倪冰慧

电话：13633834661 / 邮箱：binghui.ni02@gmail.com

教育背景

乔治城大学(US NEWS 排名第 22)　　硕士　　**2024/08--2026/05**

- **专业：**数据科学与分析　**GPA：**4.0/4.0
- **主修课程：**数据库系统与 SQL、机器学习、概率建模与统计计算、高级数据可视化、时间序列分析、大数据与云计算

中央财经大学　　**本科**　　**2020/09--2024/06**

- **专业：**财政学　**GPA：**89/100
- **主修课程：** 金融学、资产定价、财务管理、中级微观经济学、中级宏观经济学、计量经济学、公共财政学、税收学、会计学
- **奖项：**全国大学生数学竞赛三等奖、首都教育系统二十大主题征文一等奖、中央财经大学十佳志愿者

实习经历

科大讯飞， 数据分析实习生 (Python/NLP/数据库)　　**2023/11--2024/03**

- 数据挖掘：爬取医保局等网站的文本数据，利用正则表达式进行数据清洗，确保数据质量，为后续 NLP 处理做准备。
- NLP 与语义分析：使用 BGE 模型向量化文本数据，基于语义相似性分割文本，作为大语言模型的输入，增强语义理解能力。
- 数据库管理与查询优化：使用 AWS 平台创建并管理关系型(MySQL, PostgreSQL)与非关系型(MongoDB)数据库，熟练运用 SQL 进行高效的数据存储、查询优化和数据库管理，确保数据能够快速检索并支持后续分析。

方正证券， 量化投资实习生 (Python/ML/策略因子)　　**2023/07--2023/10**

- 构建回测模型，并对数据收集、数据清洗、股价预测、收益分析等方面进行模块化处理，提升策略验证的效率与准确性。
- PEAD 选股策略：使用 Python 的 jieba 库对 3 万条研报文本进行分词处理，构建 XGBoost 模型并通过交叉验证选择最优超参数，基于 log-odds 值差异作为选股因子，得到年收益率 17.62%和夏普比率 1.57 的投资组合。
- 多因子模型与因子优化：结合基本面因子、技术面因子与市场情绪因子，构建多因子选股模型；运用降维处理、回归分析和随机森林等机器学习算法，优化并筛选有效因子，实现了投资组合的预测能力与稳健性的提升。

国泰君安，固定收益实习生　　**2023/05--2023/07**

- 收集 Wind 数据库中重点板块的高频数据，进行行业周期的每周评估。
- 协助撰写月度《债券托管数据点评报告》，探究不同类型的金融机构对券种持有结构的变化，分析各机构债券杠杆率的走势。

中原信托，业务实习生　　**2022/01--2022/03**

- 从天眼查、企业预警通搜集企业信息，整理目标公司治理结构、股权结构，及财务数据。
- 分析企业业绩、财务情况及融资方式，使用 PEST 和波特五力模型进行宏观与行业分析；协助完成尽调报告和风险排查报告。

学术经历

客户反馈数据分析与智能分类 (Python)　　**2024/08--2025/01**

- 使用 K-Means 和 DBSCAN 等聚类方法，对 CFPB 客户投诉数据进行聚类分析，发现潜在投诉模式。
- 采用 KNN 和随机森林分类算法预测投诉类别和情感倾向，准确率达 85%，优化投诉自动分类与处理流程。

美元指数的波动分析与预测 (R)　　**2025/02--2025/03**

- 使用分解方法分析美元指数趋势及周期性波动，研究美元强弱对美国经济及全球市场的影响。
- 拟合 SARIMA 模型，准确率达 90%，为金融市场参与者提供可靠的美元走势预测。

互联网普及对财政透明度的影响　　**02/2023-07/2023**

- 使用 **R** 语言的 dplyr 包处理 2700 条财政透明度的面板数据，运用双向固定效应模型控制时间和城市效应，消除内生性影响。
- 使用 **Python** 爬取百度指数网站，获取各市对关键词条的搜索量作为代理变量，进行机制分析。

比赛经历

大学生创新训练项目　　**04/2022-04/2023**

- 协同团队成员对边疆地区乡村振兴的情况进行调研；使用层次分析法对地方政府绩效进行定性定量分析；
- 从政府网站等多渠道获取数据并进行交叉验证，根据广义生产理论将指标进行分类，并运用主成分分析法对指标进行降维处理；建立评估指标模型，优化乡村振兴绩效评估体系。

技能

- 计算机技能：Python (Numpy, Pandas, Sklearn, Matplotlib)、R(dplyr, ggplot, stargazer)、Stata、SQL、Git、Office
- 语言：英语（托福 105、GRE 325）

