



丝绸

Journal of Silk

ISSN 1001-7003, CN 33-1122/TS

《丝绸》网络首发论文

题目：基于 CART 决策树算法的服装可持续消费者画像构建
作者：魏山森，梁建芳，雷钦渊
收稿日期：2022-02-28
网络首发日期：2022-08-23
引用格式：魏山森，梁建芳，雷钦渊. 基于 CART 决策树算法的服装可持续消费者画像构建[J/OL]. 丝绸.
<https://kns.cnki.net/kcms/detail/33.1122.TS.20220822.1600.002.html>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

基于 CART 决策树算法的服装可持续消费者画像构建

魏山森, 梁建芳, 雷钦渊

(西安工程大学 服装与艺术设计学院, 西安 710048)

*通信作者 梁建芳, 教授, liangjianfang69@163.com。

摘要 纺织服装行业碳达峰目标的达成要求实现全供应链的绿色发展, 可持续消费行为的践行利于在消费端形成一定的绿色竞争力。目前服装消费端不可持续现象较为严重, 原因在于缺乏针对性的群体引导, 文章通过构建消费者画像来提出服装可持续消费引导策略。首先, 通过 K-means 算法对消费者进行聚类; 其次, 在聚类群体的基础上, 利用卡方检验及回归检验筛选画像指标; 最后, 利用分类与回归树 (Classification and Regression Trees, CART) CART 算法构建服装可持续消费者画像。结果表明: 服装可持续消费者可分为低迷型、能动型、活跃型 3 类; 生活习惯、代际、性别、子女状况、学历是构建能动型及活跃型服装可持续消费者画像的主要指标。进一步地, 根据画像指标对服装可持续消费行为的践行提出针对性策略。

关键词 服装可持续消费; 消费者画像; CART 算法; K-means 算法; 群体分类; 特征分析

中图分类号 TS941.1

文献标志码 A

Portrait construction of sustainable clothing consumer based on CART algorithm

WEI Shansen, LIANG Jianfang, LEI Qinyuan

(School of Apparel and Art Design, Xi'an Polytechnic University, Xi'an 710048, China)

Abstract As one of the pillar industries of China's national economy, the clothing and textile industry creates huge foreign exchange and also causes serious environmental pollution, it should bear the responsibility of reducing carbon dioxide emissions. On October 28, 2021, National Development and Reform Commission made important arrangements for the goal of achieving peak carbon emissions by 2030, which provided a clear guiding plan for the reform of the economic and social development in China, and the green transformation of the clothing and textile industry is imminent. The achievement of the goal of achieving peak carbon emissions of the clothing and textile industry requires the green development of the entire supply chain and the formation of a low-carbon supply chain system. However, although the clothing and textile industry in China has mastered certain sustainable development competitiveness in the upstream and midstream of the supply chain, there are still many unsustainable phenomena on the consumer side, due to the lack of targeted sustainable concept guidance. In order to guide consumers to develop the concept of sustainable clothing consumption, the study builds portraits of Chinese clothing consumers based on the CART algorithm to understand the basic characteristics of sustainable clothing consumers. The study designs indicators for profile construction from demographic variables and sustainable living habits. In order to get the category of sustainable clothing consumers, first of all, to divide the sustainable clothing consumption behavior into purchase, use and disposal stage, and 1835 consumers are measured in 3 stages of sustainable clothing consumption behavior through a 5-point digital scale. Then, based on the quantitative score of sustainable behavior, K-Means algorithm divides sustainable clothing consumers into "low-active type", "mid-active type" and "active type". In order to build "mid-active type" and "active type" consumer portraits, take "low-active type" consumers as a control group to filter and combine portrait indicators. First of all, use the chi-square test and Logistic regression test to screen the portrait indicators, and find that urban-rural differences and marital status are not included in the regression model of consumer attributes; Then, use the CART algorithm to construct "mid-active type" and "active type" sustainable clothing consumer portraits, which is not only extracting the main factors that affect the construction of consumer portraits, but also ranking the importance of the main factors. It is found that living habits, intergeneration and gender are important indicators for building "mid-active type"

收稿日期: 2022-02-28

基金项目: 国家社会科学基金项目(20XSH019)

作者简介: 魏山森(1998—), 男, 硕士研究生, 研究方向为服装设计与工程。

网络首发时间: 2022-08-23 09:04:42 网络首发地址: <https://kns.cnki.net/kcms/detail/33.1122.TS.20220822.1600.002.html>

sustainable clothing consumer portraits. Consumers who actively classify garbage and purchase green food are likely to become "mid-active type" sustainable clothing consumers, and the "post-95s generation" and female are also likely to become "mid-active type" sustainable clothing consumers. The study also found that living habits, the status of the number of children and educational background are important indicators for building "active type" sustainable clothing consumer portraits. Consumers who have purchased new energy vehicles/low-carbon homes are likely to become "active type" sustainable clothing consumers, and consumers with many children and consumers with low education are also likely to become "active type" sustainable clothing consumers. By building the portrait of sustainable clothing consumers in China, it is conducive to in-depth understanding of the internal mechanism of sustainable clothing consumption behavior. Under the policy of achieving peak carbon emissions, in order to achieve sustainable development in the field of clothing consumption, It is essential to pay attention to many aspects, such as the publicity of overall environmental behavior, the promotion of family environmental awareness, the education of sustainable clothing behavior, the implementation of economic subsidies for green clothing and so on. The research results can provide theoretical support for the consumer segmentation of sustainable clothing consumption and the practice of sustainable consumption behavior under the policy of achieving peak carbon emissions and carbon neutrality.

