

基于 FTRM 模型和 K-means 算法的 大学生知识付费产品使用行为研究

巫 芯 宇

西南大学 图书馆, 重庆 400715

摘要: 我国知识付费行业已进入到精细化运营时代, 而精细化运营的前提是对用户的差异化进行分类研究. 大学生是使用知识付费产品的主要群体之一, 而目前针对该群体的知识付费产品使用行为差异化分类研究还很少, 因此基于 FTRM 模型和 K-means 算法的大学生知识付费产品使用行为研究具有理论意义. 本文首先对大学生使用知识付费产品的动机和认知进行分析, 然后对大学生的基本属性与知识付费产品使用行为的关系进行了相关性探索, 最后基于传统 RFM 模型, 结合知识付费产品的特点, 提出一种改进的 FTRM 模型. 基于该模型, 利用 K-means 聚类算法, 将大学生群体聚类为睡眠型(或流失型)用户、忠实型用户、未付费转化型用户和冲动消费型用户 4 类用户群体, 并针对每一类群体提出差异化的营销建议.

关 键 词: 知识付费; FTRM 模型; K-means 算法; 行为研究

中图分类号: GP399

文献标志码: A

文章编号: 1673-9868(2021)06-0195-10

2016 年, 被称为知识付费元年, 自 2016 年以来, 我国知识付费行业始终保持快速增长趋势. 艾媒新经济产业研究中心报告显示, 2019 年我国知识付费行业用户规模达 3.6 亿人, 行业市场规模达 278 亿元, 估计 2020 年会突破 392 亿元人民币, 其中 18~30 岁的人群占比超过 50%^[1]. 然而, 知识付费行业在保持快速增长的同时, 也出现了市场竞争加剧、产品同质化、产品价格不合理、广告泛滥、缺乏知识产权保护等问题^[2], 这对知识付费行业的持续发展带来了不利影响. 此外, 尽管我国知识付费行业规模持续增长, 但其同比增速却在下降.

目前, 我国知识付费行业已进入到精细化运营时代, 需要对用户群体的差异化进行分类研究. 大学生是使用知识付费产品的主要群体之一, 但是目前针对该群体的研究主要集中在动机和影响因素等方面, 而基于大学生群体知识付费产品使用行为的用户细分研究还相对较少. 因此, 本文基于传统的 RFM 模型, 结合知识付费产品的特点, 提出一种改进的 FTRM 模型, 利用 K-means 聚类算法, 将大学生群体聚类为睡眠型(或流失型)用户、忠实型用户、未付费转化型用户、冲动消费型用户 4 类用户群体, 并针对每一类群体提出差异化的营销建议.

1 国内外研究现状

本文以“(大学生 + 高校学生) * 知识付费”作为检索词对知网数据库进行主题精确检索, 共检索到相

收稿日期: 2020-12-22

基金项目: 国家自然科学基金项目(41671291); 重庆市图书馆学会 2020 年度科研课题资助项目(CQTX202011); 重庆市高校图工委项目(203771); 2021 年 CALIS 全国农学文献信息中心研究项目(2021060); 西南大学 2020 年度继续教育研究项目(SWU2008015).

作者简介: 巫芯宇, 博士研究生, 馆员, 主要从事文化育人、科普活动推广、资源融合共享、图书馆服务与管理等方面的研究.

关文献 66 篇。通过对这 66 篇文献的分析发现,目前对大学生知识付费产品使用行为的研究主要集中在大学生使用知识付费产品的现状、付费意愿、影响因素以及知识付费产品的有用性、易用性等方面。有关国内大学生知识付费相关研究主题(前 10)如下所示(图 1)。

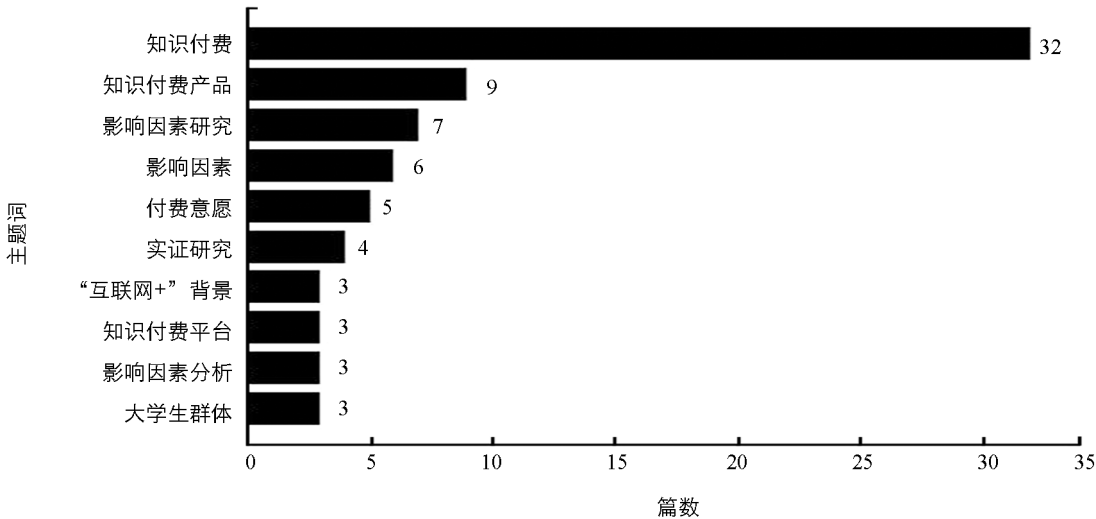


图 1 国内大学生知识付费相关研究主题(前 10)

