

基于稳定度的异常商品识别模型 研究与实现

马飞燕¹, 赵李娟¹, 林世明^{2*}, 李芳^{1*}

¹昌吉学院信息工程学院, 新疆 昌吉

²厦门大学信息学院(国家示范性软件学院)电子信息国家级实验教学示范中心, 福建 厦门

收稿日期: 2023年6月1日; 录用日期: 2023年7月1日; 发布日期: 2023年7月11日

摘要

随着经济的发展, 网购现已得到全方位的普及, 因其有方便快捷、省时省力、送货上门等优点, 越来越受到人们的青睐, 成为日常生活中不可或缺的一部分。随着人们经济能力、消费水平的提高, 对网购体验的需求也愈发上涨。同时网上各大零售业务间的竞争也愈发激烈, 为了能够吸引消费者的目光以增加商品的销量, 某些商家开始采用刷销量、刷好评、删差评等“炒信”“刷单”手段对商品进行恶意推广, 侵犯消费者权益。为保障消费者的知情权和选择权, 本项目通过浪潮卓数公司提供的数据集, 通过数据挖掘定量分析和定性分析结合的方法来剖析商品出现异常的原因, 采用数学建模和机器学习的方法, 定义部分异常商品指标, 并利用这些指标构建出查找和预测异常商品的模型。实验结果表明该模型效果较好, 具有一定的实用性。

关键词

异常商品, 数据挖掘, 稳定度, 数学建模

Research and Implementation of Abnormal Product Identification Model Based on Stability

Feiyan Ma¹, Lijuan Zhao¹, Shiming Lin^{2*}, Fang Li^{1*}

¹School of Information Engineering, Changji University, Changji Xinjiang

²National Demonstration Center for Experimental Electronic Information Education, School of Informatics
Xiamen University (National Demonstrative Software School), Xiamen Fujian

Received: Jun. 1st, 2023; accepted: Jul. 1st, 2023; published: Jul. 11th, 2023

*通讯作者。

文章引用: 马飞燕, 赵李娟, 林世明, 李芳. 基于稳定度的异常商品识别模型研究与实现[J]. 数据挖掘, 2023, 13(3): 244-253. DOI: 10.12677/hjdm.2023.133025

Abstract

With the development of the economy, online shopping has gained widespread popularity in all aspects. Due to its advantages such as convenience, speed, time and effort saving, and door-to-door delivery, it is increasingly favored by people and has become an indispensable part of daily life. With the improvement of people's economic ability and consumption level, the demand for online shopping experience is also increasing. At the same time, competition among major online retail businesses has become increasingly fierce. In order to attract consumers' attention and increase product sales, some businesses have started to use "speculation" and "order" methods such as selling, positive reviews, and negative reviews to maliciously promote products, infringing on consumers' rights and interests. To protect consumers' right to know and choose, this project uses a dataset provided by Inspur Zhuosu Company to analyze the reasons for abnormal products through a combination of quantitative and qualitative data mining analysis. Mathematical modeling and machine learning methods are used to define some abnormal product indicators, and these indicators are used to construct a model for finding and predicting abnormal products. The experimental results indicate that the model has good performance and certain practicality.

Keywords

Abnormal Product, Datamining, Stability, Mathematical Modeling

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着经济的发展,网购现已得到全方位的普及,成为了我们日常生活中不可或缺的一部分,因为其有方便快捷、省时省力、送货上门的各项优点,已经越来越受到人们的青睐,并随着人们经济能力、消费水平的提高,对网购体验的需求也愈发上涨,网上各大零售业务间的竞争也愈发激烈,在各平台规模不断扩大、商品种类数不断增加的同时。为了能够吸引消费者的目光并增加商品的销量,某些商家开始采用虚假交易数据的竞争手段进行对商品的推广,最为典型的经营行为有:虚标价格,虚假销量,刷好评[1]。本文通过网络营销公司提供的商品数据以及店铺数据,通过定量分析和定性分析结合的方法来剖析商品出现异常的数据原因,定义部分数据作为衡量商品是否出现异常的指标,并且利用这些指标来创建数学模型对和效用函数构建出查找异常商品的数字模型,能够从大量的商品数据中识别出属于商品价格或者销量异常的商品,为网络销售平台提供准确的目标商品[2] [3]。