Key words sustainable clothing consumption; consumer portrait; CART algorithm; K-Means algorithm; group classification; characteristic analysis

纺织服装产业的绿色转型与重塑是近年来的热点话题。目前, 尽管中国纺织服装产业在供应链上游和中游环节具备了一定的可持续发展竞争力^[1], 但在消费端仍存在诸多不可持续的现象, 尚未形成良好的循环再利用体系^[2]。有研究指出^[3], 对绿色可持续消费观念的引导, 能有效实现“绿色低碳全民行动”的碳达峰目标, 助力中国“碳中和”发展。

在实践操作层面, 中国服装市场消费群体基数大、范围广, 观念引导需要有针对性地了解消费者的基本特征。消费者画像作为大数据时代下精准营销的产物, 能够很好地依据客观标志区分不同消费者群体, 从而实现多元化的精准服务, 其主要包括人口统计变量、生活行为模式等基本特征^[4]。目前, 关于消费者画像的研究多以固定人群为主体, 通过对画像特点的分析, 提升该类人群的服务体验^[5]。消费者画像的应用领域广泛, 可拓展到不同学科的研究中。尽管如此, 目前在服装可持续消费的研究领域, 多数成果仍为现状调查^[6]、概念界定^[2]及影响机理分析^[7], 鲜有对消费群体特点的深入探讨。为此, 本文以服装可持续消费行为数字量表为依据, 对消费者进行群体划分, 进一步地构建群体画像, 结合画像特征的分析, 深入理解服装可持续消费行为践行的内在机理, 并有针对性地提出引导策略。

1 研究方案设计

1.1 服装可持续消费行为测量量表设计

服装可持续消费行为依照再利用、减量化等绿色意识可分为购买、使用及处置 3 个阶段^[2], 以此作为量表设计依据。为保证量表符合中国服装消费者的特征, 避免出现无效题项, 此次设计在阅读相关文献的基础上, 结合目前服装可持续消费推动的困境, 如协作消费较难推进等, 对题项进行了删减, 最终确定 4 项题目对服装可持续消费行为进行测量, 如表 1 所示。题项采用 Likert 5 点数字量表进行答案设置, 1 表示“从未做到”, 5 表示“每次做到”。

表1 量表测量题项及参考文献
Tab.1 Measurement items and references of the scale

消费行为阶段	测量题项	参考文献
购买	在生活中尽可能地购买绿色环保服装	张倩等 ^[8]
	在生活中尽可能地购买质量较好或经典时尚的服装	
使用	在生活中尽可能地通过服饰搭配达到一衣多用	王建明等 ^[7]
处置	在生活中尽可能地通过修改或再设计将旧衣服重复利用	梁建芳等 ^[9]

1.2 消费者画像指标设计

消费者画像指标应依据不同的研究对象及研究内容进行设计。以往研究的画像指标设计内容一般包括人口统计变量及与研究内容紧密相关的研究对象的其余属性, 研究对象群体越精准, 其相关属性

指标的覆盖面越广泛。但另一方面,以往研究多按照指标特征进行简单的画像描述,缺少对影响画像构建的指标进行结构关系或重要程度的深入研究,这为本研究提供了理论探索契机。

参照文献[4,10]范式和经验,本文的消费者画像指标拟包含人口统计变量及可持续生活习惯两大维度。在根据指标对服装可持续消费者进行画像描述的基础上,还对影响画像构建的指标进行筛选及重要度排序。具体地,人口统计变量包括性别、代际、居住地、城乡、职业、学历、婚姻、子女及月收入,可持续生活习惯指除服装可持续消费外的其他绿色行为,包括是否进行过垃圾分类、是否进行过旧物改造或再利用、是否购买过绿色食品、是否购买过新能源汽车/低碳住宅。

1.3 研究方法 & 研究思路设计

决策树作为机器学习中一种典型高效的分类算法,对于数据量大、分散性高的特征数据有很好的归纳能力,包括 ID3、C4.5、CART 等算法。相比其他算法,首先,CART 算法运算结果更为精准,既能对数据进行分类问题处理,也能解决连续变量的回归问题;其次,CART 算法能够直观地解释各变量间的结构关系,有利于厘清各变量间的主次关系并剔除次要变量^[11]。因此,CART 算法将被用于本研究之中。

根据 CART 算法原理,首先以服装可持续消费行为 4 项测量量表为依据,通过 K-means 算法对样本进行聚类,确定聚类个数并命名;其次,通过卡方检验与回归检验筛选影响画像构建的指标,即借助 RC 列联表对指标进行单因素统计学意义检验,将 $P<0.05$ 的变量纳入无序多元 Logistic 模型中进行回归检验,确定影响服装可持续消费者画像属性的指标因素;最后,将指标因素输入 CART 算法中进行属性归类,构建服装可持续消费者画像。

2 群体聚类与变量检验

2.1 问卷回收与样本描述

本文的调研对象为中国服装消费者,通过线上线下相结合的方式调研,调研时间为 2021 年 3—6 月,共收集问卷 2 486 份,其中有效样本 1 835 份。有效样本的人口统计信息如表 2 所示。为保证有效样本中 4 道量表题项的相关性及独立性,使用 SPSS 23.0 软件对其进行信度及 KMO 和 Bartlett 球状检验。检验结果 α 系数为 0.797>0.65,KMO 值 0.739>0.6,Bartlett 显著性水平 $Sig.=0.000<0.05$,表明测量题项有较高的一致性 & 区分性,可以进行进一步聚类分析。

