

我们提供了一个商品网站中大约 16 万名用户在四年内对商品的评分数据, 每条评分记录都有时间戳(隐匿了具体时间, 只保证顺序不变)。

评分分为 5 级, 1 分最低, 5 分最高。

我们抽取了超过 3400 万条评分记录, 作为训练集, 数据文件名为 train.csv,

字段格式为: uid,iid,score,time 用户 i, 商品 a, 评分, 相对时间

用户 j, 商品 b, 评分, 相对时间

说明: 1) 第一行为表头, 表头的格式为: uid,iid,score,time, 四个字段分别代表: 用户编号, 商品编号, 评分, 相对时间;

2) 每一行为一个用户对一个商品的评分, 行之间用“回车符”分隔;

3) 每一行各字段之间用“逗号”分隔。

我们还抽取了近 55 万条评分记录, 作为测试集。我们隐藏了用户对于商品的评分, 仅保留用户和商品的评分关系, 数据文件名为 test.csv,

字段格式为: uid,iid 用户 k, 商品 c 用户 l, 商品 d

说明信息同训练集 (train.csv)。