2. 问题分析

通常情况下,电商平台的各类商家会采用刷单的行为来提高竞争力。本文使用的数据集是某电商平台收集的海量交易商品数据,总数据分为商品数据和店铺数据,商品数据共有 21 个字段(主要的两个字段为:价格字段 ITEM_PRICE,销量字段 ITEM_SALES_VOLUME),共有 4 个月份数据,分别为 2021 年 6 月(7,677,206 例)、7 月(4,133,204 例)、8 月(4,861,351 例)、9 月份(4,771,652 例)数据,按照 DATA_MONTH

字段区分, 店铺数据和商品数据之间通过 USER_ID 字段进行关联, 采用数学建模的方式, 从全部商品数据、店铺数据中识别出可能的异常商品, 为网络零售平台查找异常商品提供目标对象[4] [5]。

3. 模型建立

3.1. 建模思路

除通过挖掘异常价格、销量的相关指标, 以建模的手段将所有的指标进行融合应用, 最终得出识别异常商品的模型, 再经过调优的操作进行效果评估, 最终得出异常商品。技术主要采用 numpy 和 pandas 模块, 对百万级别的商品数据进行分析 and 建模, 对于构建的模型需要结合人工观测多次调优, 确保模型的可行性能贴合实际应用[6]。

3.2. 特征工程

3.2.1. 价格异常特征

通过分析数据和查阅相关资料, 我们发现价格异常通常具有价格波动较大, 稳定性较差, 在同品牌同一类目中具有相对差值较大等特征, 且与店铺的好评度、开店时长等具有一定的相关性。根据现有数据, 我们构建了如下指标:

1) 商品信息缺失率

将商品特征缺失较多的归类为异常商品, 因数据格式的原因, 部分商品特征无法读取缺失, 所以对该类特征缺失的部分进行替换的操作, 以特定的值替换缺失区域, 再用该特定值代替“NaN”, 以缺失的特征个数与总特征个数的比值作为指标来评价异常商品[7]。

命名为 loss_rate:

$$\frac{\text{显示 'NaN' 的特征个数} + \text{特定值所属特征的出现个数}}{\text{总特征个数}(15)} \quad (1)$$

2) 相邻两月的共同商品数据的价格稳定度

两个相邻的商品数据中会存在共同的商品, 此类商品用两个月的价格差与较小的一个价格的比值表示这两个月的价格波动程度来作为稳定度[8]。波动程度较大的一般为异常商品($P_{i\text{月份}}$ 、 $P_{i-1\text{月份}}$ 表示相邻两个月的价格, $(P_i, P_{i-1})_{\min}$ 表示相邻两个月份价格中较小的一个月的价格)。

命名为 fluc_rate_mn (mn 表示相邻的两个月份数), 模型如下:

$$\frac{|P_{i\text{月份}} - P_{i-1\text{月份}}|}{(P_i, P_{i-1})_{\min}} \quad (2)$$

3) 四个月共同出现的商品价格稳定度

四个月份共同出现的商品, 单个价格和四个价格均值的差值的绝对值数与四个价格的均值之比来表示该商品在四个月中的稳定度指标(P_i 表示某一商品价格, P_{mean_4} 表示该商品四个月的均价)。

命名为 price_volatility_0n (n 表示月份数), 模型如下:

$$\frac{|P_i - P_{\text{mean}_4}|}{P_{\text{mean}_4}} \quad (3)$$

4) 四个月共同出现的商品价格指标稳定度

四个月共同出现的商品, 以其每个月的单个价格与四个月的均价差的平方和的二次方根值为自变量, 进行 sigmod 化, 将两端的峰值进行归一化处理, 以此来表示稳定度[9] (P_6, P_7, P_8, P_9 表示商品所属月

份对应的价格， P_{mean_4} 表示同一商品四个月的均价)。

命名为 price_stab_e，模型如下：

$$\frac{1}{1 + e^{-\sqrt{(P_6 - P_{mean_4})^2 + (P_7 - P_{mean_4})^2 + (P_8 - P_{mean_4})^2 + (P_9 - P_{mean_4})^2}}} \quad (4)$$