表2 调研对象主要人口统计信息
Tab.2 Main demographic information of survey subjects

项目	分类	频数/人	占比/%
性别	男	896	48.83
	女	939	51.17
代际	1960 年以前出生	244	13.30
	1960—1982 年出生	579	31.55
	1983—1994 年出生	287	15.64
	1995 年及以后出生	725	39.51
居住地	东部地区	522	28.45
	中部地区	505	27.52
	西部地区	596	32.48
	其他	212	11.55
学历	小学及以下	255	13.90
	初高中(含中专)	303	16.51
	本科(含大专)	1 017	55.42
	硕士及以上	260	14.17
月收入	3 000 元及以下	790	43.05
	3 001-6 000 元	258	14.06
	6 001-10 000 元	282	15.37
	10 001-15 000 元	275	14.99
	15 001 元以上	230	12.53

2.2 K-means 消费者聚类

K-means 聚类算法通过衡量欧式距离来确定样本间相似性，从而对样本进行划分，常用于大规模数据的聚类。使用 K-means 算法对 1 835 名服装可持续消费者进行聚类，聚类指标为 4 项服装可持续消费行为 5 点数字量表数值的总和。

K-means 聚类分析常用簇内误差平方和 (SSE) 作为类簇数量的评价标准，簇内误差平方和即在总样本为 k 个类簇时将各类簇中全部样本 x 与类簇中心 c_i 距离的平方相加后，各类簇相加结果之和，计算公式如下：

$$SSE = \sum_{i=1}^k \sum_{x \in c_i} dist(c_i, x)^2$$

通过设置不同的 k 值进行数据迭代，并计算簇内误差平方和。将 1 835 名服装可持续消费者分为 2 至 8 类时，SSE 值分别为 8 745.362、3 581.917、2 420.659、1 567.134、1 175.233、895.212、807.345，且各簇数量下 P 值均小于 0.05，说明聚类簇数结果均显著，故将各簇数下的 SSE 值绘制为点线图，如图 1 所示。

由图 1 可知，簇数量为 2 时，簇内误差平方和最大，簇内误差平方和随着簇数量的增多递减，簇数量从 3 开始，线段趋于平缓，因此，服装可持续消费者聚类为 3 类时是最合适的。

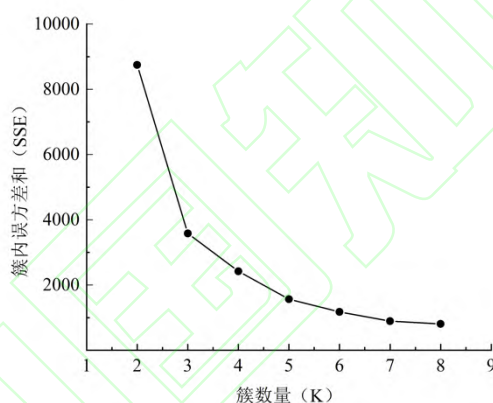


图 1 不同簇数量下簇内误差平方和
Fig.1 SSE value under different number of clusters

通过聚类分析，确定将 1 835 名服装可持续消费者分为 3 类，将每一类簇服装可持续消费评价量表值的平均值作为依据，类簇的平均值越高，表明该类簇的消费者越容易践行可持续行为。分值由低到高分别命名为“低迷型服装可持续消费者”“能动型服装可持续消费者”及“活跃型服装可持续消费者”，如表 3 所示。

表3 各类簇样本状况及命名
Tab.3 Status and naming of various cluster samples

聚类类簇	平均值	数量/人	类簇命名
第一类簇	7.51	433	低迷型服装可持续消费者
第二类簇	11.81	827	能动型服装可持续消费者
第三类簇	16.95	575	活跃型服装可持续消费者

2.3 消费者属性 RC 列联表卡方检验

卡方检验用于判断不同群体在单个变量上是否存在显著差异，以便对研究变量进行初步筛选。本次检验样本量大于 40 且各单元理论数均大于 5，故可进行 RC 列联表卡方检验，结果如表 4 所示。

表4 服装可持续消费者属性卡方检验结果
Tab.4 Chi-square test results of the attributes of sustainable clothing consumers

指标	性别 (X_1)	代际 (X_2)	居住地 (X_3)	城乡 (X_4)	职业 (X_5)	学历 (X_6)	婚姻状况 (X_7)	子女状况 (X_8)
χ^2 值	130.721	244.865	184.252	74.228	166.024	233.169	242.913	269.552
P 值	***	***	***	***	***	***	***	***
最小期望	211.430	57.580	50.030	100.050	25.250	60.170	56.630	43.650
指标	月收入 (X_9)	进行垃圾分类 (X_{10})	进行旧物改造或 再利用 (X_{11})	购买绿色食品 (X_{12})	购买新能源汽车/ 低碳住宅 (X_{13})			
χ^2 值	156.573	48.784	22.658	89.445	166.870			
P 值	***	***	***	***	***			
最小期望	54.270	126.010	117.040	116.800	117.750			

