**附件1**

入选编号： （学生不填）

**东华大学数学与统计学院“树苗”计划**

**——创新实践育人项目申请表**

负责人姓名： 刘星宇

负责人班级： 数学2401

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 东华大学数学与统计学院制表 | | | | | | |
| 填表日期： | 2025 | 年 | 5 | 月 | 22 | 日 |

1. **项目团队成员信息**

*队伍人数3~8人，第一栏填写负责人信息*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **序号** | **姓名** | **班级** | **学号** | **联系电话** | **邮箱** | **团队分工** |
| **1** | 刘星宇 | 数学2401 | 240110119 | 13995638560 | 3273744268@qq.com | 组长 |
| **2** | 李宜珊 | 统计 2402 | 240110405 | 18622618036 | yishanli\_emily@qq.com | 副组长 |
| **3** | 黄昱洲 | 统计 2401 | 240110118 | 19519355086 | 2425714851@qq.com | 组员 |
| **4** | 胡玮刚 | 统计 2402 | 240110414 | 13767354848 | 2646373243@qq.com | 组员 |
| **5** | 黄旭涵 | 统计 2402 | 240110417 | 13588972081 | 2076879579@qq.com | 组员 |
| **6** | 于 梦 | 统计 2402 | 240110411 | 15893120807 | 15893120807@163.com | 组员 |
| **7** | 郝尉辰 | 数学2401 | 240110116 | 18742259649 | 356416371@qq.com | 组员 |
| **8** | 周心语 | 数学2402 | 240110412 | 17701520356 | zxy13636357806@qq.com | 组员 |

**注：队伍中必须包含大一或大二学生**

1. **项目志愿情况**

*请填写“序号+项目名称”，最后列“是”请打勾,“否”则留空。*

|  |  |  |
| --- | --- | --- |
| **第一志愿** | 2中国人口及其上海房价预测 | * **是否已和老师初步沟通** |
| **第二志愿** | 15基于在线评价和用户偏好的消费决策 | * **是否已和老师初步沟通** |
| **第三志愿** | 16基于大语言模型微调的零售优惠信息提取 | * **是否已和老师初步沟通** |
| **□是否服从调剂** | |  |