5) 商品单价在同类商品均价中的稳定度

同类商品的价格波动程度较其他维度的价格波动程度更为准确，选取缺失率最低的三级类目为分类指标，以偏离同类商品均价的程度来表示商品价格的异常程度，用计算标准差中每个商品价格的占比大小来表示该商品偏离均价的高低[10] (P_i 商品价格， P_{mean_1} 同类商品价格均值)。

命名为 price_offset_0n (n 表示月份数)，模型如下：

$$\frac{\sqrt{(P_i - P_{mean_1})^2}}{\sqrt{\sum_{i=1}^n \frac{(P_i - P_{mean_1})^2}{n}}} \quad (5)$$

6) 补差价商品价格的稳定度

日常生活中，网购时补差价数值一般不会很高，如果补差价值很高，可视为异常商品[11] (P_i 表示补差价的商品价格， P_{mean} 表示该类商品的平均价格， n 表示商品数量)。

命名为 price_difference_deviate，模型如下：

$$\frac{\sqrt{(P - P_{mean})^2}}{\sqrt{\sum_{i=1}^n \frac{(P_i - P_{mean})^2}{n}}} \quad (6)$$

7) 店铺的评分评价店铺的优劣

以店铺数据中的商品描述得分、物流得分、服务得分占总分值的比例来评价店铺的优劣(S_1 商品描述得分， S_2 服务得分， S_3 物流得分)。

命名为 score，模型如下：

$$\frac{S_1 + S_2 + S_3}{S_{总}} \quad (7)$$

3.2.2. 销量异常特征

通过数据分析和查阅资料，我们发现销量异常具有销量通常较大，不同月份之间销量的波动较大、稳定性差，普通店销量高于旗舰店销量，存在刷单刷好评情况等。因此构建如下指标：

1) 绝对销量

为了降低销量在高点的斜率，以销量得二次方根值来表示销量得绝对销量值(S_i 表示某商品的销量)

命名为 abs_sales，模型如下：

$$\sqrt{S_i} \quad (8)$$

2) 两个相邻月份的商品数据的销量稳定度

两个相邻的商品数据中会存在共同的商品，此类商品用两个月的销量差与前一个月份的销量的比值

表示这两个月的销量波动程度。波动程度较大的一般为异常商品($S_{i\text{月份}}$ 、 $S_{i-1\text{月份}}$ 表示某商品相邻两个月的销量)。

命名为 sales_fluc_mn (m, n 表示相邻月份数), 模型如下:

$$\frac{S_{i\text{月份}} - S_{i-1\text{月份}}}{S_{i-1\text{月份}}} \quad (9)$$

3) 四个月共同出现的商品销量稳定度

四月份共同出现的商品, 其中之一的销量和四个销量均值的差值与四个销量的均值之比作为稳定度的指标(S_i 表示某商品的销量, S_{mean_4} 表示该商品四个月销量的均值)

命名为 sales_volatility_0n (n 代表月份数), 模型如下:

$$\frac{S_i - S_{\text{mean}_4}}{S_{\text{mean}_4}} \quad (10)$$

4) 四个月共同出现的商品销量波动稳定度

根据幂函数的函数特征, 指数越低, 函数值越靠近 x 轴, 适合波动程度较大的自变量的处理, 四个月共同出现的商品, 以其销量十分之一次方来将销量的波动程度降低, 用以表示商品的销量稳定程度(S_i 表示商品某一月的销量, S_{mean_4} 表示该商品四个月销量的均值, n 表示月份数)。

命名为 sales_stab_e, 模型如下:

$$\sqrt[20]{\frac{\sum_{i=1}^n (S_i - S_{\text{mean}_4})^2}{4}} \quad (11)$$

5) 商品销量在同类目商品销量中的稳定度

以三级类目为分类参照, 同类商品的均销量与单个商品的销量差再与同类销量的比值代表这一稳定度的指标(S_i 表示商品的销量, $S_{i\text{mean}_lv3}$ 表示同类商品的均销量)。