注：***表示 $P<0.001$ 。

由表 4 可知，最小期望值均大于 5，表明本次卡方检验结果成立， P 值均小于 0.05，说明 3 类服装可持续消费群体在 13 项单因素方面的差异均存在统计学意义。因此，13 项单因素变量均可纳入 Logistic 回归模型，通过回归分析对变量进一步筛选。

2.4 消费者属性无序多元 Logistic 回归检验

无序多元 Logistic 回归检验适用于数量在 2 项以上且不存在数量级关系的分类变量，以 1 项变量作为参照组，构建其余变量模型。为检验能动型与活跃型服装可持续消费者属性的影响因素的显著性，本文将低迷型服装可持续消费者作为对照组。将卡方分析中 P 值小于 0.05 的 13 项单因素变量分别代入能动型与活跃型的回归模型，并使用后退法进行变量筛选，排除标准为 $\alpha=0.1$ ^[12]，SPSS 23.0 软件输出结果如表 5 所示。

表5 服装可持续消费者属性Logistic回归检验结果
Tab.5 Logistic regression test results of the attributes of sustainable clothing consumers

单因素变量	能动型服装可持续消费者			活跃型服装可持续消费者		
	β	OR (95%CI)	P	β	OR (95%CI)	P
常量	0.948	-	***	2.105	-	***
性别为女(X_{1b})	0.388	1.474(1.136—1.911)	0.003	-	-	-
代际为 1960—1982 年(X_{2a})	-0.605	0.546(0.338—0.881)	0.013	-	-	-
代际为 1960—1982 年(X_{2b})	-0.357	0.700(0.514—0.953)	0.023	-	-	-
代际为 1983—1994 年(X_{3c})	-0.336	0.715(0.508—1.006)	0.054	-	-	-
居住地为东部地区(X_{3a})	-	-	-	0.730	2.075(1.238—3.476)	0.006
居住地为中部地区(X_{3b})	-	-	-	0.481	1.618(0.972—2.694)	0.064
居住地为西部地区(X_{3c})	-	-	-	0.059	1.061(0.644—1.749)	0.817
职业为企业一般职员(X_{5a})	-	-	-	-0.582	0.559(0.322—0.970)	0.039
职业为自由职业者(X_{5b})	-	-	-	-0.527	0.590(0.283—1.232)	0.160
职业为政府机关或事业单位(X_{5c})	-	-	-	-0.280	0.756(0.430—1.328)	0.330
职业为企业中高级管理者(X_{5d})	-	-	-	-0.682	0.506(0.250—1.024)	0.058
职业为学生(X_{5e})	-	-	-	0.296	1.345(0.599—3.022)	0.473
学历为小学及以下(X_{6a})	-	-	-	0.717	2.049(1.182—3.553)	0.011
学历为初高中（含中专）(X_{6b})	-	-	-	0.207	1.230(0.747—2.026)	0.415
学历为本科（含大专）(X_{6c})	-	-	-	-0.164	0.849(0.551—1.306)	0.456
子女状况为无子女(X_{8a})	-	-	-	-1.472	0.229(0.109—0.481)	***
子女状况为 1~2 个(X_{8b})	-	-	-	-0.784	0.457(0.284—0.734)	0.001
月收入为 3 000 元及以下(X_{9a})	-	-	-	-0.273	0.761(0.435—1.333)	0.340
月收入为 3 001~6 000 元(X_{9b})	-	-	-	-0.593	0.553(0.324—0.944)	0.030
月收入为 6 001~10 000 元(X_{9c})	-	-	-	-0.545	0.580(0.352—0.955)	0.032
月收入为 10 001~15 000 元(X_{9d})	-	-	-	-0.123	0.885(0.540—1.450)	0.627
进行垃圾分类为否(X_{10})	-0.536	0.585(0.446—0.767)	***	-	-	-
进行旧物改造或再利用为否(X_{11})	-	-	-	-0.505	0.603(0.444—0.821)	0.001
购买绿色食品为否(X_{12})	-0.631	0.532(0.396—0.716)	***	-	-	-
购买新能源汽车/低碳住宅为否(X_{13})	-	-	-	-1.041	0.353(0.253—0.492)	***

注：***表示 $P<0.001$

由表 5 可见，城乡差异、婚姻状况不会影响能动型与活跃型服装可持续消费者的属性，最终纳入

能动型服装可持续消费者模型的变量有性别、代际、进行垃圾分类情况、购买绿色食品情况，该模型拟合信息中 $Sig.=0.000<0.05$ ，具有显著性；最终纳入活跃型服装可持续消费者模型的变量有居住地、职业、学历、子女情况、月收入、进行旧物改造或再利用情况、购买新能源汽车/低碳住宅情况，该模型拟合信息中 $Sig.=0.000<0.05$ ，具有显著性。

通过 RC 列联表卡方检验及 Logistic 回归检验，分别将影响能动型与活跃型服装可持续消费者属性的影响因素进行筛选，最终确定 4 项能动型消费者属性影响因素及 7 项活跃型消费者属性影响因素，分别使用 CART 算法对其进行画像构建。

