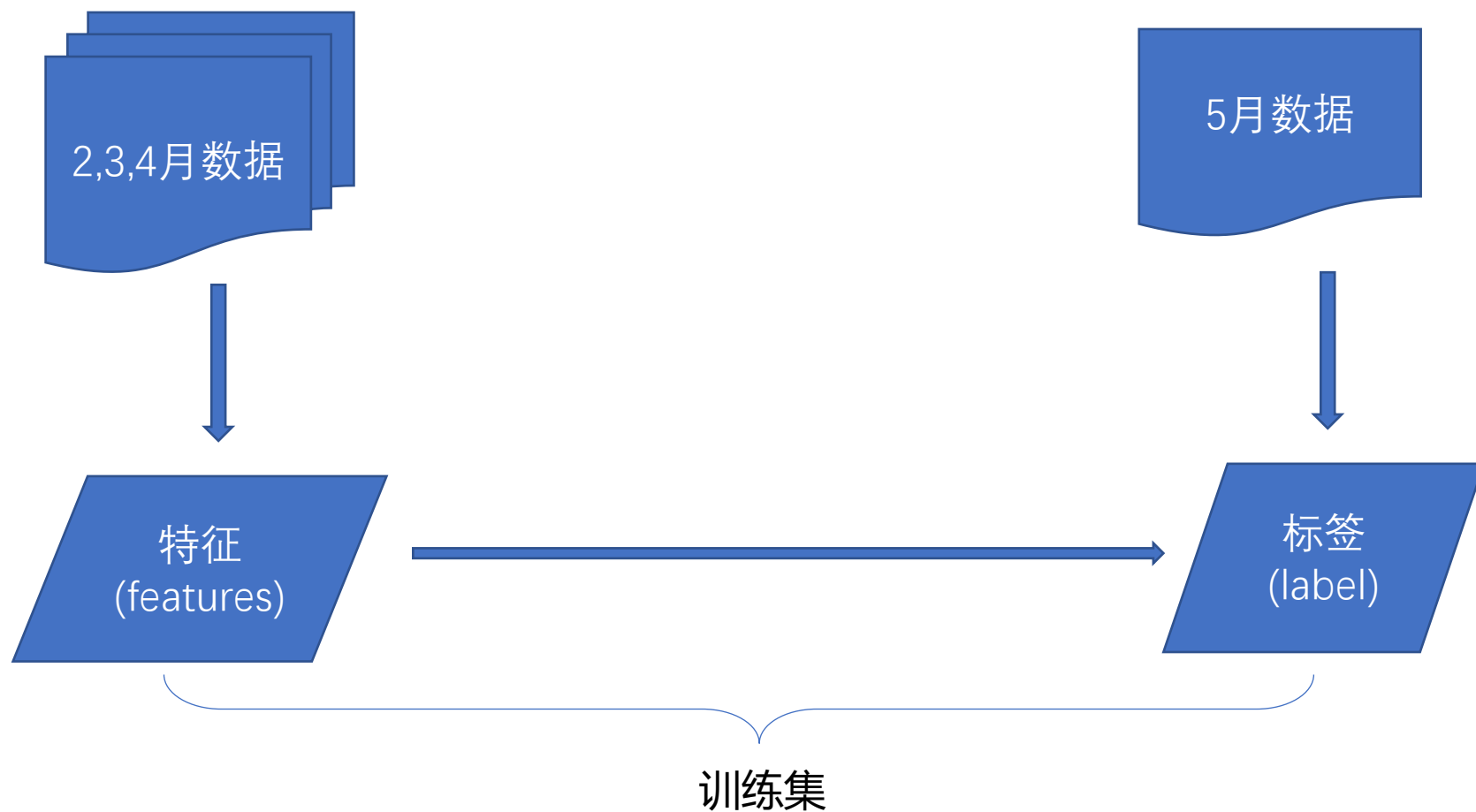


DAM2018

HW3-part2作业说明

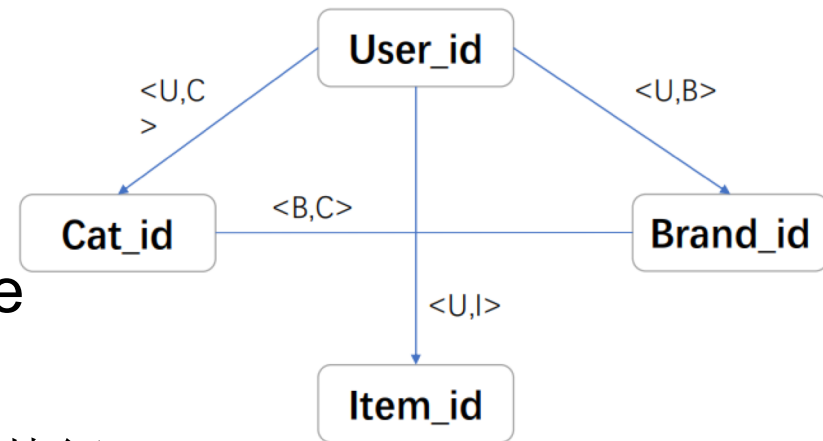
问题描述

- 训练过程示例



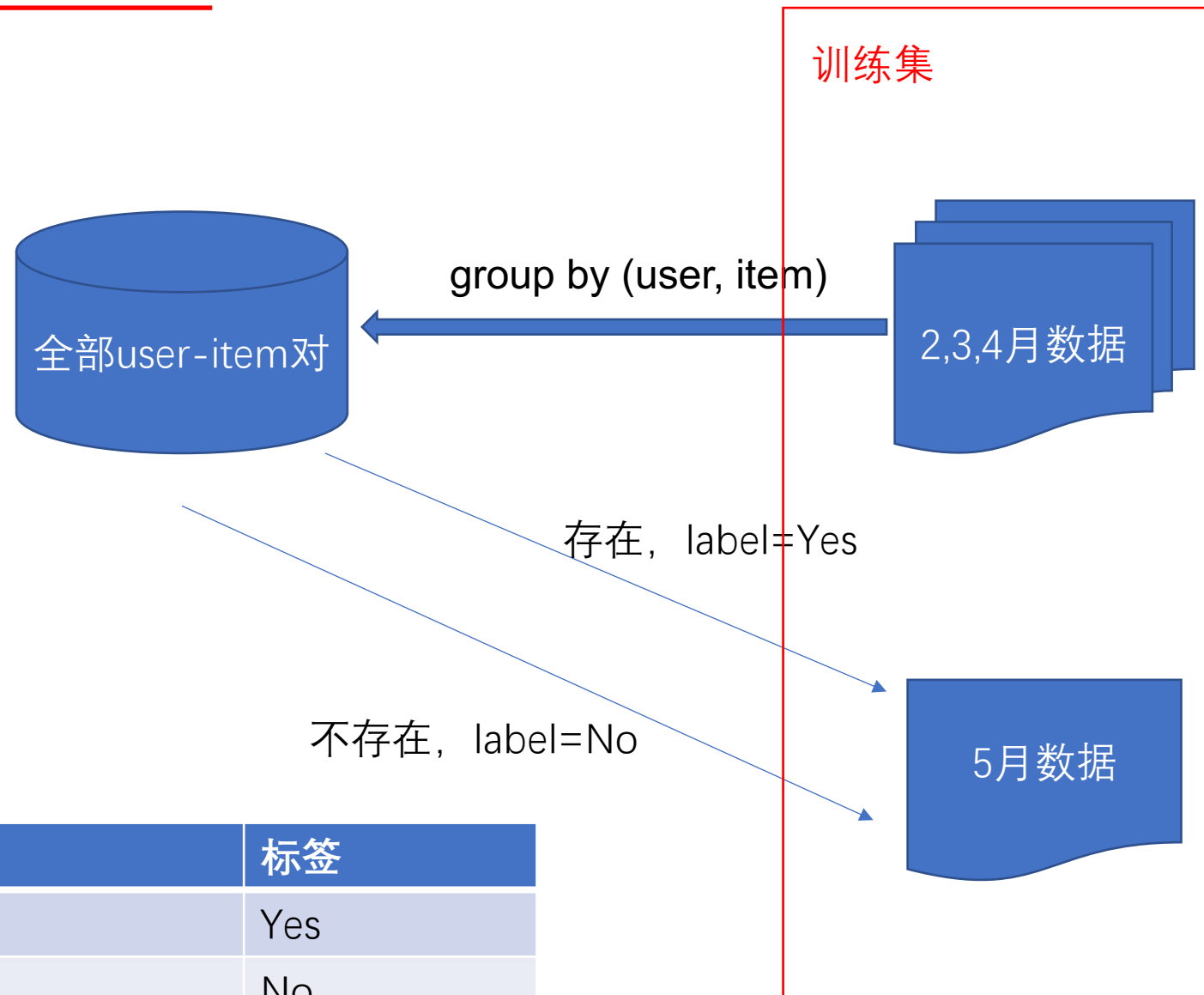
特征提取

- 参考论文：Repeat Buyer Prediction for E-Commerce
- 特征提取对象：
 1. 单个对象的特征：针对用户、品牌、类别、商品分别提取特征
 2. 联合特征：用户-品牌之间的交互特征（用户-类别、用户-商品...）
- 特征类型：
 1. 整个时间段简单统计（购买次数、购买天数、消费金额...）
 2. 每月统计（同上）
 3. 聚合特征：按月统计的特征进行聚合...
 4. 最近时间段内的特征（最近一个月、最近一周）
 5. 复杂特征（用户购买的趋势、用户-品牌间的相似度）
 6. 时间特征（用户购买周期、最近一次购买到待预测时间段的距离...）
- 可参考feature_ref.pdf