1. **对意向项目的初步设想与理解**

*本栏未展示完全者可另附页*

|  |  |
| --- | --- |
| **第一志愿** | **一、研究背景**  中国人口形势严峻，总和生育率持续低迷，同时老龄化加速。人口系统受政策、社会观念、经济发展等多种复杂因素干扰，呈现非高斯噪声特征，传统预测模型难以精准预测。上海房价受供需、土地政策、金融政策及城市发展等多种因素影响，走势复杂。政策调控的间断性、经济波动及突发事件对购房者心理预期和购买能力的冲击，都体现出非高斯特性，现有房价预测模型亟需改进。随着人口结构变化，高校适龄人口将受显著影响。按18年教育周期推算，2035年后高校适龄人口将显著下降。高校系统受经济增速波动、教育政策调整等非高斯噪声影响，需新模型预测其规模变化。  **二、研究方法**  数据收集与整理：收集中国人口、上海房价、高校相关历史数据，涵盖人口普查、房地产市场交易、高校招生就业等多方面数据。  模型构建：运用基于Levy过程的人口预测模型、结合GARCH模型的房价预测模型、考虑非高斯噪声的高校规模预测模型，以捕捉系统中的非高斯噪声特征。  参数估计与模型验证：采用最大似然估计等方法对模型参数进行估计，并利用历史数据对模型进行验证和优化。​  预测与分析：运用优化后的模型对未来趋势进行预测，并结合政策、社会经济变化等因素对结果进行深入分析。​  **三、预期成果**​  本课题预期构建出适用于人口、房价、高校规模预测的非高斯噪声模型体系，将利用非高斯噪声模型，精确预测未来十年中国人口总量下降趋势，分析上海房价走势，同时预估高校规模收缩程度，提供未来十年中国人口、上海房价、高校规模及教师失业人数的详细预测报告。研究成果将为相关领域的学术研究提供新的思路和方法。 |
| **第二志愿** | **一、项目背景**  在如今数字化的时代，人们每天都会面临各种消费选择，然而在面对网络上 海量的信息和评价，如何找到真正适合自己的选择成了难题。因此，想通过本项 目进一步优化决策模型，将个人喜好与真实的评价相结合，从而做出更适合消费 者的消费决策。  **二、实现功能（将以旅游为例）**  理解用户需求：充分理解用户的出行目的、预算、感兴趣的旅游景点、对饮 食及住宿的要求偏好。  分析评价：从旅游网站（如携程）或社交平台（如小红书、大众点评）收集 关于景点、餐厅、酒店等的评论，提取关键词，总结出其相应的优缺点。  生成个性化方案：将景点的特点与消费者偏好相匹配，优先选取匹配度高的 景点，再根据游客的实际情况设计旅游线路。  **三、研究过程**  设计并分析问卷：制作涵盖预算、偏好禁忌等内容，邀请用户进行填写，并 根据问卷分析每个用户最关心的旅行因素。  整理景点、餐厅等数据：根据用户喜好，在网络中初步筛选出5 个景点，再 通过分析评论，记录评论中的关键词，为景点生成特点标签。  建立模型：用 Excel 或简单 Python 代码实现匹配逻辑。用户填写问卷后，系 统根据权重计算景点得分，筛选出符合条件的结果， 自动排除超出预算的项目、 餐厅、酒店，优先选择用户特别关注的的景点。  测试优化：邀请用户试用并反馈。若发现推荐景点交通不便，则调整模型， 增加“交通便利 ”的权重；若某些标签与实际不符（如某景点实际人少但评论说 人多），则重新核实数据。  **四、项 目意义**  通过这个项目，不仅能帮助用户解决选择困难，还能在实践中学习如何将数学 中的权重计算、数据分析应用于现实问题。未来，这一思路还可扩展至餐饮推荐、 商品选购等领域，甚至开发成手机小程序，让更多人享受个性化消费决策的便利。 |
| **第三志愿** | **一、项目背景**  在数字化零售快速发展的当下，零售行业每天都会产生海量的非结构化数据，包括商品详情页描述、商家发布的促销邮件、电商平台APP弹窗广告、社交媒体营销文案等。这些数据中蕴含着丰富的优惠信息，如满减活动、限时折扣、赠品福利等，但由于文本表述多样（如“买三免一”“满200立减60”等）、格式不统一，传统基于规则匹配的信息提取方法存在准确率低、泛化能力差的问题，难以高效处理复杂语义和长尾表达。与此同时，随着消费者需求日益个性化，零售企业亟需精准触达用户的推荐系统。传统推荐策略依赖固定规则或简单的协同过滤算法，无法充分利用动态的用户行为数据和实时场景信息（如位置、时段），导致推荐内容与用户兴趣匹配度低，影响消费转化率和用户体验。大语言模型（LLM）凭借强大的自然语言理解能力，为解决上述问题提供了新的技术路径，通过微调预训练模型，能够实现优惠信息的高效提取与个性化推荐。  **二、实现功能**  理解用户需求：收集商家和消费者的需求反馈，分析其在零售优惠信息处理方面的痛点和期望。 构建用户画像，收集用户的基本信息、消费行为、偏好等数据，以便更好地理解用  户对优惠信息的需求。  分析评价：对比传统规则匹配方法，评估模型在处理文本表达多样性方面的优势和不足。 通过实验和数据统计，分析模型在优惠信息提取和推荐中的准确性和效率，收集用户对推荐结果的满意度反馈。  **三、研究过程**  数据收集与预处理：确定需要收集的零售场景文本数据源，如商品描述、促销邮件、APP弹窗等。对收集的数据进行清洗，去除无关内容和噪声数据，标注优惠信息相关的关键词和  短语。  模型微调：选择 BERT、GPT 等预训练模型，确定微调模型的架构和参数。利用标注好的数据对模型进行微调训练，优化模型的损失函数，提高其对优惠信息  的理解能力。  信息提取与结构化：将微调后的模型应用于实际文本数据，提取折扣类型、有效期、适用商品等关键优惠信息。设计数据结构，将提取的优惠信息进行结构化存储，以便于后续查询和推荐。  推荐策略构建：基于用户画像，确定推荐的策略和算法，如协同过滤、深度学习模型等。根据用户的特征和行为，实时调整推荐策略，为不同用户提供高效推荐的优惠信息。  系统优化与迭代。  **四、项目意义**  提升零售企业运营效率：自动、精准提取优惠信息并结构化存储，减少人工处理成本，帮助企业快速整合促销资源，缩短活动上线周期，增强市场响应速度。  优化用户购物体验：基于用户画像和实时场景的个性化推荐，将合适的优惠信息精准推送给目标用户，降低用户筛选信息的时间成本，提升购物便捷性与满意度，增强用户粘性。  促进消费转化：通过智能推荐提高优惠信息的曝光精准度，激发用户购买意愿，直接带动销售额增长，为企业创造更高的商业价值。  推动技术创新与应用：探索大语言模型在零售场景的落地实践，验证微调技术在垂直领域的有效性，为NLP与推荐系统的结合提供案例参考，助力零售行业智能化升级。 |

