

Education

University of Connecticut

PHD IN STATISTICS

Connecticut, USA

Sep 2018 - Sep 2023

University of Wisconsin - Madison

MASTER IN STATISTICS

Wisconsin, USA

Sep 2016 - May 2018

Nanjing University

BACHELOR IN STATISTICS

Jiangsu, China

Sep 2013 - Jun 2017

Work Experience

Quantitative Trading Book in Ernst & Young U.S. LLP

SENIOR CONSULTANT

New York, USA

Oct 2023 - Present

- **Assistant AI System for Intelligent Data Retrieval and Analysis**
 - Developed an AI assistant that interprets user input and retrieves relevant data from internal databases for automated analysis and summarization.
 - Implemented a **Retrieval-Augmented Generation (RAG)** architecture: embedded structured and unstructured content—including text, tables, and images—using multi-modal encoders.
 - Transformed user queries into dense vector representations and performed semantic search to identify top-matching data entries.
 - Constructed dynamic prompts from retrieved content to interface with LLMs, enabling explainable AI-generated responses.
- **Automated Sensitive Information Detection and Classification**
 - Built a neural network system to automatically assess and classify documents based on the presence of sensitive information (e.g., customer records, internal data) at the moment of file creation or saving.
 - Embedded semi-structured documents (text, tables, metadata) using a hybrid multi-modal pipeline combining Sentence-BERT for free text, TAPAS for tabular data, and early fusion for layout features.
 - Applied a shared encoder layer followed by task-specific classification heads for each sensitivity category.
 - Combined multi-head outputs to determine the overall confidentiality level, enabling real-time access control and compliance labeling.
- **Personalized Ad Recommendation System**
 - Developed a targeted advertising system to match users with the most relevant ads based on behavioral and demographic features.
 - Applied **K-Means clustering** to segment users into behavior-based cohorts, enabling personalized downstream modeling.
 - Implemented a **DeepFM** model to capture both low-order feature interactions (via factorization machines) and high-order nonlinear patterns (via deep neural networks).
 - Accurately predicted user click-through rates (CTR), improving ad relevance and conversion efficiency across user groups.
- **Modular Redesign of Derivatives Pricing Algorithm**
 - Led the architectural overhaul by decomposing the algorithm into service class and analysis units, achieving high **decoupling** of code.
 - Enabling independent updates to each component without affecting the overall system, significantly reducing redundancy and enhancing maintainability.
 - Designed robust unit testing frameworks, improving system **debug reliability** by proactively identifying potential errors.
- **Optimization of American Options Pricing**
 - Applied the American Monte Carlo (**AMC**) method to price American options, replacing the original Monte Carlo over Monte Carlo method.
 - Achieved a substantial reduction in computational complexity from $O(n^2)$ to **$O(n)$** , cutting pricing time and saving considerable resources.
- **Equity Derivatives Pricing Algorithm Enhancement**
 - Improved the pricing framework for equity derivatives by transitioning from a market-based risk model to an underlying location-based risk analysis, enhancing accuracy and **interpretability**.
 - Integrated advanced machine learning techniques, such as **LSTM**, **random forest** models with traditional MCMC methods to price derivatives, enabling the pricing of complex toxic options with more than three underlying.

Bank of China International Holdings Limited

SECURITIES ANALYST ASSISTANT (INTERN)

Shanghai, China

Jun 2021-Sep 2021

- Predicted the short- and long-term performance of new energy industry equity based on time series model with a spike-and-slab error.
- Adjusted the prediction under a multinomial model based on the performance of correlated companies, avoiding an over-optimistic forecast.

HUATAI SECURITIES CO., LTD. (HTSC)

DATA ANALYST (INTERN)

Jiangsu, China

Jul 2017-Sep 2017

- **Unsupervised** screened visitors with a strong desire to buy products based on their records on company's APP.
- Cleaned and reshaped the **17 million** visitor records by summarizing operations from the same visitor, and grouped them by **K-means**.
- Extracted useful variables by principal component analysis (**PCA**) method used in decision tree to tag visitor in 20s while the target is 1 min.

Skill

- **Language:** Mandarin Chinese (Native), English
- **Coding/Database Languages:** Master R, Python, GitHub, Latex, Nimble, JUGS and HPC, familiar with SQL, SAS, MATLAB, C++ and Julia.
- **Certificate:** CFA level 1