3 CART 决策树分类算法画像构建

3.1 CART 算法结果输出及评价

CART 算法是以最小分割 Gini 系数的属性作为划分依据的二分递归算法，可以避免数据过分拟合，有效提高预测精度。其主要计算步骤如下：

1) 计算初始 Gini 系数值。针对样本训练集 Q ，在此训练集中的属性为 A ，根据每一属性 A ，计算此时的初始 Gini 系数，计算公式如下：

$$Gini(Q) = \sum_{i=1}^n P_i(1 - P_i) = 1 - \sum_{i=1}^n P_i^2$$

式中， n 为训练集中类别个数， P_i 为样本点属于第 n 类的概率。

2) 计算分割 Gini 系数值。针对训练集中的每一属性 A ，以阈值 a 作为属性 A 的分割依据，将训练集 Q 分为 Q_1 和 Q_2 两个子集，分别计算两个子集的分割 Gini 系数值，数值越大，表示分割后集合出错的概率越大，计算公式如下：

$$Gini(Q, A) = \frac{|Q_1|}{|Q|} Gini(Q_1) + \frac{|Q_2|}{|Q|} Gini(Q_2)$$

式中， $Gini(Q, A)$ 为属性 A 的分割阈值为 a 时，训练集 Q 分割子集的错误概率。

3) 确定最佳属性及分割阈值。对于每一属性 A ，选择最小分割 Gini 系数的属性及其阈值作为最佳划分依据，生成两个子节点，并进行样本划分。

依据上述步骤，以低迷型 (R) 服装可持续消费者为对照组，分别对能动型 (M) 与活跃型 (N) 服装可持续消费者进行画像构建，运行软件为 Matlab R2021a。

影响能动型服装可持续消费者的属性有 $X_1(X_{1a}-X_{1b})$ 、 $X_2(X_{2a}-X_{2d})$ 、 X_{10} 、 X_{12} ，“1 (≥ 0.5)”代表具备该属性，“0 (< 0.5)”代表不具备该属性。低迷型与能动型群体共计 1 260 个样本，将 1 008 个样本作为训练集，252 个样本作为测试集，发现叶子节点含有最小样本数为 28 时，交叉验证误差最小，为 0.312 5，生成的决策树如图 2 所示。由图知影响画像构建的拆分点有 4 个，分别是是否购买绿色食品、是否为 1995 年后出生、是否为女性、是否进行垃圾分类。

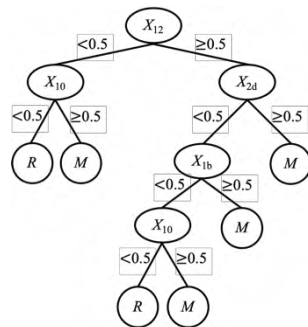


图2 能动型服装可持续消费者决策树分类图

Fig.2 Decision tree classification result map of sustainable clothing consumers of mid-active type

影响活跃型服装可持续消费者的属性有 $X_3(X_{3a}-X_{3d})$ 、 $X_5(X_{5a}-X_{5f})$ 、 $X_6(X_{6a}-X_{6d})$ 、 $X_8(X_{8a}-X_{8c})$ 、 $X_9(X_{9a}-X_{9e})$ 、 X_{11} 、 X_{13} ，“1 (≥ 0.5)”代表具备该属性，“0 (< 0.5)”代表不具备该属性。低迷型与活跃型群体共计 1 008 个样本，将 807 个样本作为训练集，201 个样本作为测试集，发现叶子节点含有最小样本数为 17 时，交叉验证误差最小，为 0.314 7，生成的决策树如图 3 所示。由图 3 知，影响画像构建的拆分点有 5 个，分别是是否购买新能源汽车/低碳住宅、是否有 2 个以上的子女、是否为小学及以下学历、是否居住在东部地区、是否有子女。

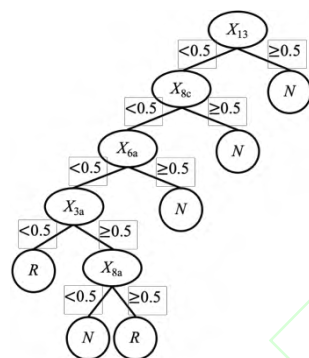


图 3 活跃型服装可持续消费者决策树分类
Fig.3 Decision tree classification result map of c sustainable clothing consumers of active type

本次实验结果可通过测试集样本预测的准确率进行评价，在 CART 算法下，能动型群体测试集识别数为 165，正确预测数为 160，识别率为 96.97%；活跃型群体测试集识别数为 116，正确预测数为 98，识别率为 84.48%。由此可见，识别率均在 80% 以上且总识别率达 90.73%，表明本次试验结果有较高的可信度。

3.2 画像构建结果分析

3.2.1 能动型服装可持续消费者

决策树中越靠近末端的拆分点，在模型构建中发挥的实际作用越小，对于画像构建的影响也较小^[11]。由图 2 可知，能动型服装可持续消费者画像构建指标重要性排序为生活习惯 (X_{10} 、 X_{12}) > 代际 (X_2) > 性别 (X_1)。通过画像逻辑描述可知，在生活习惯上，购买过绿色食品或践行垃圾分类的人群易成为能动型服装可持续消费者；在代际上，“95 后”易成为能动型服装可持续消费者；在性别上，女性易成为能动型服装可持续消费者。

