

# 朱奕舟

南京 · 江苏 / +86 139-1299-6212 / yizhou\_zhu200320@163.com

## 教育背景

### 纽约大学

美国 · 纽约

数据科学与数学（本科）

辅修计算机科学

2022年9月–2025年12月

• GPA: 3.6/4.0

• 主要课程: 数据科学原理、因果推断、责任型数据科学、数据结构、数理统计、数理分析

## 实习经历

### 中国建设银行江苏省分行

江苏 · 南京

数据管理实习生

2024年12月–2025年1月

- 参与大数据风控解决方案研发，聚焦异常检测与反欺诈算法（孤立森林、单类支持向量机）及实时数据处理技术（Apache Kafka、Flink）；
- 参与第四届中国建设银行人工智能竞赛，基于24,000份客户画像与6个月信用卡还款数据，利用CatBoost、LightGBM及AdaBoost分类器构建客户违约风险预测模型，最终模型实现0.99的AUC与0.95的F1值；
- 完成约50万潜在客户基金投资偏好分析的特征工程，构建多分类预测模型，通过下采样数据优化LightGBM模型性能，最终达到0.76的准确率和0.64的F1值。

## 项目经历

### 2024年美国总统选举预测（基于机器学习）

2024年9月–2024年11月

- 分析历史选举和民调数据以预测2024年美国总统选举结果
- 探索了多种机器学习模型，包括线性回归、随机森林、XGBoost、梯度提升，并使用堆叠回归器进行精细预测，最终通过对比历史选举结果进一步验证预测的准确性

### Spotify歌曲特征分析与分类

2024年6月–2024年7月

- 使用直方图、散点图和相关性分析研究歌曲特征与流行度之间的分布和关系。
- 开发了基于回归、分类和PCA的模型来预测歌曲流行度。并通过Mann-Whitney U检验和t检验分析歌曲流行度的统计差异。

### 垃圾邮件分类（基于机器学习和深度学习）

2024年3月–2024年5月

- 对垃圾邮件和正常邮件的数据集进行了处理，包括分词、词形还原和停用词移除，并通过词云和柱状图对词频分布进行了可视化展示。
- 开发了多种模型，包括朴素贝叶斯、决策树、逻辑回归和梯度提升，同时还构建了一个用于分类的双向LSTM深度学习模型。
- 使用混淆矩阵、分类报告和准确率等建立指标进模型性能评估。

## 素质能力

编程语言: Java, Python, SQL

工具掌握: NumPy, Matplotlib, SK learn, L<sup>A</sup>T<sub>E</sub>X

语言能力: 英语, 中文