Journal of Business Strategies

Customer Persona Analysis of Online Music Streaming Platforms with HDBSCAN Algorithm: The Case of Spotify Users

DOI: 10.22070/cs.2026.21227.1459

Ehsan Soltanifar [1] *and Navid Mohammadi* [2]

1. *College of Management, University of Tehran, Tehran, Iran. (Email: soltanifar@ut.ac.ir)*
2. *College of Management, University of Tehran, Tehran, Iran. (Corresponding Author)*

*Email: navid.m@ut.ac.ir

Received on: 10/11/2025

Accepted on: 19/01/2026

Abstract

Introduction and Purpose: Persona, a concept originating in academic psychology, was introduced by psychiatrist Carl Gustav Jung and refers to the face or social mask an individual presents to the outside world. With advances in data and personalization technologies, Alan Cooper pioneered the modern use of personas in user experience (UX) design in the 1990s (Prutt & Grudin, 2003). The persona concept has been used in various fields, including software development, health, higher education, marketing, robotics, responsive systems, video games, and system security (Salminen et al., 2022). The persona concept in the field of marketing has been used more in the field of digital marketing and digital marketing strategy, as Wanga (2024) says, the emergence of persona marketing is closely related to the growth of popular social media platforms such as Instagram, TikTok, YouTube, and Douyin, and has also been widely used in studying users of music distribution platforms such as Spotify. Accordingly, the present study examined the personas of Spotify music platform users. Next, three overlapping waves in the field of persona studies were introduced, including (1) qualitative and research-based personas; (2) quantitative approaches and algorithmic clustering; and (3) hybrid or predictive approaches based on machine learning..

Methodology: This study adopted a quantitative, clustering-based approach in an observational-analytical, data-driven manner. It used the HDBSCAN density-based clustering algorithm to discover behavioral patterns and extract customer personas among 8,000 active Spotify users. The data included 10 behavioral and demographic indicators: gender, age, country, subscription type, daily listening time, number of songs played per day, song skip rate, device type, number of ads heard per week, and offline listening. This work was carried out in three stages: data description, data preprocessing, and HDBSCAN clustering.

Finding: By aggregating identified clusters and analyzing user behavior within each, six audience personas were identified on online music streaming platforms.

1. Premium consumers: listeners with high attention span, high daily listening time, high number of songs played, and stricter selection standards that can lead to a moderate to slightly high dropout rate, but this group of customers is accompanied by frequent returns to selected playlists.
2. Background-free customers: music is the “background” of other activities for this group. Low to moderate daily listening time and number of songs played, low dropout rate (due to semi-active listening), high ad exposure (free subscription type), and low offline listening. Their listening behavior is related to the context of their daily life (work, study, household chores).

Journal of Business Strategies Shahed University Thirtysecond Year No.26 Autumn & Winter 2025-26

Journal of Business Strategies

Received on: 10/11/2025

Accepted on: 19/01/2026

Journal of Business Strategies Shahed University Thirtysecond Year No.26 Autumn & Winter 2025-26

3. Mobile Trend Hunters: The main characteristics of this customer group are young, mobile, and sensitive to trends; a high number of songs played, a high skip rate (fast discovery), and average daily listening time.
4. Quality and Offline Traditionalists: The main characteristics of this customer group are older age, high loyalty to established playlists and familiar genres, low skip rate, high daily listening time, high offline listening, and low exposure to advertising due to the tendency towards premium.
5. Fast-paced, ad-enabled youth: This customer group is free-based with a "slice-by-slice" model. This means they have a high number of songs played daily but relatively low daily listening time, a high skip rate, high advertising exposure, and low offline listening.
6. Global Multi-Device Professionals: This user group is characterized by moderate but profound consumption with significant switching between mobile and desktop. This user group has relatively high daily listening time, moderate skip rate, moderate to high offline listening, and high geographic diversity.