命名为 sales_differ_0n (n 表示月份数), 模型如下:

$$\frac{S_i - S_{i\text{mean}_lv3}}{S_{i\text{mean}_lv3}} \quad (12)$$

6) 评销比

评论数与销量的比值为评销比, 代表每件商品的评论数, 如果销量数据是刷单刷出来的, 那该商品的评销比值一定异常, 但因为评论数采用历史沿用数据, 所以只能用开店时长算出商品的月均评论数来计算评销比, 计算出的评销比波动较大不宜观察及总结, 所以我们采用 sigmoid 函数处理(其中 $T_{\text{评论}}$ 表示的是销量的数量, $time_{\text{月份}}$ 表示的是开店时长的月份数) [12]。

命名为 ev, 模型如下:

$$\text{eval/sales} = \frac{T_{\text{评论}}/time_{\text{月份}}}{S}$$

$$\text{ev} = \frac{1}{1 + e^{-\text{eval/sales}}} \quad (13)$$

7) 与官方旗舰店的销量差

消费者在选择同品牌的商品时, 考虑到真假的问题, 更多会优先选择旗舰店的商品, 所以一般旗舰

店的商品较普通店铺卖的更好，但商品的种类不同，所对应的销量也各不相同，所以我们最终采用同品牌的商品中销量最高的值为参照，以同品牌普通店铺的商品销量与之做对比作为一个评价商品异常的指标[13] (S 表示销量)。

命名为 differ，模型如下：

$$\frac{S - S_{\text{同品牌官方旗舰店的最大值}}}{S_{\text{同品牌官方旗舰店的最大值}}} \quad (14)$$

3.2.3. 特征归一化

所有评判商品异常的指标计算出来以后，应对最终模型的需求，需要将所有指标进行归一化处理[14] (其中 x 表示所需要的任何一个指标， x_{\min} 表示所选指标的最小值， x_{\max} 表示所选用的指标的最大值)。

$$\frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (15)$$

3.2.4. 特征指标总结

所有评判商品异常的指标如下表 1 和表 2 所示。

Table 1. Price anomaly judging characteristics

表 1. 价格异常评判特征

特征	含义
loss_rate	商品信息缺失率
fluc_rate_mn	相邻两月的共同商品数据的价格稳定度
price_volatility_0n	四个月共同出现的商品价格稳定度
price_stab_e	四个月共同出现的商品价格指标稳定度
price_offset_0n	商品单价在同类商品均价中的稳定度
price_diference_deviat	补差价商品价格的稳定度
score	店铺的评分评价店铺的优劣

Table 2. Sales anomaly judging characteristics

表 2. 销量异常评判特征

特征	含义
abs_sales	绝对销量
sales_fluc_mn	两个相邻月份的商品数据的销量稳定度
sales_volatility_0n	四个月共同出现的商品销量稳定度
sales_stab_e	四个月共同出现的商品销量波动稳定度
sales_differ_0n	商品销量在同类目商品销量中的稳定度
ev	评销比
differ	与官方旗舰店的销量差

3.3. 模型预测与评判

3.3.1. 预测模型

预测模型包括价格异常预测模型和销量异常预测模型。价格异常预测模型将同类商品的价格稳定度、

补差价商品的价格稳定度、多次出现商品的价格稳定度、相邻月份共同出现商品的价格稳定度进行线性融合,不同的稳定度指标对结果的贡献度不同,因此通过权重因子 W 进行调节。同时考虑到一般情况下开店时长越长、店铺评分越高出现价格异常的情况应该是越少。当然存在店铺刷分的情况,因此利用指数模型进行调节。一般情况下商品的信息缺失率高,商品出现异常的情况会增加。最终得出价格异常商品的模型如公式(16)。