1. **本项目团队优势**

*团队成员学生工作经历、科研经历、获奖情况、论文发表情况等*

|  |
| --- |
| 刘星宇：“树苗计划”数学竞赛二等奖  李宜珊：数学类 2404 班职业发展委员、数学与统计学院党建中心新媒体中心干事、东华大学网球协会社长  黄昱洲：智力助残志愿者社团干事、学生职业发展中心组织部成员  胡玮刚：数学与统计学院组织部部长、数学与统计学院党建组织中心干事、数学与统计学院易班干事  黄旭涵：数学与统计学院组织部部长、数梦织愿实践部部长、校易班文秘部干事、智力助残志愿者社团干事  于梦：数学与统计学院秘书部部长、数学与统计学院调研部干事、数学与统计学院第二届学生常任代表委员  郝尉辰：数梦织愿志愿实践部成员 |

1. **项目开展计划**

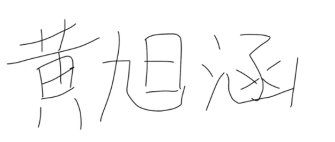
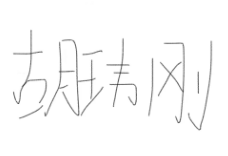
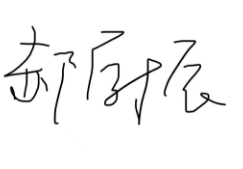
*项目开展计划时间表、预计项目成果等*

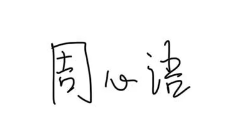
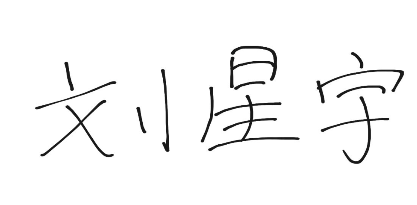
|  |
| --- |
| 项目开展计划时间表（以课题组成立算第一周）  第1周：项目启动与团队组建：课题与组老师取得联系，确定研究方向为基于非高斯噪声预测中国未来人口及其上海房价走势，并探究未来高校规模与高校教师失业人数估计。明确各成员分工，制定时间计划、表研究目标和预期成果。  第2-3周：人口与房价预测模型构建：回顾相关人口预测理论，结合非高斯噪声特点，构建人口预测模型。收集上海房价历史数据，分析影响房价因素，建立房价预测初步模型。  第4-5周：数据收集与整理：收集中国人口数据，包括出生率、死亡率、迁移等多维度数据。收集上海房价相关数据，如房价走势、交易量、供需情况等，对收集到的数据进行清洗、整理和预处理。  第6-7周：人口预测模型优化与验证：运用历史数据对构建的人口预测模型进行验证，分析预测误差，优化模型参数，提高模型的准确性。  第8-9周：房价预测模型优化与高校规模、教师失业研究：利用人口预测数据，结合上海房价影响因素，优化房价预测模型。开始收集高校规模数据，包括招生人数、毕业生人数、教师数量等，分析与教师失业人数的相关性。  第10-11周：综合分析与策略制定：通过人口和房价预测结果，分析相互影响关系。结合高校规模数据，估计未来高校教师失业人数，提出应对策略建议。  第12周：撰写研究报告与论文：整理项目研究成果，撰写研究报告初稿，审核修改报告，准备学术论文投稿。  第13周：总结与成果展示：组织总结会议，展示项目成果，完成课题总结报告，整理归档课题资料，准备参与相关学术竞赛和项目申报。 |

1. **承诺**

我们承诺：若成功入选，我们一定根据项目要求完成相关文献综述、论文、专利等，并参加国家级、上海市级大学生创新创业项目申报，且参加当年中国国际大学生创新大赛、“挑战杯”创新创业大赛等双创相关赛事。

**团队全体成员签字：（电子版请插入电子签名，纸质版请手签）**



*w*

*注：上表所填资料必须真实、完整、合法。*