孙晓臻等^[3]以青岛大学为例,对大学生使用知识付费产品的动机、态度等进行了调研分析,提出了知识分享平台应坚持内容为王的发展战略和低收费、多受众的营销策略。何慧雪等^[4]基于技术接受模型和期望确认理论,提出影响大学生持续使用知识付费产品意愿的 8 个因素。陈娟等^[5]基于修正后的技术接受模型 TAM2 和感知价值接受模型,构建了知识付费产品使用偏好模型。刘诗卉等^[6]以性别、年级、月均消费额(Average monthly consumption,简称 A)等为自变量,以对知识付费的态度为因变量进行二元 Logit 回归分析,认为经济条件和内容质量是影响大学生知识付费的主要因素,发现大学生知识付费产品使用行为注重实用性和娱乐性等特点。刘澜等^[7]对大学生知识付费产品使用行为、付费需求和付费态度等进行了分析,并结合文化产业政策、数字出版市场、图书馆资源组织和阅读服务等方面提出了数字付费阅读使用对策。

此外,笔者以类似的方法对 EI 数据库进行检索,检索到“online paid knowledge OR pay for knowledge”与“college student OR university student”关键词组相关文献 149 篇。通过对这些文献进行分析发现,外文期刊中对大学生知识付费产品使用行为的研究并不多,且不少文献是来自中国学者的研究,现有的研究也主要集中在知识付费意愿(Willingness to pay)、使用决策模式(Decision Making)、在线教育(E-Learning)等方面(图 2)。尽管知识付费相关概念和模式起源于国外,但是由于知识产权保护等原因,知识付费产品在国外发展并不那么迅速^[8]。而且,外文期刊中还没有比较明确的知识付费概念,相关研究成果较少,仅有少量国外媒体采用“online paid knowledge”来报道知识付费相关内容。

Xie Wei 等^[9]认为知识付费产品的功能价值、情感价值、社会价值均对购买意愿产生积极且显著的正向影响。Moutaz Khouja 等^[10]指出不断提升服务是影响付费成员支付意愿的重要因素。另外,雷兵等^[11]在对国内外知识付费研究进展的文献综述中指出,国内外现有研究较多关注平台商业模式及个人行为的影响因素,缺乏更深层次研究,且相关文献关注点较为集中,使研究存在一定的同质性。

目前,国内外学术界对大学生知识付费主题的研究主要集中在知识付费动机及行为影响因素两方面。但大学生群体因性别、专业、月均消费额等不同,在知识付费产品使用行为上存在差异。杨舒怡^[12]认为,知识付费平台应开展用户研究,并深入挖掘用户需求,而挖掘用户需求不能单单以年龄、学历、职业等标签划分用户群,而要找出更具个性化的标签来区分用户。因此,本文根据大学生使用知识付费产品的行为差异,将大学生群体划分为不同类别,并制定差异化的营销策略,实现精准营销,有助于进一步推动知识付费行业持续发展。以“(大学生 + 高校学生)*知识付费*(聚类 + 分类 + 细分)”作为检索词在知网数

数据库中进行检索，仅检索到 1 篇相关论文，该论文对大学生知识付费行为进行了调研分析，并运用 SWOT 分析方法对知识付费市场需求提出了建议^[13]。尽管该文使用了 K-means 聚类算法，但文章聚类的对象是影响知识付费产品的变量，并非大学生用户群体。

