Newark, NJ
(848)252-9195
jg95@njit.edu
Homepage
Github Linkedin

Jingyi Gu

Summary

– Research interests lie in Applied Machine Learning and Data Mining, especially Adversarial Learning, Reinforcement Learning (RL), Graph Neural Networks (GNN), and Natural Language Processing (NLP) with applications to real-world scenarios (e.g. Finance, Traffic), such as Portfolio Management, Time Series Forecasting, and Anomaly Detection.

– 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

Oct 2022 – Present: ***Portfolio Management with Margin by Reinforcement Learning***.
– Establish an RL-based model to maximize the profit and hedge the risk in margin trading.

Jan 2022 – Sep 2022: ***Risk-Aware Generative Adversarial Model for Stock Interval Construction***.
– 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 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 – Sep 2022: ***ADPP, A Novel Anomaly Detection and Privacy-Preserving Framework in Tokenomics***.
– 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 – Jun 2022: ***Scene Segmentation of Storybook***.
– 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 – May 2022: ***SafeLight: Collision-Free Traffic Signal Control with Reinforcement Learning***.
- 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.

CAIT, Rutgers University, New Brunswick

Nov 2018 – Apr 2019: ***Long Term Bridge Performance (LTBP)*** .
- 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

Peer-reviewed journal and conference papers (* equal contributions)

– **Jingyi Gu**, Fadi P. Deek, Guiling Wang. "Stock Broad-Index Trend Patterns Learning via Domain Knowledge Informed Generative Network." *International Journal of Artificial Intelligence & Applications 14(2):11-28*, March 2023.

– Junyi Ye, **Jingyi Gu**, Ankan Dash, Fadi P. Deek, Guiling Wang. "Prediction with Time-Series Mixer for the S&P500 Index." *IEEE 39th International Conference on Data Engineering (**ICDE**) Workshops on Big Data Analytics in Finance and Commerce (BDAFC)*, 2023

– Wenlu Du, Junyi Ye, **Jingyi Gu**, Jing Li, Hua Wei, Guiling Wang. "SafeLight: A Reinforcement Learning Method toward Collision-free Traffic Signal Control." *The 37th AAAI Conference on Artificial Intelligence (**AAAI**)*, 2023.

– Wei Yao*, **Jingyi Gu***, Wenlu Du*, Fadi P. Deek, Guiling Wang. "ADPP: A Novel Anomaly Detection and Privacy-Preserving Framework in Tokenomics." *International Journal of Artificial Intelligence & Applications 13(6):17-32*, November 2022.

Under Review

– **Jingyi Gu**, Wenlu Du, Guiling Wang. "Risk-Aware Generative Adversarial Model for Stock Interval Construction." *ACM SIGKDD Conference on Knowledge Discovery & Data Mining (KDD)*, 2023

– **Jingyi Gu**, Sarvesh Shukla, Junyi Ye, Ajim Uddin, Guiling Wang. "Deep Learning Model with Sentiment Score and Weekend Effect in Stock Price Prediction."

– Wei Yao, Wenlu Du, **Jingyi Gu**, Junyi Ye, Fadi P. Deek, Guiling Wang, Martin Kellogg. "Establishing a Baseline for Evaluating Blockchain-Based Self-Sovereign Identity Systems: A Systematic Approach to Assess Capability, Compatibility, and Interoperability."

In Preparation

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

– Ankan Dash, **Jingyi Gu**, Guiling Wang. "Label-aware Image Inpainting"

Working Experience

May 2019 – Jul 2020: **Data Scientist**, *Plymouth Rock Assurance*, Woodbridge, NJ.
- 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

Skills

Programming: Python (PyTorch, Tensorflow, HuggingFace), R, SQL, Java, LaTex, Git

Language English (Fluent), Chinese (Native), French (Basic proficiency)

Service

Reviewer IEEE Transactions on Knowledge and Data Engineering (TKDE)

Presenter The 1st International Workshop on Big Data Analytics in Finance and Commerce (BDAFC), The 39th IEEE International Conference on Data Engineering (ICDE) 2023

Honors and Leadership

2023: **Excellence in Award, Teaching Assistant**, *New Jersey Institute of Technology*.

2021 – Present: **Chair**, *ACM-W*, NJIT.
– Organized the event "Roundtable Conversation with Dr. Jeff Ullman, Turing Award Laureate 2020"

2021 – 2022: **President**, *Graduate Woman in Computer Society (GWiCS)*, NJIT.

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