Binghui NI

185-1830-1588 / binghui.ni02@gmail.com

EDUCATION BACKGROUND

Georgetown University Expected Graduation Date: May 2026
Master of Science in Data Science & Analytics

- **GPA:** 4.0/4.0
- **Coursework:** Machine Learning, Database Systems and SQL, Data Science and Analytics, Probabilistic Modeling and Statistical Computing, Advanced Data Visualization, Applied Time Series for DS, Big Data and Cloud Computing

Central University of Finance and Economics Sep 2020 -- Jun 2024
Bachelor of Economics in Public Finance

- **GPA:** 89/100 (WES: 3.86)
- **Coursework:** Finance, Asset Pricing, Financial Management, Intermediate Microeconomics, Intermediate Macroeconomics, Econometrics, Public Finance, Principle of Taxation, Principles of Accounting
- **Awards:** Third Prize in the 14th Mathematics Competition of Chinese College Students, Top Ten Volunteer at CUFE

INTERNSHIP EXPERIENCE

iFLYTEK Co., Ltd Nov 2023 -- Mar 2024
Data Analyst Intern (Python)

- Crawled text data and applied regular expressions for data cleaning to ensure data quality.
- Employed the BAAI Embedding model to vectorize text data, and segmented long texts into fragments and QA pairs based on semantic similarity, providing input for large language models (LLM) and enhancing semantic understanding..
- Created relational (MySQL, PostgreSQL) and non-relational (MongoDB) databases on the AWS platform using SQL, performing efficient data storage, query optimization, and database management to ensure fast data retrieval and support subsequent analysis.

Founder Securities Co., Ltd Jul 2023 -- Oct 2023
Quantitative Investment Intern (Python)

- Built a backtesting system, including modules for data collection & cleaning, stock price prediction, and performance analysis, improving the efficiency and accuracy of strategy validation.
- Transformed 30,000 research report texts into a keyword frequency matrix, developed a XGBoost classifier for stock movement predictions, and optimized hyperparameters through cross-validation. The log-odds difference between the upward and downward categories was used as a stock selection factor, resulting in an annual return of 17.62% and a Sharpe ratio of 1.57.
- Combined fundamental factors, technical factors, and market sentiment factors to build a multi-factor stock selection model. Utilized dimensionality reduction to optimize effective factors, improving the predictive power and robustness of the investment portfolio.

Guotai Junan Securities Co., Ltd May 2023 -- Jul 2023
Fixed Income Research Intern

- Collected high-frequency data from key sectors in Wind database for weekly industry cycle assessments.
- Assisted in drafting the monthly *Bond Custody Data Report*, analyzing bond holdings and leverage ratios of financial institutions.

Zhongyuan Trust Co., Ltd Jan 2022 -- Mar 2022
Business Operations Intern

- Collected company information from Bloomberg, and organized governance, shareholding, and financial data.
- Analyzed company performance, financing methods, and conducted macro and industry analysis using PEST and Porter's Five Forces; assisted with due diligence and risk reports.

RESEARCH EXPERIENCE

Customer Feedback Data Analysis & Smart Classification *(Python)* Aug 2024 -- Jan 2025

- Used K-Means and DBSCAN clustering to analyze CFPB customer complaint data, uncovering potential complaint patterns.
- Applied KNN and Random Forest algorithms for classifying complaint categories and sentiment, achieving 85% accuracy and optimizing the automated classification process.

USD Index Volatility Analysis & Forecasting *(R)* Feb 2025 -- Mar 2025

- Analyzed the trend and seasonality of the USD index using decomposition methods, discovering its impact on the economy.
- Developed an optimal SARIMA model with 90% accuracy, providing reliable USD trend forecasts for market participants.

Impact of Internet on Fiscal Transparency *(R, Python)* Feb 2023 -- Jul 2023

- Constructed a two-way fixed-effects model to control time and city effects for mitigating the endogeneity problem.
- Scraped data from Baidu Index to obtain keyword search volume across cities as a proxy variable for mechanism analysis.

PROJECT EXPERIENCE

National Undergraduate Innovation Training Program *(R)* Apr 2022 -- Jun 2023

- Conducted research on rural development in China; used the Analytic Hierarchy Process to analyze local government performance.
- Used PCA for dimensionality reduction, and established an evaluation model to improve the rural revitalization performance system.

SKILLS

Data Analytics, Data Science, Machine Learning, Data Visualization, Python, R, STATA, Bash, Git, SQL