Informatics fellowship pages utilizing the Screaming Frog **web-scraping** API and coordinated with Massachusetts General Hospital physicians to develope data extraction template

Learngle, *Founding Engineer*, Boston, MA Feb 2023 - Mar 2024

- Built **Retrieval-Augmented Generation (RAG)** pipeline with **Open-AI API** generating adaptive AI/ML education content
- Designed SuperMemo 2 based **adaptive learning algorithm** for personalized content presentation using MongoDB database
- Achieved **100% success rate** across **43 users** and **$10,000 in revenue** during the 2023 exam cycles

PROJECTS

Self-Supervised Sequential Recommendation with Graphs

- Engineered **self-supervised graph neural network** framework for **sequential recommendation** with temporal user-item interaction graphs and multi-level sequential encoders, using structural and temporal signals for future interaction prediction
- Designed a **hard negative sampling** mechanism to select non-interacted items most similar to a user's short-term embedding using **cosine similarity**, integrating **InfoNCE contrastive loss** to optimize short-term user representations

LLM Bias in Clinical Reasoning

- Conducted medical bias analysis of a **LLM** (Llama 3.1) utilizing its publicly available API through **HuggingFace,** evaluating performance against 10,179 STEP-1,2, and 3 medical examination questions
- **Tuned LLM** performance using progressive **prompt engineering**, identifying prompt structures with reduced bias