

2023 年全国大学生数据分析大赛

某电商平台用户行为分析与挖掘

背景：电商是当今用户最大的交易市场之一，电商行业也逐渐成熟，所有市场中可售卖的商品全都在平台中存在，并且在网络和疫情的影响下，在线上的消费行为满足全年龄段用户。用户的交易行为推动了电商的发展。对于电商平台来说，通过用户的交易行为、交易金额、交易频率、交易时间等信息，分析用户的行为，可促进二次营销，精准营销，促活等。请根据提供的数据进行数据的分析与建模工作。

一、数据说明：

字段	注释
订单 ID	订单的唯一标识
商家 ID	商家的唯一标识
用户 ID	用户的唯一标识
付款日期	22 年 1 月-22 年 6 月
使用状态	1: 已使用该商家优惠券；0: 领取未使用；None: 未领取
实付金额	实际支付的金额，单位元
邮费	邮寄费用
省份	用户所在省份

城市	用户所在城市
数量	购买商品数量

二、任务：

1. 对电商平台用户的消费行为进行统计分析并可视化。（15 分）
2. 有这么一句话在业界广泛流传：“数据和特征决定了机器学习的上限，而模型和算法只是逼近这个上限而已”。特征工程是指利用领域知识从原始数据中提取有效特征，从而达到提高机器学习过程的结果的质量的目的。请根据用户的交易行为数据构造用户、商户和优惠券等有效特征。（20 分）
3. 用户画像是指根据用户的属性、用户偏好、用户行为等信息而抽象出来的标签化用户模型，通过给用户贴标签的方式来高度概括和描述用户，可以让人更直观地了解用户特征。请构建模型进行客户价值分析，并绘制用户画像分析用户特点。（25 分）
4. 平台推出不同商家的代金券，用于发放给全量用户。请根据用户的交易行为构建模型并训练，预测是否为该用户发放商家代金券，最后评估模型性能。（25 分）
5. 请设计一份优惠券投放策略。（15 分）