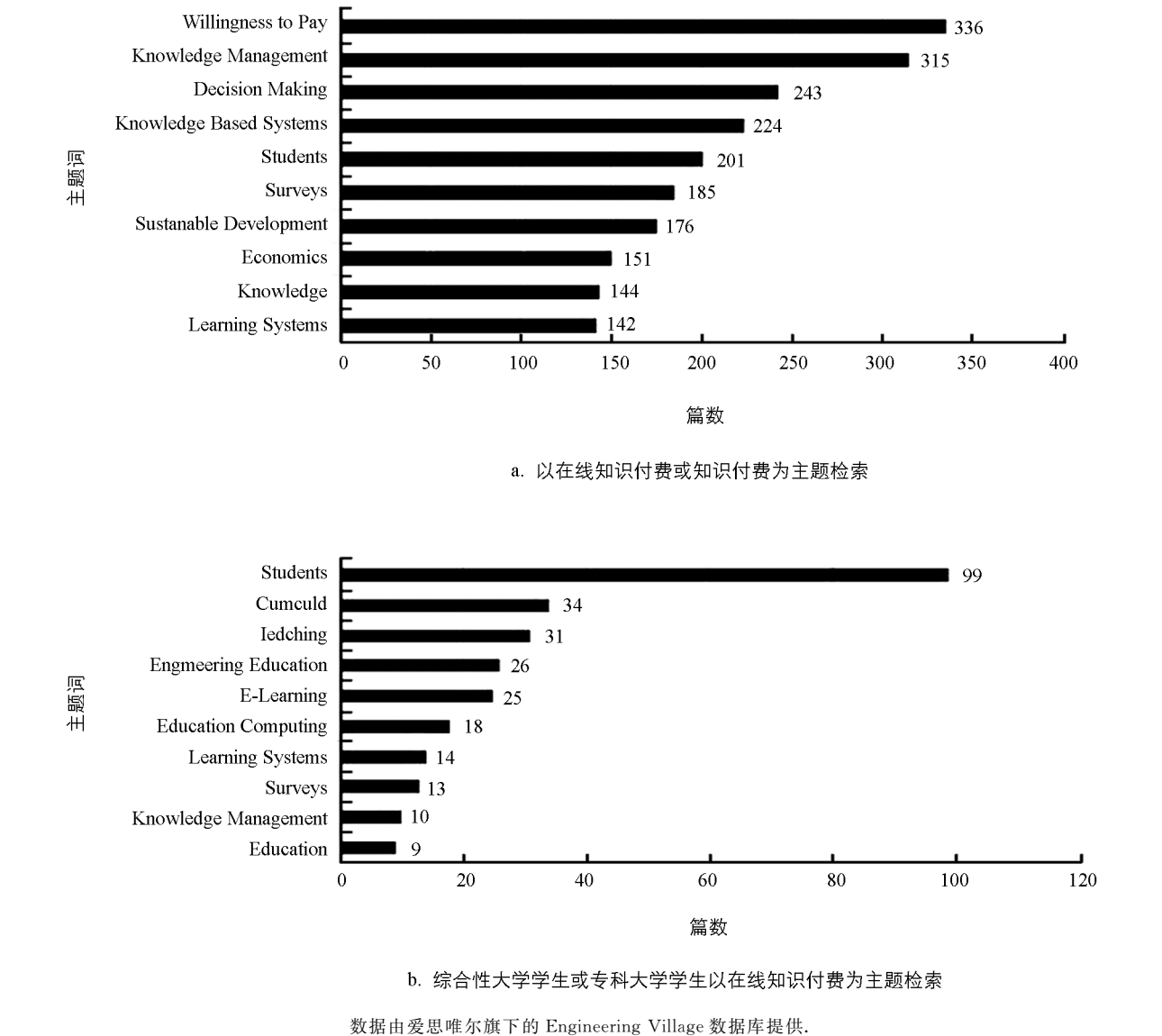


图 2 国外大学生知识付费相关研究主题

综上所述，通过对当前国内外学者关于大学生知识付费主题的研究现状进行分析，笔者认为：目前学术界对大学生群体使用知识付费行为进行的用户聚类、用户细分等研究相对较少。目前，基于用户分类的精准营销、精细化运营已成为各类互联网产品运营必须采取的措施。因此，本文基于传统的 RFM 模型，结合知识付费产品的特点，提出了 FTRM 模型，并基于该模型利用 K-means 聚类算法^[14]，将大学生群体细分为不同类别，并针对不同用户特点提出差异化的营销建议，以提高知识付费产品的营销效率，实现精准营销。

2 FTRM 模型与 K-means 算法

2.1 FTRM 模型

RFM 模型是衡量客户价值和客户创利能力的重要工具和手段，被广泛应用于用户价值衡量、用户细分等领域中。在 RFM 模型中， R (Recency) 表示最近一次消费时间到当前时间的时间间隔， R 值越大表示用户价值越低； F (Frequency) 表示某段时间内消费的次数， F 值越大表示用户价值越高； M (Monetary) 表

示某段时间内消费的总金额, M 值越大表示用户价值越高^[15].

传统的 RFM 模型主要用于直销产品, 通过用户付费行为对用户价值进行衡量. 然而, 知识付费产品作为一种基于互联网的软件产品, 区别于传统的消费品, 用户日常活跃行为也是衡量用户价值的重要指标. 在 App 产品数据分析中, DAU(Daily Active Users, 日活跃用户数)、MAU(Monthly Active Users, 月活跃用户数)也是最重要的分析指标. 因此, 本文基于传统的 RFM 模型, 结合知识付费产品的特点, 提出针对知识付费产品行为分析的 FTRM 模型. 其中, F, T 用于衡量用户的日常活跃行为, R, M 用于衡量用户的付费购买行为. FTRM 具体含义如下:

F (Frequency): 使用频率, 表示用户使用某 App 的月平均次数.

T (Time Length): 使用时长, 表示用户每次使用某 App 的时间长度(min).

R (Recency): 最近一次付费, 表示用户最近一次购买时间到当前时间的时间间隔.

M (Monetary): 付费金额, 表示用户在某段时间内付费购买的总金额.

2.2 K-means 算法

K-means 聚类算法是一种经典的、最常用的无监督学习聚类算法. 通过 K-means 算法, 可以将数据按照距离分为不同的簇, 簇内样本相似度高, 簇间样本相似度低. 其核心步骤如下:

S1: 从数据集中选取 K 个数据对象作为初始聚类中心.

S2: 逐一计算数据集中每个样本到 K 个初始聚类中心的距离, 并根据距离远近, 将每个样本聚类到距离其最近的聚类中心, 形成不同的簇.

S3: 计算每个簇中样本的均值, 将均值作为新的聚类中心.

S4: 重复 S1~S3, 通过不断迭代, 直至聚类中心不再变化.

S5: 评估聚类效果. 由于聚类数量 K 是人为确定的, 不同 K 值下聚类的结果不同, 一般常用的 K 值确定方法是计算不同 K 值下的簇内误方差(SSE)和轮廓系数(Silhouette Coefficient), 当 K 值位于 SSE 学习曲线变化的“肘部”位置或使得轮廓系数取最大值时, 即为最佳 K 值.

通过 FTRM 模型对使用行为进行建模, 再利用 K-means 算法对数据进行聚类, 对用户使用情况进行分析, 挖掘不同用户的使用特点.

3 数据的获取与预处理

3.1 用户行为的界定

用户行为复杂多样, 主要受内在心理和外在环境因素的影响, 其研究涉及范围较广, 囊括了经济学、传播学、心理学、计算机技术等学科, 课题研究综合性较强. 因此在知识付费用户行为研究中, 对其范围进行界定也极为重要. 董富强等^[16]认为用户行为是指用户在使用资源时所呈现出的规律. 杨善林等^[17]对用户行为进行了细致分类, 将其分为消费行为、参与行为和创造行为 3 类. 消费行为是指用户对知识付费产品购买, 参与行为则是用户与内容生产者的交互, 创造行为是指用户独立生产和创造一定内容. 目前学术界并未对知识付费行为有一个统一的界定, 而是将其统称为用户对知识付费产品的购买使用.