在生活习惯上，进行垃圾分类及购买绿色食品属于受到宏观及微观因素共同影响的可持续行为，同时具有利己及利环境属性^[13-14]，相比旧物再利用等较难践行的可持续生活习惯，其行为驱动机理较易，且容易获得个体利益及环境保护层面的满足感^[13]，因此，在生活中践行垃圾分类及购买绿色食品的消费者在服装可持续消费行为层面被划分为能动型较为合理。在代际方面，“95 后”一代的价值观形成于综合国力快速上升时期，飞速的经济发展、稳定的社会基础及高涨的民族自信与其成长历程紧密相连，巨大的发展福利使“95 后”一代对保障国力提升的方针政策等保有较强的拥护意愿^[15]，因此更愿意践行国家所倡导的低碳绿色生活方式；但同时，深受消费主义文化影响的“95 后”一代也带来了较大的服装消费与浪费^[15-16]，因此，虽然“95 后”相比其他代际群体更愿意践行服装可持续消费行为，但其仍属于能动型。在性别上，已有许多研究表明环境行为存在性别差异，男性对泛化环境问题和公共领域的环境问题表现出更多的关心，女性则对具体环境问题和私人或家庭领域的环境问题表现出更多的行为意向^[17]。作为常见消费品，服装本身的属性使其与个体生活息息相关，服装可持续消费行为也属于与个人与家庭密切关联的具体的环境行为，因此从理论上，女性相比男性更愿意践行服装可持续消费行为，本文结论与之相契合，但同时，与“95 后”一代类似，女性在服装上的高消费特性决定了其属于能动型群体。

3.2.2 活跃型服装可持续消费者

由图 3 可知, 活跃型服装可持续消费者画像构建指标重要性排序为生活习惯 (X_{13}) > 子女状况 (X_8) > 学历 (X_6), 其余指标在决策树模型中未发挥实际作用或发挥作用较小。通过画像逻辑描述可知, 在生活习惯上, 购买过新能源汽车/低碳住宅的人群易成为活跃型服装可持续消费者; 在子女状况上, 子女数量较多或有子女的人群易成为活跃型服装可持续消费者; 在学历上, 低学历人群更容易成为活跃型服装可持续消费者。

在生活习惯上, 新能源汽车与低碳住宅都属于具有高环保价值的新兴消费品, 而较高的溢价支出及使用功能的可替代性导致其市场竞争力较低, 因此, 购买新能源汽车或低碳住宅的消费者多数看中其“绿色”的商品属性, 该类人群往往具有更强的环境意识与社会责任感, 在生活各方面会积极践行环境行为^[18-19]。对照本研究结论, 认为购买过新能源汽车或低碳住宅的消费者在服装购买和使用方面也具有更强的可持续意识, 故将此类消费者划分为活跃型。在子女状况上, 家庭的环境行为引导作为将个体环境意识上升至公共环境关心的有效路径, 对子女有着深远的影响, 通常出于对子女良好的教育与子女未来福祉等因素的考量, 拥有子女的家庭通常更加关心环境问题并积极践行环境行为^[20]。对照本文研究, 认为子女数量较多或有子女的人群有较强的服装可持续消费意识。此外, 服装较强的再利用属性导致其会在多子女家庭群体中进行代际内的多次使用, 有效延长了其使用寿命, 故将此类群体划分为活跃型。在学历上, 已有许多研究表明学历水平与可持续行为的践行程度呈正相关性, 而本文发现低学历群体易践行服装可持续消费行为, 这一结论表明服装的可持续行为与其他层面的可持续行为存在一定的差异。服装生命周期各阶段的环境行为与大众生活环境的改善间存在“感知鸿沟”, 例如新能源汽车带来的大气环境改善、垃圾分类带来的空间环境改善等, 都能在生活中进行直观地感受, 而就目前消费者在服装购买、使用、处理等阶段可持续行为认知的模糊性^[2], 导致其对于服装的可持续行为对提升环境水平的感知并不强烈; 同时, 不同学历人群所处社会阶层的差异间接导致了服装购买与使用等行为的差异, 低学历人群的服装消费水平往往较低, 在购买及使用服装时更多注重其实用性, 而较长的使用时间则延长了服装的使用寿命, 导致服装可持续消费行为量化分数较高, 成为活跃型群体。

4 建议对策

通过聚类分析, 发现目前中国消费者在服装可持续消费践行程度方面可被划分为低中高 3 个层次的群体, 即低迷型、能动型与活跃型。通过卡方检验、回归检验及 CART 算法分析, 发现影响能动型群体画像构建的指标按照重要程度排序为生活习惯、代际及性别, 购买过绿色食品或进行垃圾分类的人群, “95 后”及女性是能动型服装可持续消费者的主要组成; 影响活跃型群体画像构建的指标按照重要程度排序为生活习惯、子女状况及学历, 购买过新能源汽车/低碳住宅的人群, 子女较多或有子女的人群, 低学历人群是活跃型服装可持续消费者的主要组成。据此, 提出践行服装可持续消费行为的对策。

1) 规避单一宣传服装方面的可持续行为, 着眼全渠道环境行为宣传。研究表明, 不论对于能动型还是活跃型服装可持续消费者, 生活习惯均是画像构建最重要的影响指标, 说明具有可持续生活习惯的消费者同样会在服装上践行可持续行为。目前, 服装可持续消费行为相比其他环境行为来讲, 存在意识不足、效力欠缺等推进困境, 其践行不能一蹴而就, 需要从消费者环境感知较为清晰的公共环境领域切入, 在环境意识提升后逐渐过渡, 避免仅宣传服装可持续行为造成环境利益感知不足。在碳达峰全民行动的背景下, 充分利用新媒体优势, 通过多种渠道对生活中常见的环境行为进行宣传教育, 在利用传统媒介的同时, 着重利用移动互联网、手机客户端等新媒体, 通过微博、公众号等移动社交应用对常见环境行为进行宣传; 此外, 定期通过学校、单位、社区等进行宣传教育, 通过群体效应来提升总体环境意识, 进而达到各方面可持续行为的提升。

