

# 基于天猫交易数据的分析与挖掘\*

王希梅

清华大学 软件学院

北京 100084

wxm17@mails.tsinghua.edu.cn

于千山

清华大学 软件学院

北京 100084

yqs17@mails.tsinghua.edu.cn

陳善宇

清华大学 软件学院

北京 100084

email

## 摘要

描述所提出的问题。

## 关键词

数据挖掘，聚类分析，行为预测

## 1. 问题描述

详细描述所提出的问题。

阐述问题的合理性或价值。

本文组织形式如下，第。。章介绍。。

## 2. 方法设计

我们考虑对如下问题进行方法设计：

1. 商品平凡模式挖掘；
2. 同一商家用户的分类；
3. 用户重复购买的预测。

Todo: 实验环境

### 2.1 商品平凡模式挖掘

我们使用 FPGrowth 算法进行商品关联规则挖掘，

该算法...Todo: 算法 review

数据...Todo: 描述使用到的数据

\*具体数据集信息参见 <https://tianchi.aliyun.com/datalab/d-ataSet.htm?id=5>

### 2.2 同一商家用户的分类

...

### 2.3 用户重复购买的预测

我们提取了 `\user_gender` 等... 个特征，...

Todo: 特征工程

Todo: 算法框架

Todo: 模型、融合

### 2.4 第四个问题？

Todo: 思考第四个问题 或者直接从前面归结出一个问题

## 3. 结果与分析

## 4. 总结

本文我们用。。。数据，。。。方法，分析了。。。问题，得到了。。。结果

## 5. 参考文献

- [1] M. Bowman, S. K. Debray, and L. L. Peterson. Reasoning about naming systems. *ACM Trans. Program. Lang. Syst.*, 15(5):795–825, November 1993.
- [2] J. Braams. Babel, a multilingual style-option system for use with latex's standard document styles. *TUGboat*, 12(2):291–301, June 1991.
- [3] M. Clark. Post congress tristesse. In *TeX90 Conference Proceedings*, pages 84–89. TeX Users Group, March 1991.
- [4] M. Herlihy. A methodology for implementing highly concurrent data objects. *ACM Trans. Program. Lang. Syst.*, 15(5):745–770, November 1993.
- [5] L. Lamport. *LaTeX User's Guide and Document Reference Manual*. Addison-Wesley Publishing Company, Reading, Massachusetts, 1986.
- [6] S. Salas and E. Hille. *Calculus: One and Several Variable*. John Wiley and Sons, New York, 1978.