基于此, 笔者对知识付费产品行为做了一个简单的归纳: 对用户购买和使用知识付费产品等行为进行搜集和分析, 发掘出不同类型的用户因自身的特点和需求在知识付费产品使用上的不同特点, 从而为其打造专属的营销策略. 本文的用户仅限大学生群体, 用户使用行为包括付款、购买等消费行为、参与行为和创造行为. 而活跃行为则是对使用行为状态的描述, 与潜水行为相对应, 具体指用户在使用知识付费产品时积极或消极的态度. 用户使用知识付费产品越频繁, 体验感越好, 购买欲望则越强.

并不是所有用户在使用知识付费产品时都存在购买的意愿, 因此商家在宣传时往往会打出“打折”“促销”等口号, 并采用优惠拼团或 7 天试用等方式, 降低用户感知价值的风险. 商家还会提供按篇购买、积分兑换、试读部分章节等行为, 通过较低的入场费促进用户最终进行购买消费. 用户获得了优质的知识内容或解决了实际问题后, 商家还设置了包月行为、包年行为、普通用户行为、vip 行为等供用户选择, 最大限

度地提升用户的使用体验，以便吸引更多的用户群体。用户行为实际上是由商家提供的支撑行为决定的，不是独立存在的，因此这些行为也将被转换为用户行为。

3.2 问卷设计与数据获取

为了深入研究大学生使用知识付费产品的情况，本文设计了调研问卷，对大学生基本属性、动机、认知及行为等方面进行调研，并利用问卷星平台获取调研数据(只针对大学生群体进行随机发放)。

基本属性包括性别、学历、专业类别、月均消费额等，其中专业类别按照教育部普通高校本科专业类别划分标准设置；动机主要是使用知识付费产品的原因；认知包括对知识付费产品的认知程度，喜欢哪种类型的知识付费产品，喜欢哪种知识提供者，目前知识付费产品存在的问题等；行为主要包括日常活跃行为和付费购买行为，日常活跃行为包括最近一年内，月均使用频率和每次使用的平均时长，付费行为包括最近一次付费时间和最近一年付费的总金额。

3.3 数据预处理

本次调研共获得问卷 5 219 份，数据预处理时首先剔除完全不了解知识付费产品的问卷 805 份，最终获得有效问卷 4 414 份。由于 F, T, R, M 这 4 个指标量纲不同，为了避免量纲不同带来的偏差，在数据预处理时，对这 4 个指标进行了归一化处理，归一化公式如下：

$$x = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}}$$

(1)

式(1)中， X_i 为各指标样本集中每个样本的值， $X_{\max}$ 和 $X_{\min}$ 分别为各指标样本集的最大值和最小值。

4 大学生知识付费产品使用行为描述性分析

4.1 动机与认知描述性分析

在动机层面，大学生使用知识付费产品的主要原因有获取专业知识，获取更有价值的信息以及满足一定的社交需求等，也有少量原因是新鲜感、好奇心、从众等心理因素(图 3)。

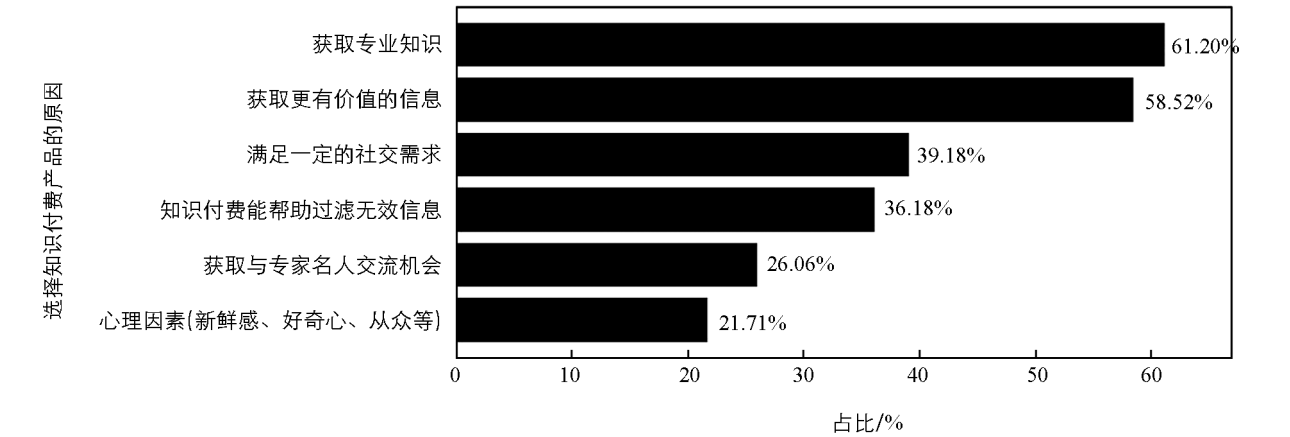


图 3 大学生使用知识付费产品动机

在总体认知层面，知识付费产品在大学生群体中普及率达 85%，平均熟悉程度为 6.74 分(总分 10 分)。其中，41% 的大学生用户表示对知识付费产品熟悉程度达到 8 分及以上。而在不同类型知识付费产品的喜好上，大学生群体更倾向于知识平台类，如喜马拉雅、得到、蜻蜓 FM、豆瓣时间等，以及付费文档类，如百度文库、豆丁文库、道客巴巴等；而对于社区直播，如知乎 Live、荔枝微课、千聊，以及社群类，如知识星球、小红圈等使用相对较少。此外，对不同类型知识付费内容提供者的喜好上，大学生群体更倾向于专家学者和行业大咖(图 4)。此外，针对目前知识付费产品存在的问题，63% 的大学生用户表示知识付费产品定价较高或存在不合理现象；60% 的大学生用户表示知识付费产品内容质量混杂，部分质量较低；也有部分大学生用户表示知识付费产品广告较多，降低了用户体验度。

