

陈霖

(+86) 18950336780 | 1969773923@qq.com | 上海

教育经历

上海交通大学 数学科学学院	2021 年 8 月 - 2025 年 6 月
理学学士, 统计学专业	
GPA: 3.75/4.3; 排名: 4/17	
导师: Samuel Drapeau 副教授, 林一青副教授	
荣誉奖项: 上海交通大学优秀毕业生, 上海交通大学数学科学学院优异学士学位论文, 上海交通大学三好学生, 上海交通大学数学科学学院统计学直博优秀营员 (排名 2/26), 上海交通大学本科生优秀奖学金 B 等 (2021-2022 学年与 2022-2023 学年两次获得), 美国大学生数学建模竞赛 H 奖, 上海交通大学数学建模校内赛二等奖, 军训优秀学员	
复旦大学 管理学院	2025 年 9 月 (预计) - 2027 年 6 月 (预计)
数据科学与商务分析硕士候选人	
荣誉奖项: 复旦大学管理学院新生奖学金二等	

论文

- [1] Lin Chen, Samuel Drapeau, Fanghao Shao, Xuekai Zhu, Bo Xue, Yunchong Song, Mathieu Lauriere, and Zhouhan Lin. α -GFN: Generalizing GFlowNets via Markov Chains for Better Forward and Backward Mixing. Under Review.
- [2] Bo Xue, Yunchong Song, Xuekai Zhu, Lin Chen, Fanghao Shao, Luoyi Fu, Xinbing Wang, and Zhouhan Lin. Flow of Spans: Generalizing Language Models to Dynamic Span-Vocabulary via GFlowNets. Under Review.

实习经历

上海乾矿商务咨询有限公司 (Trexquant)	2025 年 1 月 - 2025 年 1 月
part-time, Global Alpha Researcher	远程实习
分析 TheFly 新闻网站数据构建新闻情绪因子, 基于 BBG 分析师数据构建分析师因子, 基于日内流动性构建量价因子。	
使用 RidgeRegression, XGBoost, LGBM, XGBRF, GradientBoostRegressor 等模型, 基于财报公布日, 使用技术面、期权、新闻、分析师、基本面因子构建美股流动性前 3000 支股票的多空选股策略。	
九坤投资	2024 年 9 月 - 2024 年 12 月
量化实现部, IT 实习生	上海
基于 Kafka 与 PyArrow 优化风控任务的流式计算: 为风控系统增加了主备数据源切换功能, 并修复了盘中数据回放异常问题。	
根据交易合规条款, 迭代已有风控项目, 增加虚假报单等其他交易系统的盘后流式检测功能。	
使用 ClickHouse、MySQL 数据库与 Requests, 基于 Python 为已部署任务和风控指标配置增加检查与统计报告, 并部署上线。	
上海悬铃私募基金管理有限公司	2024 年 4 月 - 2024 年 7 月
研究部, 量化实习生	上海
使用 Python, 使用分钟级数据与 level2 行情数据, 基于交易逻辑构建量价选股因子。共构建并计算 562 个因子, 其中入库 28 个。	
基于因子数据分布及因子表达式的组合与筛选, 搭建自动化因子挖掘框架。共构建因子 37612 个, 计算约 1000 个, 入库 26 个。	
所有入库因子中, IC 绝对值最高可达 5.6%, IR 绝对值最高可达 0.7, 头组收益最高可达 21.4%, 横截面相关性绝对值不超过 0.4。	
上海乾上泉私募基金管理有限公司	2023 年 7 月 - 2024 年 1 月
研究部, 量化研究助理	上海
在 Linux 环境下, 使用 Python 与 SQL 脚本维护 SQL 数据库 (PostgreSQL) 中的现有因子。	
使用 Python 搭建因子回测系统, 并使用并行计算技巧优化因子的计算速度 (如 joblib 库, 向量化计算)。	
基于分钟级数据构建因子, 并计算 IC、RankIC、多空收益等指标, 并使用正交化、加权平均、回归等方法构建复合因子。	

项目经历

基于随机森林模型的地震后建筑受灾等级预测	2024 年 1 月 - 2024 年 1 月
《大数据分析》课程大作业, 获本课程大作业最高分, 指导教师: 毛晓军副教授	
使用 Python 进行探索性数据分析, 并完成特征工程与预处理。	
使用朴素贝叶斯、决策树、随机森林模型预测建筑物受灾等级, 并通过 SMOTE 重采样和超参数调优提升模型准确率约 10%。	
针对不同地区的受灾等级分布进行层次聚类, 并分聚类区训练模型以优化预测效果, 提升模型准确率约 5%。	
神经网络的量化	2023 年 12 月 - 2024 年 1 月
《并行与分布式程序设计》课程大作业, 获本课程大作业最高分之一, 指导教师: 宋卓然助理教授	
使用 Pytorch, 在 Cora, CiteSeer, PubMed 数据集对 GIN, GCN, GAT 神经网络进行伪量化实验。	
结合多种量化方法 (均匀量化、分层量化、分组件量化) 优化存储大小, 实现以 25% 的模型与数据大小达到原模型 95% 的精度。	

组织经历

上海交通大学 LUMIA 实验室	2024 年 10 月 - 至今
研究助理, 指导教师: 林洲汉副教授	
参与 GFlowNet 的理论与应用研究, 并探索新的语言模型架构与训练方法。	
上海交通大学学生交响乐团	2021 年 9 月 - 2024 年 9 月
低音提琴首席, 指挥: 吴一波副教授	
参与上海交通大学迎新音乐会、全国大学生艺术展演 (获省级一等奖)、上海市夏季音乐节、上海交通大学校庆等多场演出。	
作为乐团成员获上海交通大学校长奖、上海市青年五四奖章集体等荣誉。	

技能

编程技能: Python, C++, R, LaTeX, SQL
语言: 普通话 (二级乙等), 英语 (CET-6: 618, CET-4: 606)