销量异常预测模型将绝对销量值、相邻月共同商品的销量波动程度、多次出现商品的销量波动程度进行线性叠加,同时设置调节因子 W_1 ,将缺失率、评销比、销量稳定程度、销量的对数值进行乘性叠加。同样的,一般情况下开店时长越长、店铺评分越高销量出现异常的情况应该是越少,因此将其作为分母。最终得到销量异常的预测模型如公式(17)。

$$x = W_1 \times \text{price_offset_0n} + W_2 \times \text{price_difference_deviate} + W_3 \times \text{price_volatility_0n} + W_4 \times \text{fluc_rate_mn}$$

$$Y = \frac{x \times W_5 \times \text{loss_rate}}{a^{\text{score}} \times \ln(\text{duration_shop} + b)} \quad (16)$$

$$x = W_1 \times (\text{abs_sales} + \text{sales_fluc_mn} + \text{sales_volatility_0n}) \times W_2 \times \text{loss_rate} \times W_3 \times \text{ev} \times W_4 \times \text{sales_stab_e} \times \log_{W_5}(\text{ITEM_SALES_VOLUME})$$

$$Y = \frac{x}{a^{\text{score}} \times \ln(\text{duration_shop} + b)} \quad (17)$$

3.3.2. 参数选择和模型评价

对筛选后的价格异常商品进行按照总样本的 2% 进行抽样,与日常生活中类似的网络交易商品价格比对其进行人工添加标签确认作为验证集。模型测试多组不同权重值计算出来的异常概率值,依据 F1 分数作为评判标准选出最优的权重组合,之所以选择 F1 分数作为评判标准是因为 F1 分数作为准确率和召回率的加权平均值,可同时考虑准确率和召回率,以便重新计算新的分数, F1 分数计算[15] (18):

(TP : 真阳性, FP : 假阳性, TN : 真阴性, FN : 假阴性)

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (18)$$

价格异常分数计算所需指标的权重最优组合是通过遍历网格(表 3)最终得出 F1 分数最高的组合 $W_{et_comb_1}$ (如图 1),计算最优权重组合的确认后的标签与预测标签计算准确率与召回率:准确率为 0.90,召回率为 0.88。

对筛选后的销量异常商品进行按照总样本的 1% 进行抽样,与电商平台的同类商品实际观测到的数据对比再次确认标签作为验证集,模型测试多组不同权重值计算出来的异常概率值,同样依据 F1 分数作为评判标准选出最优的权重组合,价格异常分数计算所需指标的权重最优组合是通过遍历网格(表 4)最终得出 F1 分数最高的组合 $W_{et_comb_1}$ (如图 2),计算最优权重组合的预测标签与确认后的标签的准确率与召回率[16]:准确率为 0.874,召回率为 0.875。

Table 3. Price index weight test grid
表 3. 价格指标权重测试网格

组合	指标权重					a	b
	W_1	W_2	W_3	W_4	W_5		
Wet_comb_1	0.2	2	0.5	9	1	2	2
Wet_comb_2	0.3	3	1	10	3	1	1
Wet_comb_3	0.4	4	1.5	11	5	0.8	0
Wet_comb_4	0.5	5	2	12	7	0.6	0.9
Wet_comb_5	0.6	6	2.	13	9	0.4	0.8
Wet_comb_6	0.7	7	3	14	11	0.2	0.7
Wet_comb_7	0.8	8	3.5	15	13	0.1	0.6
Wet_comb_8	0.9	9	4	16	15	0.08	0.5
Wet_comb_9	1.0	10	4.5	17	17	0.06	0.4
Wet_comb_10	1.2	11	5	18	19	0.04	0.3

Table 4. Sales index weight test grid
表 4. 销量指标权重测试网格

组合	指标权重					a	b
	W_1	W_2	W_3	W_4	W_5		
Wet_comb_1	1	1	1	1	2	2	2
Wet_comb_2	0.8	0.8	1.2	2	2	5	6
Wet_comb_3	0.6	0.8	1.2	2	2	5	5
Wet_comb_4	0.4	0.6	0.8	4	2	4	4
Wet_comb_5	0.4	0.4	0.8	5	3	4	3
Wet_comb_6	0.2	0.2	0.6	6	3	3	1
Wet_comb_7	0.2	1	0.4	7	3	6	0.8
Wet_comb_8	1.2	1.2	0.2	8	4	7	0.6
Wet_comb_9	1.4	1.4	1.4	1.2	5	8	0.4
Wet_comb_10	1	1.6	1.6	1.4	5	9	0.2

