

## 2.3【A03】网络零售平台异常商品（价格、销量）识别【浪潮卓数】

### 1. 命题方向

智能计算

### 2. 题目类别

应用类

### 3. 题目名称

网络零售平台异常商品（价格、销量）识别

### 4. 背景说明

#### 【整体背景】

近年来，互联网技术发展迅速，电子商务行业也踏上了发展的快车道。“网购”因其方便快捷、省时省力、送货上门的特点越来越受到人们的青睐。在各平台规模不断扩大、商品数不断增加的同时，一些不正当的经营行为，例如虚标价格、刷单行为也随之出现，严重违反了电商法，需要对这类商品数据进行准确识别。针对如此庞大的商品数量，如果单纯人工检查筛选，不仅工作量巨大，还会出现遗漏和错误的情况。本赛题旨在寻找一种识别方法，能够实现对异常商品的准确定位，减少人工干预成本和降低出错率。

#### 【公司背景】

浪潮卓数大数据产业发展有限公司作为浪潮集团旗下的大数据板块，致力于成为数据资源提供商、数据资产运营商和数据交易服务商，以大数据时代的“数商”为发展目标，促进数据社会化。网络零售分析及电子商务监测，是当前卓数大数据公司的优势板块，其拥有丰富的数据量和案例，可供相关机构和学院进行实训。

#### 【业务背景】

浪潮卓数大数据产业发展有限公司多年来深耕电商大数据行业，作为行业内领先的数据服务提供商，浪潮卓数有责任为市场提供及时有效、高质量的电商数据。通过发布该命题，期待看到一种新颖的异常商品识别方法，能够在较短时间内准确定位异常商品，极大的提高数据产出质量。

### 5. 项目说明

### 【问题说明】

对网上零售平台提供的商品，利用其提供的商品价格、销量、类目、商品参数、店铺等信息，对这些商品标注其中价格异常和销量异常的商品。价格异常主要是标价与商品实际价值相差过大等，可以参考但不限于离群值等，在不同细分类中（如店铺、品牌、类目）中，离散程度不同。销量异常主要为商品的销量与实际的成交量有明显差距，可能是店铺存在刷单行为，或者其他修改展示销量的行为，可考虑商品销售稳定性与店铺经营稳定性等。针对存在的标价与销量与实际情况不符的现象，旨在建立一种异常商品识别模型，对给定的商品数据，进行异常标注（价格异常、销量异常）。

### 【用户期望】

追求异常商品识别模型的准确性与识别模型对大规模数据应用的高效性。

## 6. 任务要求

### 【开发说明】

识别异常存在生活中的方方面面，在及时预警方面表现的尤为突出。电商行业积累了大量的数据，识别其中的异常备受关注。本次拟定给定参赛者电商经营商品数据，旨在建立合适的异常商品识别模型，对其中的标价异常和销量异常的商品进行标注识别。

### 【技术要求与指标】

在此项目中，要求详细阐述识别模型使用的数据指标以及衍生指标，建立识别异常商品模型的逻辑思路，建模过程中对识别异常商品模型涉及到的参数如何调优，最终模型的效率（识别给定商品中异常商品所需时间）等。

主要评价指标：

- （1）异常商品识别比例（识别出的异常商品数/异常商品总数）；
- （2）识别正确率（异常商品识别正确率）；
- （3）模型或方法的时间复杂度和空间复杂度。

### 【提交材料】

- （1）项目概要介绍；
- （2）项目详细方案；
- （3）项目演示视频；

- (4) 项目简介 PPT;
- (5) 企业要求提交材料;
- ①异常商品识别模型建立的方法原理;
- ②异常商品标注(标注哪些商品为异常商品);
- ③建立后的模型,模型的应用效率(使用何种设备,执行商品标注所需时间);
- ④可视化 web 端开发,展示异常商品分布情况与特征描述;
- 6) 团队自愿提交的其他补充材料。

### 【任务清单】

- (1) 模型的概述与简介, 应对问题, 模型建立的思想逻辑;
- (2) 模型的参数调优过程(如果包含调优过程);
- (3)模型的效率与准确率(包括处理文本的效率与识别异常商品的正确率)。

### 【开发工具与数据接口】

- (1) 开发工具: 开发工具以及开发平台不限, 可以借助开源的工具;
- (2) 数据接口: 企业会通过网盘提供。

## 7. 其他



赛题介绍说明.doc

x

## 8. 参考信息

Python、java 等开发工具。

## 9. 评分要点

赛题评分要点见附件四: A 类初决赛阶段评分表。