

数据挖掘与商务分析-第 2 次平时作业

姓名 + 学号

请于 5 月 12 日晚 10 点前提交

1 信用卡用户聚类

利用信用卡用户行为数据 (credit_card.csv, 字段描述见表1), 运用 K 均值方法, 开展用户聚类分析, 根据聚类结果制定营销策略。

表 1: 数据字段描述

字段	描述
CUSTID	信用卡持有人 ID
BALANCE	信用卡余额
BALANCEFREQUENCY	余额更新的频率, 0 到 1 之间的得分 (1= 频繁, 0= 不频繁)
PURCHASES	交易金额
ONEOFFPURCHASES	一次性购买的最大交易金额
INSTALLMENTSPURCHASES	分期购买的交易金额
CASHADVANCE	用户预先支付的现金
PURCHASESFREQUENCY	进行购买的频率, 得分在 0 到 1 之间 (1= 频繁, 0= 不频繁)
ONEOFFPURCHASESFREQUENCY	一次性购买的频率 (1= 频繁, 0= 不频繁)
PURCHASESINSTALLMENTSFREQUENCY	分期购买的频率 (1= 频繁, 0= 不频繁)
CASHADVANCEFREQUENCY	预先支付现金的频率
CASHADVANCECTR	使用“预先支付现金”进行的交易数量
PURCHASESTRX	进行的购买交易数量
CREDITLIMIT	用户的信用卡限额
PAYMENTS	用户还款金额
MINIMUM_PAYMENTS	用户最低还款金额
PRCFULLPAYMENT	用户还款占全额账单的百分比
TENURE	用户信用卡服务使用期限

请同时提交 *.ipynb 和 *.html 文件, 其中 *.html 是将 *.ipynb 导出后的版本。