

# 庞秋

17878055223 | 2181236626@qq.com



## 教育经历

- 东北大学

2018年08月 - 2022年06月
- 信息管理与信息系统 本科
- 沈阳
- 相关课程：数据分析与数据挖掘、商务决策分析与模拟、商务大数据分析实践
- 大连理工大学

2022年09月 - 2025年06月
- 管理科学与工程 硕士
- 大连
- 相关课程：高级应用统计、数据库技术与应用、系统工程原理

## 实习经历

- 平安科技深圳有限公司 分析师（实习）

2022年06月 - 2022年08月
- 文本数据标注：**在宠物险（健康险）项目中首先需要根据用户朋友圈数据判断用户养猫狗宠物（健康）的可能性，再使用标注好的数据进行建模训练，实现宠物险（健康险）精准化推荐营销。期间主要负责在AI标注平台中对大量的的用户数据进行标注，累计标注数据9万多条。数据标注准确率达90%以上，为算法部门提供高质量数据，支持实现保险用户细分和精准化营销。
  - 图像数据标注：**在宠物险项目中，除了文本类的数据还有图像类数据，需要同时根据用户发表的文本和图像数据判断其养猫狗的可能性。期间主要负责对模型识别后的猫狗图像数据进行复核，累计复核图像数据2000多条。对模型所识别的图像数据质量进行把控，助力模型优化。
  - 模型训练结果统计：**使用vlookup对模型结果，各个特征的召回率、准确度、正确分类数、未召回数等指标进行统计。
  - 数据扩写：**根据给定的关键词（老人、小孩等）扩写具体场景下的相关语句，累计扩写数据400多条。

## 项目经历

- 天猫复购预测挑战赛

2022年08月 - 2022年09月
- 项目背景：**商家会在一些特定的日期开展大型的促销活动，比如618，双11。但促销吸引来的很多都是一次性的消费者，对销售业绩的增长没有长远的帮助。为此，商家需要预测双11中哪些新消费者会进行复购，再精准地投放广告，将其转换为忠诚客户。
  - 数据处理：**对5000多万条用户购物行为数据进行数据类型转换以压缩内存。为了保证数据质量，查看数据集空值、重复值、异常值情况，删除少量空值，重复值可认为用户同一天多次点击或购买某一商品，将其保留。
  - 特征处理：**在业务上，从用户、商家、用户与商家三个方面，结合现有数据集构建特征。例如用户年龄、性别、商家流量、产品评价等，累计构建16个特征。并对类别型数据进行独热编码。
  - 建模预测：**划分训练集和测试集数据，使用StandardScaler进行数据标准化，利用逻辑回归模型和随机森林进行建模预测。最后建立混淆矩阵，使用accuracy\_score、precision、recall、f1\_score对模型预测结果进行评分，选取随机森林为最终预测模型，分类准确率达95%。

## 专业技能

- 英语能力：**通过CET—4，CET—6，良好的英语听说读写能力。
- 办公能力：**通过全国计算机二级考试 MS Office高级应用，熟练使用word,excel,ppt等办公软件。
- 数据能力：**能够熟练使用Mysql对数据库中的数据进行增删查改操作；使用Python对数据进行清洗处理、数据分析、模型训练；使用Tebluea、FineBi等软件进行数据可视化操作。
- 编程能力：**掌握Python的基本操作，能够熟练使用Python中的numpy、pandas、matplotlib、seaborn、sklearn等库进行数据清洗、分析、建模、可视化。能够实现应用常见的机器学习学习算法（线性回归、逻辑回归、KNN、K-means、随机森林、决策树等）。