

## Summary

- My research interests lie in Applied Machine Learning and Data Mining, especially Adversarial Learning, Reinforcement Learning, Graph Neural Networks, and Natural Language Processing with applications to real-world scenarios (e.g. Finance, Traffic).
- Fast learner, self-motivated, collaborative, with ability to think critically and deliver analytical solutions clearly, concisely, and in a timely manner.

## Education

- 2020 – 2025: **PhD, Computer Science**, *New Jersey Institute of Technology*, Newark.  
Advisor: Dr. Guiling (Grace) Wang, Distinguished Professor
- 2017 – 2019: **Master, Data Science**, *Rutgers University*, New Brunswick.
- 2013 – 2017: **Bachelor, Economics & Statistics**, *Xiamen University*, China.

## Research Experience

FinTech Lab, AI Center for Research, New Jersey Institute of Technology

- Jun 2022 – **Portfolio Management with Margin by Reinforcement Learning**.  
Present: – Establish an RL-based model to maximize the profit and hedge the risk in margin trading.
- Jan 2022 – **Risk-Aware Generative Adversarial Model for Stock Interval Construction**.  
Sep 2022: – Employed a novel sequence generative model and statistical inference for *interval prediction* to quantify uncertainty of stock price, which is more informative than *point prediction*.  
– Adapted a Generative Adversarial Network (GAN) with temporal and risk modules to simulate a set of time series of future stock price with artificial noise from the finance market and horizon-wise information.  
– Incorporated the volatility index to capture the risk perceived by smart money, and designed a risk-sensitive interval to detect market volatility and potential needs for hedging.  
– Conducted extensive experiments on multiple stock indices worldwide, achieved 95% of coverage of true next-day price with precise range, and outperformed existing state-of-the-art benchmarks.
- Jun 2022 – **ADPP, A Novel Anomaly Detection and Privacy-Preserving Framework in Tokenomics**.  
Sep 2022: – Proposed a novel framework integrating a privacy-preserving authentication platform and deep learning techniques to secure users' privacy and regulate illicit behaviors for crypto assets trade.  
– Designed an anomaly detection system that utilizes a Graph Convolutional Network (GCN), Gated Recurrent Unit (GRU), and imbalanced learning on topological crypto asset flow among users to identify anomalous addresses and maintain a sanction list repository, detected 60%+ of illicit addresses in real-world crypto transaction data under dark market.
- Apr 2022 – **Scene Segmentation of Storybook**.  
Jun 2022: – Annotated sentences in the storybook by BIO tag to split the book into scenes.  
– Fine-tuned BERT, with Next Sentence Prediction (NSP), on consecutive sentence pairs in the storybook, utilized the NSP probability from BERT to detect scene boundaries.  
– Implemented SentenceBERT to segment the book according to movie script scene boundaries.
- Feb 2022 – **SafeLight: Collision-Free Traffic Signal Control with Reinforcement Learning**.  
May 2022: – Enhanced existing RL-based traffic signal control methods which focus on minimizing traffic delay.  
– Designed reward shaping considering both efficiency (delay) and safety (collision) in Deep Q Networks (DQN), evaluated model in real-world and synthetic datasets, achieved 85% reduction in collisions, and retained minimum traffic delay.

Sep 2021 – **Imbalanced Learning in Churn Prediction.**

Dec 2021: – Built classification models (SVM) with imbalance learning (sampling, cost-sensitive, and ensemble methods) to predict whether customers will subscribe to the service, achieved 0.7 of F1 score.

Mar 2021 – **Stock Price Prediction by Neural Networks and Sentiment Analysis.**

May 2021: – Designed a sentiment score utilizing VADER to extract public opinions from news text.  
– Integrated sentiment score with RNN (LSTM, GRU) to predict future stock price.

CAIT, Rutgers University, New Brunswick

Nov 2018 – **Long Term Bridge Performance (LTBP) .**

Apr 2019: – Developed deterioration model based on survival analysis and Markov Chain to forecast long-term survival performance of 150k Mid-Atlantic bridge components, simulated condition rating trend by Monte Carlo method.  
– Utilized Random Forest to select variables on raw bridge data and improved 15% efficiency.

---

## Publications

### Papers

AAAI'23 Wenlu Du, Junyi Ye, **Jingyi Gu**, Jing Li, Hua Wei, Guiling Wang. "SafeLight: A Reinforcement Learning Method toward Collision-free Traffic Signal Control."

IJAIA Wei Yao\*, **Jingyi Gu\***, Wenlu Du\*, Fadi P. Deek, Guiling Wang. "ADPP: A Novel Anomaly Detection and Privacy-Preserving Framework in Tokenomics." (\* indicates equal contributions.)

### Submitted & Under Review

KDD'23 **Jingyi Gu**, Wenlu Du, Guiling Wang. "Risk-Aware Generative Adversarial Model for Stock Interval Construction."

IJAIA **Jingyi Gu**, Guiling Wang. "Learning Stock Broad-Index Trend Patterns via Domain Knowledge-Informed Generative Network."

SN Business & Economics **Jingyi Gu**, Sarvesh Shukla, Junyi Ye, Guiling Wang. "Deep Learning Model with VADER Sentiment Score in Stock Price Prediction."

### In Preparation

**Jingyi Gu**, Wenlu Du, Guiling Wang. "MarginTrader: A Risk-Profit Balanced Reinforcement Learning Framework for Portfolio Management with Margin."

---

## Working Experience

May 2019 – **Data Scientist**, *Plymouth Rock Assurance*, Woodbridge, NJ.

Jul 2020: – Built tree-based models (GBM, XGBoost) to predict loss on 3 million policies in auto insurance, improved model performance by 4% regarding customer segmentation ability  
– Conducted Logistic Regression to predict policy lifetime expectancy, monitored customer retention trend change over time to allow early detection

Jun 2018 – **Clinical Data Analyst Intern**, *Roche*, Shanghai.

Aug 2018: – Generated data processing pipelines on clinical trial study data to detect discrepancies between drug administrating rules and protocol, guaranteed safety results.

---

## Skills

Programming Python (PyTorch, Tensorflow, HuggingFace), R, SQL, Java, LaTeX, Git  
Language English (Fluent), Chinese (Native), French (Basic proficiency)

---

## Honors and Leadership

2021 – 2022: **President**, *ACM-W, Graduate Woman in Computer Society (GWICS)*, NJIT.

– Organized the event "Roundtable Conversation with Dr. Jeff Ullman, Turing Award Laureate 2020"

2017: **Dean's List Award**, *The School of Economics, Xiamen University*.