2) 注重家庭层面环境意识与行为的培养, 提倡理论教育与实践并行。研究表明, 多子女或有子女家庭、“95 后”及女性均是构建服装可持续消费者画像的重要影响指标, 这 3 项特征均指向了家

庭这一重要教育途径。女性母亲身份的特殊性奠定了其对子女教育的重要地位,同时“95 后”群体父母一代普遍接受过一定的教育,“95 后”较高的可持续意识也与良好的家庭教育有着紧密的联系。由于服装可持续行为属于典型的私域环境行为,家庭成为其教育的重要阵地。家庭环境行为教育应首先归正家长的可持续意识,宣传应以亲子教育、未来环境福祉等为切入点,将可持续理念逐渐渗入到宣传教育当中,使家长意识到家庭环境行为教育与学校教育同等重要;同时要可持续意识应用到实践中,利用学校、社区、社会团体等组织的亲子环保活动,共同践行诸如旧衣改造等可持续行为,将环保教育纳入到国民日常生活及教育体系中。

3) 加强国家政府及龙头企业的引导作用,增强政策的弹性与灵活性。研究表明,低学历群体是构建服装可持续消费者画像的重要指标,这与较低的服装消费水平及较长的使用周期等有着直接关系,造成一种“假绿”的现象,而这种“假绿”现象也恰恰反映出服装可持续消费行为推进的艰难性。从源头来看,环保健康的绿色服装市场竞争力较差,同其他行业环保企业类似,政府应适当对绿色服装生产企业施行经济减税政策,使绿色服装形成一定的价格优势,并根据不同地区经济、社会的差异实施差异化激励政策,鼓励消费者购买绿色服装;在服装购后阶段,服装相关的环保事务不同于其他环保事业可以得到政府的有力监管,龙头企业应肩负起一定的社会责任,如在线下零售店建立相关回收设施,鼓励消费者积极参与到服装回收的可持续行为中,优化服装回收体系,形成政府引导、行业协助、民众参与等多位一体的环境协同模式。

5 结 语

目前,中国服装可持续消费者按照可持续行为践行程度可分为高中低 3 个层次,影响服装可持续消费者画像构建的指标有生活习惯、代际、性别、家庭子女状况及学历。在“双碳”政策背景下,实现服装可持续消费行为的途径主要有重视总体环境行为宣传、提升家庭环境意识并注重服装可持续行为教育、实行绿色服装经济补贴等。本文以 1 835 份问卷数据为研究基础,通过聚类分析、回归分析、CART 算法分析等构建中国服装可持续消费者画像,为服装可持续消费的消费细分及“双碳”政策下服装可持续消费的践行提供理论支撑。与此同时,本次研究针对所有的服装消费者,没有群体或角色的限制,因此画像指标仅从人口统计变量与可持续生活习惯两方面进行了设计,在设计维度上存在一定的局限性。在后续服装消费者画像的研究中,需进一步细化研究对象:在群体上,可根据性别、代际等属性将研究群体清晰化;在内容上,可聚焦在某一消费阶段,如回收阶段,从而丰富与细化画像构建的指标,以得到更有针对性的研究结论。

参考文献:

- [1] 陈李红, 严新锋, 丁雪梅, 等. 基于网络层次分析法的纺织服装产业可持续竞争力评价[J]. 纺织学报, 2018, 39(10): 162-167.
CHEN Lihong, YAN Xinfeng, DING Xuemei, et al. Evaluation of sustainable competitiveness of textile and costume industry based on analytic network process theory[J]. Journal of Textile Research, 2018,39(10):162-167.
- [2] 梁建芳, 程婉莹. 服装可持续消费行为的研究现状及困境分析[J]. 丝绸, 2020, 57(6): 18-25.
LIANG Jianfang, CHENG Wanying. Research on status and dilemma of sustainable clothing consumption behavior[J]. Journal of Silk, 2020, 57(6): 18-25.
- [3] 胡鞍钢. 中国实现 2030 年前碳达峰目标及主要途径[J]. 北京工业大学学报(社会科学版), 2021, 21(3): 1-15.
HU Angang. China's Goal of Achieving Carbon Peak by 2030 and Its Main Approaches[J]. Journal of Beijing University of Technology(Social Sciences Edition), 2021,21(3): 1-15.
- [4] 方方, 蒋蒙蒙, 刘苗. 基于生活习惯量表的上海地区孕妇用户画像[J]. 东华大学学报(自然科学版), 2021, 47(3): 68-74.
FANG Fang, JIANG Mengmeng, LIU Miao. Pregnant women's user portrait and analysis of consumption behavior in Shanghai based on lifestyle scale[J]. Journal of Donghua University (Natural Science), 2021,47(3):68-74.
- [5] AN J, KWAK H, JUNG S G, et al. Customer segmentation using online platforms: isolating behavioral and demographic segments for persona creation via aggregated user data[J]. Social Network Analysis and Mining, 2018,8(1): 54-54.
- [6] IRAN S, GEIGER S M, SCHRADER U. Collaborative fashion consumption: A cross-cultural study between Tehran and Berlin[J]. Journal of Cleaner Production, 2019, 212: 313-323.
- [7] 王健明, 武落冰, 何正霞. 慢时尚博主特性对服装可持续消费行为的影响机制: 网红经济时代“近朱者赤”的“认同-责任”模型[J]. 中国地质大学学报(社会科学版), 2021, 21(4): 90-104.
WANG Jianming, WU Luobing, HE Zhengxia. The impact of ethical fashion blogger's characteristics on sustainable consumption in eco-apparel: The "identification-responsibility" model of "keep good men company and you shall be of the number" in internet celebrity