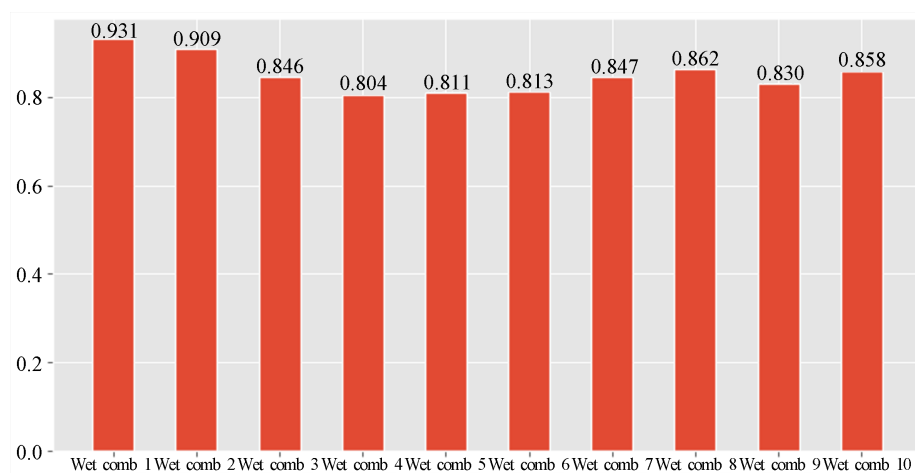


Figure 1. Price anomaly indicator weight recombination F1 value

图 1. 价格异常指标权重重组 F1 值

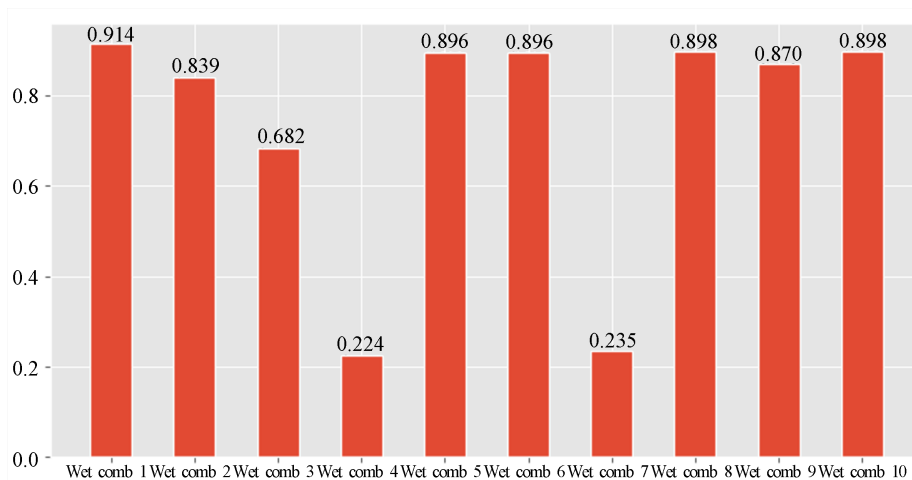


Figure 2. Sales exception index weight reorganization F1 value

图 2. 销量异常指标权重重组 F1 值

4. 结果分析

经过模型的筛选,从总样本千万的数据集中筛选出 6 月价格异常 4365 例,销量异常 12,569 例,7 月价格异常 2350 例,销量异常 6767 例,8 月价格异常 2764 例,销量异常 7959 例,9 月价格异常 2713 例,销量异常 7813 例,异常的总占比为 2.2%,五金/工具类的商品更容易出现异常的情况,电子元件或电子产品的异常占比也相对较高,这一种类的商品更容易有价格或者销量的波动,所以商家利用这一特征钻空子,因此需要重视该类问题。

通过计算模型的准确性可以看出,基于稳定度的异常商品(价格、销量)识别模型拥有较高的识别能力,对于复杂的数据集也能合理的应对,能够从海量的商品数据中快速的计算出商品异常的概率,辅助相关电商监管机构准确、快速的完成筛选处理。需要注意的是该模型需要足够的历史数据来辅助模型的调优与测试,确保其可扩展性和可维护性,以便将来随着数据量的增长进行优化和改进。

