

崔一帆

188-1066-9189 | 18810669189@163.com

教育经历

- 对外经济贸易大学
- 信息管理与信息系统 本科
- 2017.09 - 2021.06
- GPA: 3.17/4.0
 - 主修课程: 数据库系统, 数据库设计与实践, 智能计算, Python与大数据分析, 系统分析与设计
 - 荣誉与奖项: 综合三等奖学金, 全国大学生数学建模竞赛北京市一等奖

项目经历

- 基于蜡烛图和AutoEncoder的股价走势预测
- 2019.12 - 2020.02
- 项目描述: 将股票交易数据转换为蜡烛图, 使用AutoEncoder提取图像特征, 对特征聚类以选出投资组合
 - 项目职责:
 - 使用Tushare接口采集沪深两市的股票交易数据, 并使用Matplotlib转化为蜡烛图
 - 使用AutoEncoder将蜡烛图转化为100维的特征, 对特征进行K-means聚类, 得到走势相似的股票
 - 根据夏普指数对股票进行排序, 选取高收益低风险的股票投资组合

- 众安黑客松大赛-健康险理赔风控
- 2020.09
- 项目描述: 使用保险用户在理赔报案场景的案件相关数据, 建立算法模型以预测每个案件的欺诈风险水平
 - 项目职责:
 - 使用Python进行数据预处理及探索性分析, 并结合实际业务场景构建新的特征, 如事件发生与申请理赔的日期差
 - 融合经过调参之后的XGBoost和LightGBM模型, 对2w条样本进行预测, 最终micro F1 score为线下0.88, 线上0.82

- 基于自然语言处理的广告文章识别算法
- 2020.11 - 2021.01
- 项目描述: 从微信APP中爬取文章, 提取文本特征用于BERT模型的训练, 最终模型的AUC为0.84
 - 项目职责:
 - 编写微信APP爬虫, 爬取并下载2w篇文章。使用TF-IDF算法提取文章关键词, 对少样本进行回译增强以平衡样本
 - 使用BERT预训练模型将词语转化为向量, 添加softmax层用于分类, 判断文章是否为广告

技能/证书及其他

- 技能: C++, SQL, Python (数据分析, 机器学习), Excel
- 科学计算与机器学习: 精通Pandas, Numpy, Matplotlib, 熟悉数据结构和机器学习算法, 熟练使用Keras搭建CV和NLP模型
- 语言: 通过英语四级和六级