- economy age[J]. Journal of China University of Geosciences(Social Sciences Edition), 2021,21(4):90-104.
- [8] 张倩, 韩燕. 绿色服装购买意愿及影响因素研究[J]. 丝绸, 2013, 50(12): 41-45.
ZHANG Qian, HAN Yan. Study on purchasing intention of green apparel and influencing factors[J]. Journal of Silk, 2013,50(12): 41-45.
- [9] 梁建芳, 和嘉伟. 基于行为意向中介效应的可持续消费认知与服装再利用行为的关系[J]. 东华大学学报(自然科学版), 2020, 46(3): 463-471.
LIANG Jianfang, HE Jiawei. Relationship between Chinese consumers' sustainable consumption cognition and clothing reusing behavior based on mediating effect of behavioral intention[J]. Journal of Donghua University(Natural Science), 2020, 46(3): 463-471.
- [10] 唐晓波, 高和璇. 基于特征分析和标签提取的医生画像构建研究[J]. 情报科学, 2020, 38(5): 3-10.
TANG Xiaobo, GAO Hexuan. Study of the doctor portrait based on feature analysis and label extraction[J]. Information Science, 2020, 38(5): 3-10.
- [11] 李鹏, 雷雨秋, 刘宗杰, 等. 基于决策树算法的断路器弹簧操动机构振动诊断技术[J]. 高压电器, 2021, 57(9): 1-8.
LI Peng, LEI Yuqiu, LIU Zongjie, et al. Vibration diagnosis technology of technology of spring operating mechanical for circuit breaker-based on decision tree algorithm[J]. High Voltage Apparatus, 2021, 57(9): 1-8.
- [12] 李小强, 莫淼, 吴菲, 等. 基于问卷调查的上海女性乳腺癌人工神经网络预测模型[J]. 肿瘤, 2018, 38(9): 883-893.
LI Xiaoqiang, MO Miao, WU Fei, et al. Artificial neural network models based on questionnaire survey for prediction of breast cancer risk among Chinese women in Shanghai[J]. Tumor, 2018, 38(9): 883-893.
- [13] 李玮, 王志浩, 刘效广. 宣传教育对城市居民垃圾分类意愿的影响机制: 环境情感的中介作用及道德认同的调节作用[J]. 干旱区资源与环境, 2021, 35(3): 21-28.
LI Wei, WANG Zhihao, LIU Xiaoguang. Mechanisms of public education influencing waste sorting willingness of urban residents: The mediation of environment emotion and moderation of moral identity[J]. Journal of Arid Land Resources and Environment, 2021,35(3):21-28.
- [14] WANG J M, NGUYEN N, BU X Z. Exploring the roles of green food consumption and social trust in the relationship between perceived consumer effectiveness and psychological wellbeing[J]. International Journal of Environmental Research and Public Health, 2020, 17: 4676.
- [15] 陈立明, 刘炳辉. 复兴一代: “95 后”大学生的时代特性剖析[J]. 华东理工大学学报(社会科学版), 2019, 34(4): 107-116.
CHEN Liming, LIU Binghui. Revival generation: An analysis of the characteristics of the post-95 college students[J]. Journal of East China University of Science and Technology(Social Science Edition), 2019,34(4):107-116.
- [16] 封竹, 梁建芳, 王雨田. 生活方式细分下的 90 后群体服装可持续消费行为[J]. 纺织高校基础科学学报, 2021, 34(2): 74-80.
FENG Zhu, LIANG Jianfang, WANG Yutian. Sustainable clothing consumption behavior of the post-90s group under the lifestyle segmentation[J]. Basic Sciences Journal of Textile Universities, 2021,34(2):74-80.
- [17] 杜平, 张林虢. 性别化的亲环境行为: 性别平等意识与环境问题感知的中介效应分析[J]. 社会学评论, 2020, 8(2): 47-60.
DU Ping, ZHANG Linxiao. Gendered pro-environmental behavior: Analysis of the mediation effects of gender equality awareness and perception of environmental problem[J]. Sociological Review of China, 2020, 8(2): 47-60.
- [18] 李苏秀, 刘颖琦, ARI K. 新能源汽车产业公众意识培育策略: 北京数据与国际经验[J]. 北京理工大学学报(社会科学版), 2017, 19(3): 57-66.
LI Suxiu, LIU Yingqi, ARI K. Strategies for public swariness cultivation of new energy vehicles industry: Beijing data and international experiences[J]. Journal of Beijing Institute of Technology (Social Sciences Edition), 2017, 19(3): 57-66.
- [19] 杨雪锋, 梁邦利. 低碳住宅购买行为的影响因素: 基于杭州市的实证研究[J]. 城市问题, 2013(7): 9-17.
YANG Xuefeng, LIANG Bangli. Influencing factors of low-carbon housing purchase behavior: An empirical study based on Hangzhou city[J]. Urban Problems, 2013(7): 9-17.
- [20] RICKINSON M. Learners and learning in environmental education: A critical review of the evidence[J]. Environmental Education Research, 2001, 7(3): 207-320.