从动机与认知层面可以看出，知识付费产品在大学生群体中使用率相对较高；从动机和喜好看，知识

付费平台的内容质量、专业程度是其关注的核心。

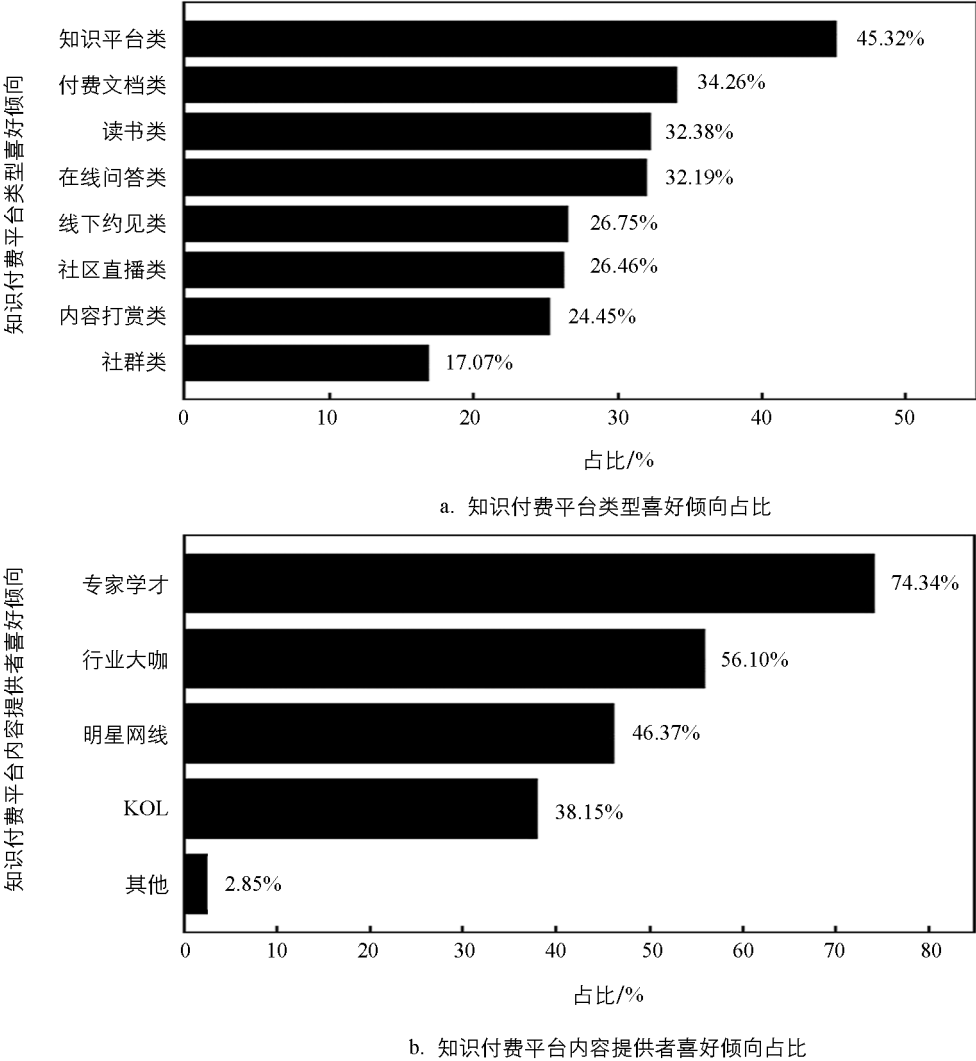


图 4 大学生使用知识付费产品倾向与喜好

4.2 基本属性与使用行为关系分析

一般而言，用户属性的差异会导致不同的使用行为。因此，本文进一步挖掘了大学生群体的基本属性与知识付费产品使用行为之间的相关性，即探讨性别、学历、专业、月均消费额 A 与 F, T, M 之间的相关性。

1) 性别、学历、专业与 F, T, M 的关系分析

由于性别、学历、专业均为定性数据，而 F, T, M 均为定量数据，因此，本文使用方差分析方法，分析基本属性的不同分类(如性别男、女)在特定使用行为指标(如 F)下的差异，如表 1 所示。

表 1 性别、学历、专业与 F, T, M 的关系分析

类别	性别 P 值	学历 P 值	专业 P 值
F	0.28	0.37	0.37
T	0.41	0.25	0.19
M	0.00	0.00	0.00

表格中 P 值为某基本属性的不同分类在特定使用行为指标下的差异，如 F 项的性别 P 值，表示性别基本属性中，不同分类(男、女)在 F 上是否存在差异。如果 $p < 0.05$ ，则说明差异有统计学意义，即表现出一定的相关性。因此，由表 1 可知，性别、学历、专业均对 F, T 差异无统计学意义，而对 M 均差异有统计

学意义($p<0.05$),也就是说,性别、学历、专业与 F,T 之间没有相关性,而与 M 之间存在一定相关性.进一步分析这种相关性可以发现,大学生群体中,男性知识付费金额比女性高 3%,而博士、硕士、本科知识付费金额依次为 3%,3%,10%.对不同专业大类,文学、历史学、医学等专业知识付费金额相对较高,而工学、理学、经济学则相对较低(图 5).

