

个人信息

姓名：邹帅帅
电话：19946210139
政治面貌：中共党员

民族：汉
邮箱：1505213523@qq.com



教育背景

2015.09-2019.06 郑州轻工业大学 建筑电气与智能化

专业技能

- 熟练掌握自然语言处理相关技术，如命名实体识别、意图识别、语义理解等。
- 精通深度学习框架，如 PyTorch、TensorFlow等，能够运用其进行模型的训练和优化。
- 熟悉各种语言模型，如 DeepSeek、Qwen、BERT；以及模型微调推理框架，如Vllm、LLamafactory，能够根据项目的特点对预训练模型进行微调，提高模型的性能和泛化能力。
- 具备良好的编程能力，熟练使用 Python进行项目开发。
- 精通 Dataworks、Oceanbase、Oracle、Mysql、Hive 数据库，熟练使用 SQL 完成各项指标计算。

工作经历

- 2024.04-至今 上海海万信息科技股份有限公司 算法工程师
1. 根据业务需求，用 BERT架构模型对文本进行分类，提升分类准确率。
 2. 应用 BERT+ CRF 架构准确识别关键实体信息。
 3. 完成语义理解工作，对大语言模型进行微调训练用以理解文本语义，支撑智能应用。
- 2021.05-2024.03 上海新致软件股份有限公司 算法工程师
1. 对文本数据预处理，包括分词、词性标注、数据清洗等，保证数据质量。
 2. 模型训练与优化，借助深度学习框架搭建合适模型并提升性能。
- 2019.07-2021.04 安徽力瀚科技有限公司 数据开发工程师
1. 根据业务需求编写存储过程及MaxCompute流程；
 2. 使用数据库 SQL 语句编写查询；
 3. 参与数据治理工作，提升数据的易用性和数据质量。

项目经历

项目一：逾期客户智能催收项目

项目背景：

消费金融发展致逾期问题突出，传统催收效率低且缺个性，故开展智能催收项目提升效果与效率。

项目职责：

●文本预处理

运用规则与统计结合的分词，结合五折交叉法和混淆度分析，为催收文本数据预处理。

●意图分析

基于 Bert 模型微调构建多维度分类模型，结合数据增强、主题模型，高精度识别客户意图，支持催收策略优化。

●自然语言生成

使用Qwen模型，融合多种模型构建混合模型，根据不同的催收场景和客户特点，结合识别出的意图以及流程话术，生成针对于客户的催收话术。

项目成果：

- 效率提升：自动化催收大幅减少人工工作量，处理逾期客户数量较传统方式提升 13% 左右。
- 风险降低：还款率提升 2%，逾期率降低 5%，降低逾期风险。
- 满意度提高：客户满意度升 6%，投诉率降 11%。

项目二：签约合同质检项目

项目背景：

公司现有的合同单证存在签名与姓名不统一，个人信息图片模糊不清，盖章出现错误，视频有卡顿花屏等问题。

项目职责：

●问题分类

对现有合同单证中的内容和所要提取的数据字段进行分析，将整体分为三类，第一类直接输出校验结果，第二类输出校验内容信息，第三类分割模型核查识别。

●关键信息提取

使用Qwen2.5_VL模型进行微调，对合同中特定信息进行提取，按照指定格式输出，与系统中所记录的数据对比校验。

●提示词编写

证件种类繁多，针对不同类型的单证，编写适合的提示词，使其可以更加精准的提取所需要的数据字段。

项目成果：

- 降低风险：对于单证信息不统一或者不正确的案件进行拦截，降低了放款时的风险。
- 人工校验减负：项目上线后，减少了人工每日校验，可专注复杂问题。

项目三：违规关键词检测项目

项目背景：

本项目旨在检测坐席人员与客户沟通期间是否有不合规话术。

项目职责：

●数据采集与净化

通过语音识别和深度学习文本清洗算法处理对话记录，并对数据进行初标和格式整理。

●意图识别

使用Qwen模型对标注后的对话文本进行微调。用微调后的模型对坐席人员的话术进行意图识别，个别意图使用正则匹配，校验话术是否合规。

项目四：智能坐席辅助项目

项目背景：

本项目旨在为坐席人员打造一款智能辅助系统，通过运用先进的自然语言处理技术，提升坐席与客户沟通的效率和质量，优化业务流程。

项目职责：

●数据采集与净化

运用 NLP 信息抽取技术，结合模板和机器学习方法从保险文档提取信息，通过深度学习文本清洗算法处理对话记录。

●意图识别

基于词向量方法将文本转特征向量，匹配专家话术，为坐席提供意图决策支持。

项目成果：

- 沟通效率提升：坐席回答客户问题平均时间缩短约 20% - 30%，业务处理速度加快。
- 沟通质量提高：回答准确性与规范性提升，客户投诉率降低 15% - 25%。