基金项目

2020 年新疆维吾尔自治区自然科学基金创新环境建设专项(2020D01C001),新疆维吾尔自治区高等学校科学研究计划资助(XJEDU2023P127)。

参考文献

- [1] 李士琪. 网络消费者权益保护的政府监管研究——以上海市静安区大宁路街道为例[D]: [硕士学位论文]. 上海: 华东政法大学, 2022. <https://doi.org/10.27150/d.cnki.ghdzc.2022.000750>
- [2] 卢小宾, 朱庆华, 查先进, 朝乐门. 信息分析导论[M]. 武汉: 武汉大学出版社.
- [3] 刘宝锺. 大数据分类模型和算法研究[M]. 云南大学出版社.
- [4] 谢昌锟, 赵明琪, 林世明. 基于体检大数据的健康指数建模[J]. 数据挖掘, 2021, 11(1): 1-10.
- [5] Zhao, M., Song, C., Luo, T., Huang, T. and Lin, S. (2021) Fatty Liver Disease Prediction Model Based on Big Data of Electronic Physical Examination Records. *Frontiers in Public Health*, 9, Article 668351. <https://doi.org/10.3389/fpubh.2021.668351>
- [6] 贝振东. 大数据分析引擎性能自动优化关键技术研究[D]: [博士学位论文]. 北京: 中国科学院大学(中国科学院深圳先进技术研究院), 2017. <https://doi.org/10.27822/d.cnki.gszxj.2017.000002>
- [7] 王琅. 基于数据挖掘的电商平台促销选品研究[D]: [硕士学位论文]. 南京: 南京大学, 2020. <https://doi.org/10.27235/d.cnki.gnjj.2020.001023>

-
- [8] 楼迎军. 我国期货价格行为与市场稳定机制研究——以大宗农产品期货为例[D]: [博士学位论文]. 杭州: 浙江大学, 2005.
- [9] 孙纪舟. 数据微观不一致性问题的研究[D]: [博士学位论文]. 哈尔滨: 哈尔滨工业大学, 2020.
<https://doi.org/10.27061/d.cnki.ghgdu.2020.001960>
- [10] 陈雄兵. 基于支持向量机和卷积神经网络的匹配交易研究[D]: [博士学位论文]. 武汉: 武汉大学, 2019.
<https://doi.org/10.27379/d.cnki.gwhdu.2019.001582>
- [11] 祝文祥. B2C 电子商务中数据挖掘技术的研究与应用[D]: [硕士学位论文]. 合肥: 中国科学技术大学, 2011.
- [12] 李喆. 基于函数收敛和模糊聚类的评论质量检测方法的研究[D]: [硕士学位论文]. 哈尔滨: 哈尔滨理工大学, 2021. <https://doi.org/10.27063/d.cnki.ghlgu.2021.000123>
- [13] 《中国信用》杂志编辑部电子商务领域信用建设研究课题组. “双十一”网购综合信用评价报告[J]. 中国信用, 2020(1): 16-33.
- [14] Teng, Y., Ren, H., Zhu, J., Jiang, C., Ye, X. and Zeng, H. (2023) A Practical Method for Angular Normalization on Land Surface Temperature Using Space between Thermal Radiance and Fraction of Vegetation Cover. *Remote Sensing of Environment*, **291**, Article ID: 113558. <https://doi.org/10.1016/j.rse.2023.113558>
- [15] Wang, B., Liu, H., Han, X., *et al.* (2021) Image-Based Ransomware Classification with Classifier Combination. *Proceedings of 2021 3rd International Conference on Advanced Information Science and System (AISS 2021)*, Sanya, 26-28 November 2021, 1-6. <https://doi.org/10.1145/3503047.3503083>
- [16] Shin, Y., Mohanty, B.P., Kim, J. and Lee, T. (2023) Multi-Model Based Soil Moisture Simulation Approach under Contrasting Weather Conditions. *Journal of Hydrology*, **617**, Article ID: 129112.
<https://doi.org/10.1016/j.jhydrol.2023.129112>