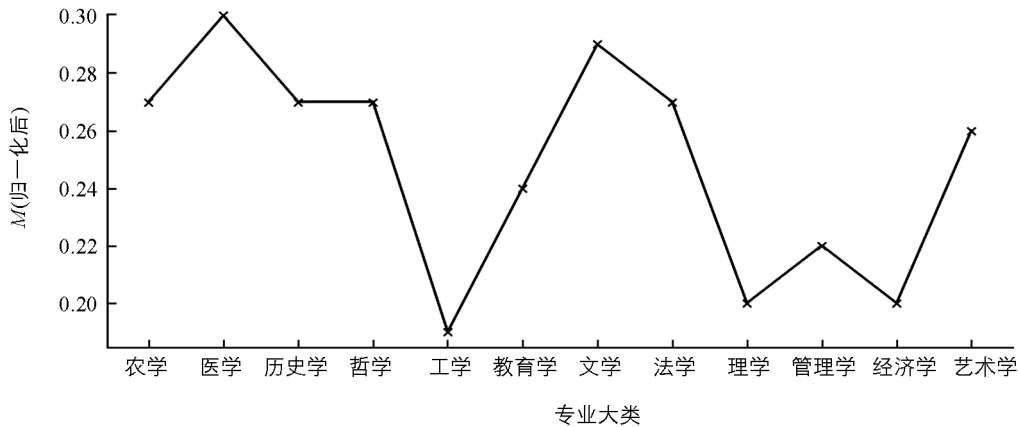


图 5 不同专业大类知识付费金额(M 值, 归一化后)

2) 月均消费额 A 与 F,T,M 的关系分析

月均消费额 A 与 F,T,M 均为定量数据,可使用相关性分析方法,分析这 4 个变量之间是否存在相关性.在进行相关性分析之前,首先对定量数据进行正态性检验,本文采用 K-S 检验,经检验, A,F,T,M 差异有统计学意义($p<0.05$),即均不具有正态性特质.

由于 Pearson 相关系数一般适用于正态数据分布,而以上数据指标均不具有正态性特质,因此,本文采用 Spearman 系数进行相关性分析.

经过计算, A 与 F,T,M 之间的 Spearman 系数依次为 0.049, 0.057, 0.073,说明 A 与 F,T,M 之间存在较弱的相关性(表 2).

表 2 A 与 F,T,M 关系分析

样 本	F	T	M
Spearman 系数	0.049	0.057	0.073

通过以上研究发现,大学生用户的基本属性,如性别、学历、专业均与付费金额等付费购买行为之间存在一定相关关系,与使用频率、使用时长等日常活跃行为不存在显著关系,而月均消费额与付费金额、使用频率、使用时长等使用行为均不存在显著相关关系.因此,知识付费平台在进行营销时,可针对特定人群进行差异化营销.

5 大学生知识付费产品使用行为聚类分析

5.1 聚类结果实现

K-means 算法中, SSE 和轮廓系数是评估 K 值最常用的指标. SSE 即簇内误差平方和,即经过多次迭代后,数据集中每个样本点到各簇质心的距离平方的总和,一般来说, SSE 学习曲线的“肘部” K 值最佳. SSE 公式如公式(2)所示,其中, $x(i)$ 表示样本点 i , $u(j)$ 表示第 j 簇的质心:

$$S_{SSE} = \sum_{i=1}^n \sum_{j=1}^m w(i,j) = \|x(i) - u(j)\|_2^2$$

(2)

轮廓系数结合内聚度和分离度 2 种因素,其取值为 $[-1, 1]$,一般来说,轮廓系数的值越接近 1,聚类越好.轮廓系数公式如公式(3)所示,其中, $a(i)$ 为样本 i 到同簇样本点的平均距离, $b(i)$ 为样本点 i 到其他每个簇距离平均值的最小值^[16]:

$$S(i)=\frac{b(i)-a(i)}{\max(b(i),a(i))}$$

(3)

本文采用 sklearn.cluster 中的 Kmeans 类进行 K-means 聚类分析, 并采用 sklearn.metrics 的 silhouette score 类计算轮廓系数. 然后将 K 值从 2~10 逐次循环, 并绘制不同 K 值下的学习曲线. 由图 6 可知, 当 $K=4$ 时, SSE 达到“肘部”, 且轮廓系数达到最大值, 即表示在 2 种评价指标下, K 取 4 时聚类效果最佳.