题目说明

b:
训练集特征数据中的所有
用户(vipno) – 商品(pluno)对
在训练集标签数据中是否出现



用户	商品	特征1	...	标签
vipno1	pluno1	feature1	...	Yes
vipno2	pluno2	feature1	...	No

题目说明.续

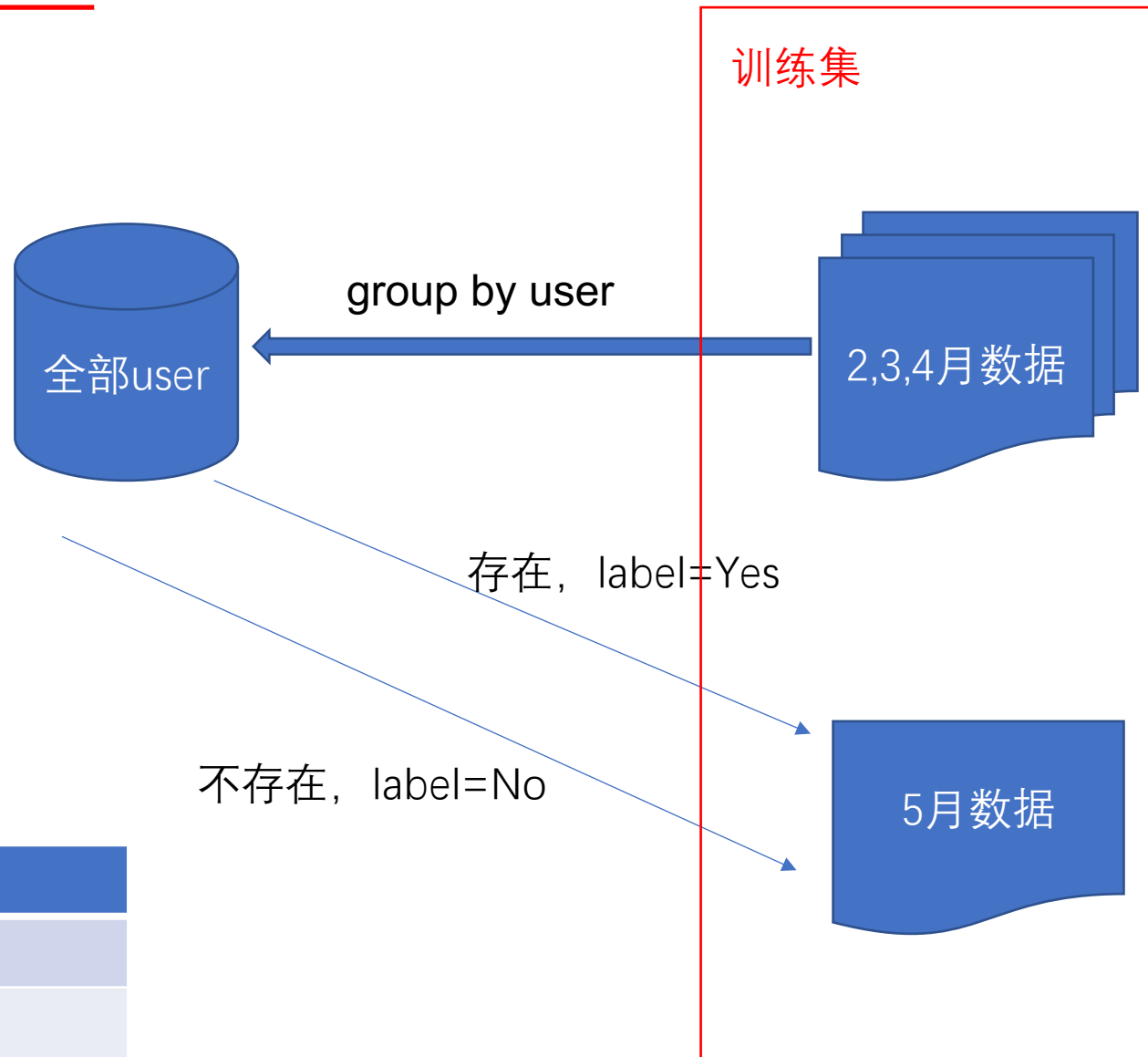
C-i:

训练集特征数据中的所有

用户(vipno)

在训练集标签数据中是否出现

用户	特征1	...	标签
vipno1	feature1	...	Yes
vipno2	feature1	...	No



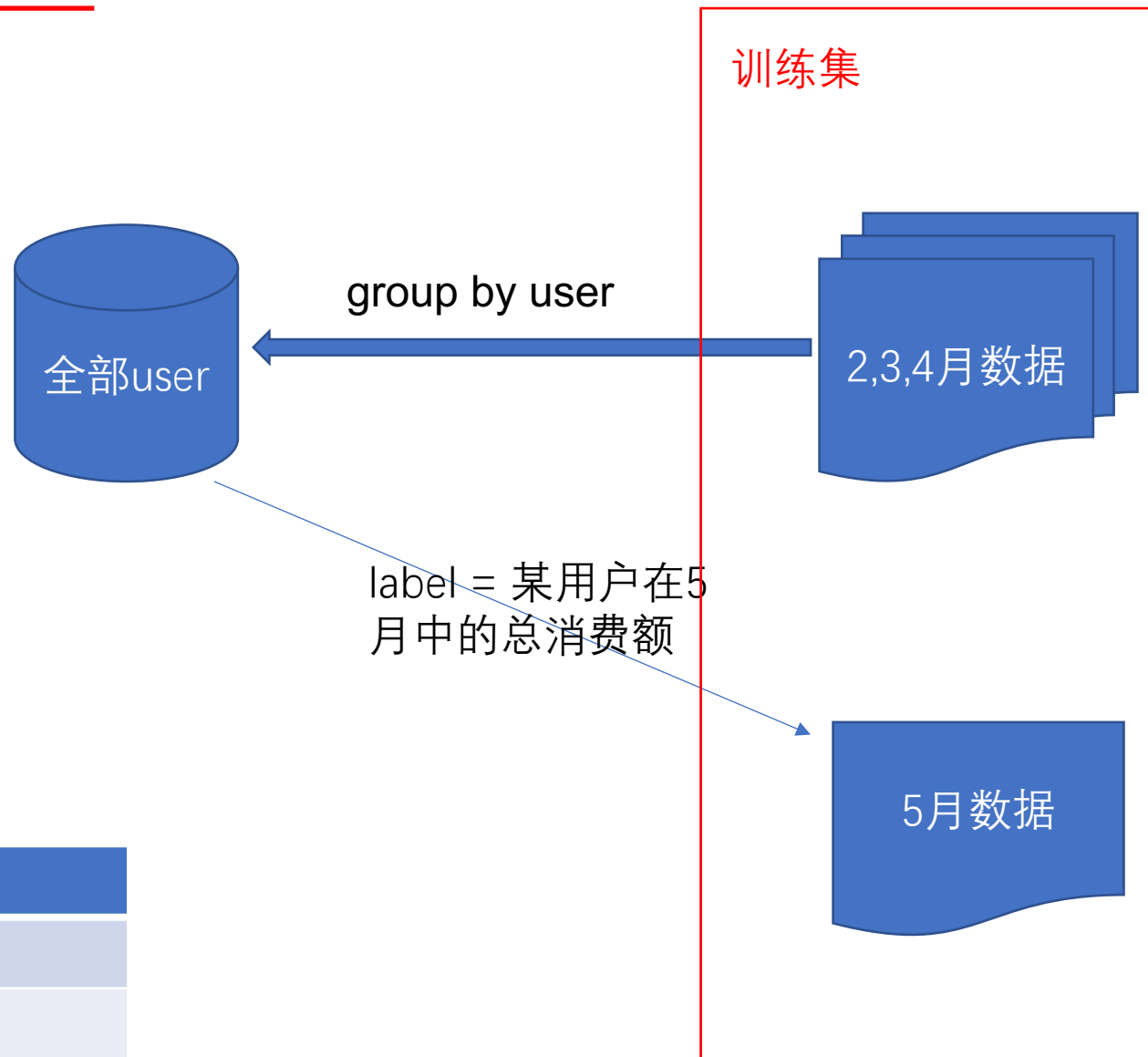
题目说明.续

c-ii & c-iii:

与b类似

c-iv: 对金额的回归预测

用户	特征1	...	标签
vipno1	feature1	...	100.0
vipno2	feature1	...	12.2



提交内容

- 代码、文档、预测结果
- 预测结果格式详见hw3中**第2题**的说明
- 预测结果请务必注意答案提交格式
- 最顶层文件夹和zip文件务必带上学号