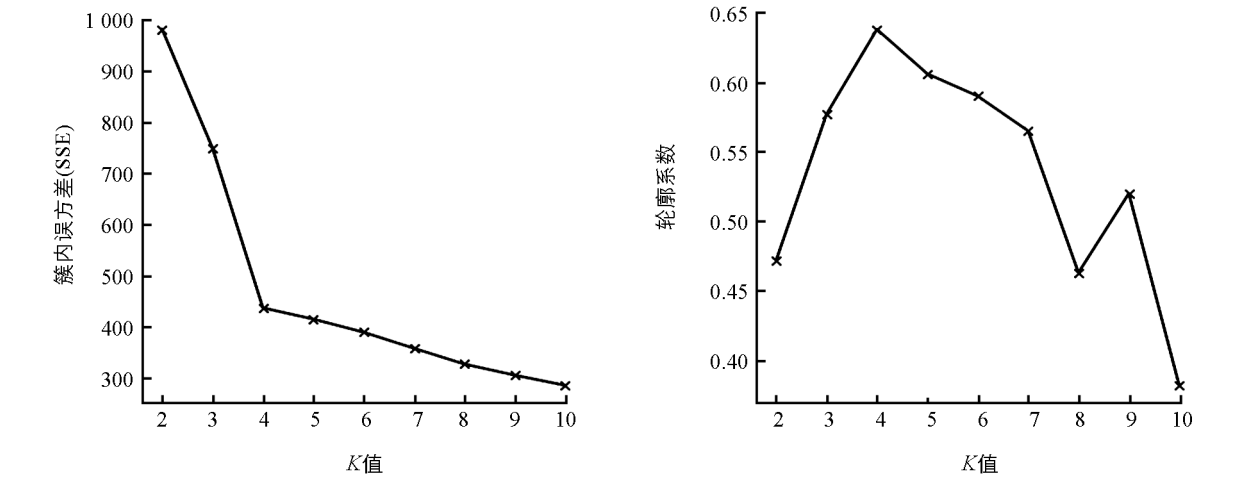


图 6 不同 K 值下的簇内误差差 (SSE) 和轮廓系数

最后, 基于 FTRM 模型和 K-means 算法对清洗后的数据进行聚类, 得到 4 类数据, 然后对每一类数据的 F, T, R, M 值分别计算均值和标准差, 最终所得数据如表 3 所示, 其中样本数量和样本比例以 N 和 P' 两个指标代替.

表 3 每一类数据的 F, T, R, M 值的均值、标准差及 N, P'

样本	cluster_1	cluster_2	cluster_3	cluster_4
	均值(标准差)	均值(标准差)	均值(标准差)	均值(标准差)
F	0.11(0.12)	0.61(0.12)	0.51(0.11)	0.14(0.16)
T	0.24(0.09)	0.47(0.18)	0.39(0.17)	0.29(0.19)
R	0.83(0.19)	0.07(0.14)	0.68(0.10)	0.06(0.10)
M	0.09(0.19)	0.33(0.13)	0.12(0.10)	0.51(0.17)
N	858	610	2192	754
P'	19.44%	13.82%	49.66%	17.08%

注: N 表示样本数量, P' 表示样本比例.

5.2 聚类结果分析

通过对 4 类群体聚类结果的 F, T, R, M 值分析可以发现:

- 1) cluster_1, 此分类群体的 F, T, M 值均较小, R 值最大, 说明该群体在知识付费产品的日常活跃行为和付费购买行为上都表现“较差”, 尚未真正养成使用知识付费产品的习惯或曾经使用知识付费产品, 但基于一些原因已经流失, 属于睡眠型(或流失型)用户, 这类用户占总体比例的 19.44%.
- 2) cluster_2, 此分类群体的 F, T 值均较大, M 值次之, R 值较小, 说明该群体在知识付费产品的日常活跃行为和付费购买行为上都表现“优异”, 已经养成使用知识付费产品的习惯, 属于忠实型用户, 这类用户占总体比例的 13.82%.
- 3) cluster_3, 此分类群体的 F, T, R 值均较大, M 值很小, 说明该群体在知识付费产品的日常活跃行为中表现“较优”, 但付费购买行为表现“较差”, 即经常使用知识付费产品, 但尚未转换成真正的付费用户, 属于未付费转化型用户, 这类用户占总体比例的 49.66%.
- 4) cluster_4, 此分类群体的 F, T, R 值相对较小, M 值最大, 说明该群体在知识付费产品的日常活跃

行为中表现“较差”，但付费购买行为表现“优异”，即经常付费购买知识付费产品，但并未真正有效使用，属于冲动消费型用户，这类用户占总体比例的 17.08%。

5.3 营销建议

1) 睡眠型(或流失型)用户。针对此类用户，需要针对其睡眠或流失的原因进一步深入挖掘，然后制定针对性措施。如首先改善其日常活跃行为，再制定促销措施，刺激其付费购买。

2) 忠实型用户。针对此类用户，可以按照其日常活跃行为、付费购买行为等进一步划分忠实等级，并给予一定的奖励，运用粉丝经济、社群经济的思维进行精细化运营。如 B 站在 UP 主达到一定粉丝数量后，会逐次开放一些特权，如头像加标识、颁发实体勋章等，用以增强用户的认同感；此外，还可以制定一些转发、介绍奖励政策，进一步刺激忠实用户介绍新的用户，如拼购、转发有礼、拉新有礼等。

3) 未付费转化型用户。此类客户是整个用户群体中占比最大的，达到 49.66%，企业应该重点针对此类用户制定营销策略，促进其转化购买，如根据使用时长赠送积分，并采用积分+现金的组合方式优惠购买付费产品；利用 Apriori 算法进行付费产品大数据挖掘，发现关联知识，提高付费单价等。

4) 冲动消费型用户。针对此类用户，应该制定有效的活跃政策，提升其日常活跃行为，加强其学习的效果，增强其有用性体验。否则长期购买但不使用，会使用户产生产品无用、浪费钱等观点，不利于品牌口碑传播，因此需要特别注意。提升用户日常活跃行为的方法有设置定期、连续打卡，并对打卡行为给予奖励；一些付费内容设置学习时间期限，到期前定期提醒；日常签到、学习时长积分，积分换礼等；此外，调研结果显示，部分大学生群体使用知识付费产品的动机之一是满足社交需求，因此，可以建立学习社区，社区选拔 KOL，建立完善的社区管理机制，加强日常管理，增加用户活跃度。

此外，大学生用户的基本属性，如性别、学历、专业均与付费金额等付费购买行为之间存在一定相关关系，因此，知识付费平台在进行营销时，可针对特定人群进行差异化营销。

6 结 语

本文首先对大学生使用知识付费产品的动机和认知进行了详细的分析，通过分析发现，知识付费产品在大学生群体中使用比较普及，从喜好上看，大学生群体最关注的是知识付费平台的内容质量和专业程度。另外通过分析发现，大学生用户的基本属性，如性别、学历、专业与付费金额等付费购买行为之间存在相关关系，与使用频率、使用时长等日常活跃行为之间不存在显著关系，同时发现月均消费额与付费金额、使用频率、使用时长等使用行为之间均不存在显著相关关系。最后，运用本文提出的 FTRM 模型，再采用 K-means 聚类算法，将大学生群体聚类为睡眠型(或流失型)用户、忠实性用户、未付费转化型用户、冲动消费型用户 4 类用户群体，并针对不同用户群体的特点提出差异化的营销建议。

本文基于传统的 RFM 模型，结合知识付费产品的特点，提出了一种改进的 FTRM 模型，但研究中并未对该模型的权重进行深入探讨，因此，模型权重是下一步的研究方向之一。此外，针对睡眠型用户或流失型用户产生的具体原因进一步深入挖掘，也是下一步的研究方向之一。

参考文献：

- [1] 艾媒新经济产业研究中心. 2020 年中国知识付费行业运行发展及用户行为调研分析报告 [R/OL]. 艾媒新经济产业研究中心, (2020-02-14) [2020-06-18]. <https://www.iimedia.cn/c400/69029.html>.
- [2] 艾媒大文娱产业研究中心. 2018-2019 中国知识付费行业研究报告 [R/OL]. 艾媒大文娱产业研究中心, (2019-01-18) [2020-06-18]. <https://www.iimedia.cn/c400/63439.html>.
- [3] 孙晓臻, 董晓, 李颖, 等. 大学生知识付费状况调查与分析 [J]. 知识经济, 2017(19): 162-164.
- [4] 何慧雪, 董治华, 赵慧敏, 等. 知识付费平台用户持续使用意愿影响因素探究——以大学生为例 [J]. 中国市场, 2018(27): 192-193.
- [5] 陈娟, 李金旭. 大学生知识付费产品使用偏好研究——以广州大学城十所高校为例 [J]. 现代传播(中国传媒大学学报), 2019, 41(12): 119-126.

- [6] 刘诗卉, 周世军. 大学生网络知识付费状况调查与分析 [J]. 安徽工业大学学报(社会科学版), 2020, 37(2): 117-119.
- [7] 刘 澜, 陈 晨. 大学生数字阅读付费行为实证研究 [J]. 图书馆理论与实践, 2020(5): 110-114.
- [8] 周 洁. 冰火两重天, 知识付费在国外 [J]. 新民周刊, 2018(2): 34-35.
- [9] XIE W, ZHANG Z H, XIANG M. Research on the Perceived Value of College Students for Knowledge-Paying Products [J]. Journal of Physics: Conference Series, 2019, 1168(3): 032027.
- [10] KHOUJA M, SMITH M A. Optimal Pricing for Information Goods with Piracy and Saturation Effect [J]. European Journal of Operational Research, 2007, 176(1): 482-497.
- [11] 雷 兵, 王巧霞. 国内外知识付费研究进展: 一个文献综述 [J]. 科技和产业, 2019, 19(11): 102-107, 155.
- [12] 杨舒怡. 国内外知识付费平台的对比分析 [J]. 情报探索, 2019(6): 83-89.
- [13] 郑松进, 何鹏飞, 汪思晨, 等. 关于大学生“知识付费”行为的调查与分析 [J]. 科技风, 2020(17): 231-232.
- [14] 任春华, 孙林夫, 吴奇石. 基于 LRFAT 模型和改进 K-means 的汽车忠诚客户细分方法 [J]. 计算机集成制造系统, 2019, 25(12): 3267-3278.
- [15] ROUSSEUW P J. Silhouettes: a Graphical Aid to the Interpretation and Validation of Cluster Analysis [J]. Journal of Computational and Applied Mathematics, 1987, 20: 53-65.
- [16] 董富强, 马 力, 武 波. 一种基于 Internet 的用户行为分类方法与模型的研究 [J]. 现代电子技术, 2004, 27(22): 19-21.
- [17] 杨善林, 王佳佳, 代 宝, 等. 在线社交网络用户行为研究现状与展望 [J]. 中国科学院院刊, 2015, 30(2): 200-215.

Research on the Use Behavior of Knowledge Payment Products of University Students Based on FTRM Model and K-means

WU Xin-yu

Southwest University Library, Chongqing 400715, China

Abstract: China's knowledge payment industry has entered the era of fine operation, and the premise of fine operation is to study the differentiated classification of users. At present, there are few researches on the behavior differentiation classification of university students, which is one of the main user groups of knowledge payment products. Therefore, it is of theoretical significance to study the use behavior of university students' knowledge payment products based on the FTRM model and K-means algorithm. First, this paper analyzes the motivation and cognition of university students to use knowledge paid products, and then explores the relationship between the basic attributes of college students and the use behavior of knowledge paid products. Finally, based on the traditional RFM model and the characteristics of knowledge paid products, an improved FTRM model is proposed. Based on the model, the K-means clustering algorithm is used to cluster college students into four groups: sleeping or losing users, loyal users, unpaid conversion users and impulsive consumption users. Marketing suggestions on business alienation are put forward according to the characteristics of the four different user groups.

Key words: knowledge payment; FTRM model; K-means algorithm; behavior research

责任编辑 崔玉洁