Discussion and Coclusion: As we mentioned in previous studies by Fuller et al. (2016), seven personas were identified: active curator, addict, guided listener, discerning listener, isolated user, distrustful, and wanderer or free music explorer. Kim (2016) also identified three personas among music professionals, including: 1- the listener who listens to high-resolution music recreationally as a new hobby, 2- the music seeker who listens to high-resolution music as a way to learn about music, and 3- the artist lover who is excited when his favorite artist releases a high-resolution version of music. Comparing the results of this study with previous studies shows relative correspondences, such that the premium consumer, background music for other tasks, mobile predator, fast-paced advertising, and global multi-device careers from the present study have complete or relative conceptual correspondence with the personas of the listener with a specific moment, music addict, wanderer or free music explorer, guided listener, and active curator from the study of Fuller et al. (2016), respectively. No correspondence was found for the persona of the traditionalist with online quality in this study. On the other hand, the distrustful and isolated user persona in the study by Fuller et al. (2016) was not found in the present study. In general, after examining and describing the personas identified in this study, it can be stated that the distinction of personas is mainly based on three underlying vectors: which include depth of attention (combination of daily listening time, number of songs per day, and song skip rate), level of service (covariation of subscription type with number of ads heard per week and offline listening rate), and consumption context (device type and country). Additional studies using quantitative data-based approaches across other platforms for selling goods and services can lead to the development of more effective marketing strategies. The researchers are ready to conduct similar research and collaborate with other researchers interested in this field.

Keywords: Marketing strategy; Online music streaming platforms; Persona; Platform business models; Clustering.

Journal of Business Strategies

Received on: 10/11/2025

Accepted on: 19/01/2026

Journal of
Business Strategies
Shahed University
Thirtysecond Year
No.26
Autumn & Winter
2025-26

نشریه علمی
راهبردهای
بازرگانی

(دانشور رفتار)

Journal of
Business
Strategies

مقاله پژوهشی
صفحه ۴۸-۳۱

● دریافت مقاله: ۱۴۰۴/۰۸/۱۹
● پذیرش مقاله: ۱۴۰۴/۱۰/۲۹

*Journal of
Business Strategies
Shahed University
Thirtysecond Year
No.26
Autumn & Winter
2025-26*

**نشریه علمی
دانشگاه شاهد
سال سی و دوم – دوره ۲۲
شماره ۲۶
پاییز و زمستان ۱۴۰۴**

تحلیل پرسونای مشتریان سکوهای پخش آنلاین موسیقی با الگوریتم HDBSCAN، مورد مطالعه: کاربران اسپاتیفای

نویسندگان: احسان سلطانی فر [۱] و نوید محمدی [۲*]

۱. استادیار، دانشکدگان مدیریت، دانشگاه تهران، تهران، ایران.
(soltanifar@ut.ac.ir)

۲. استادیار، دانشکدگان مدیریت، دانشگاه تهران، تهران، ایران (نویسنده مسئول).

* Email: navid.m@ut.ac.ir

DOI: 10.22070/cs.2026.21227.1459

چکیده

تحلیل پرسونای مشتریان به‌عنوان یکی از ابزارهای کلیدی در طراحی تجربهٔ کاربری و تدوین استراتژی‌های بازاریابی داده‌محور، نقش فزاینده‌ای در موفقیت سکوهای دیجیتال ایفا می‌کند. با گسترش سکوهای پخش آنلاین موسیقی و افزایش تنوع الگوهای مصرف کاربران، ضرورت بهره‌گیری از رویکردهای مقیاس‌پذیر و تجربی برای درک رفتار شنیداری مخاطبان بیش‌ازپیش احساس می‌شود. هدف این پژوهش، شناسایی و تبیین پرسونای کاربران سکوهای پخش آنلاین موسیقی با استفاده از رویکردی کمی و داده‌محور است. بدین منظور، داده‌های رفتاری و جمعیت‌شناختی ۸۰۰۰ کاربر فعال سکوی اسپاتیفای شامل ده شاخص اصلی، ازجمله سن، جنسیت، کشور، نوع اشتراک، زمان گوش دادن روزانه، تعداد آهنگ‌های پخش‌شده در روز، نرخ رد کردن آهنگ، نوع دستگاه، میزان مواجهه با تبلیغات هفتگی و سطح گوش دادن آفلاین، مورد تحلیل قرار گرفت. برای استخراج الگوهای رفتاری پایدار و تمایز کاربران با تراکم‌های متفاوت، از الگوریتم خوشه‌بندی مبتنی بر تراکم HDBSCAN استفاده شد که بدون نیاز به تعیین تعداد خوشه‌ها از پیش، امکان شناسایی ساختارهای پیچیده داده را فراهم می‌سازد. نتایج خوشه‌بندی به شناسایی شش پرسونای متمایز شامل مشتریان پُرمصرفِ پریمیوم، مشتریان رایگان پس‌زمینه‌ای، شکارچیان ترند موبایلی، سنت‌گراهای باکیفیت و آفلاین، جوانان سریع‌گذرِ تبلیغ‌پذیر و حرفه‌ای‌های چنددستگاهی جهانی انجامید. این پرسوناها بیانگر تفاوت‌های معنادار در عمق توجه، سطح خدمت و بافت مصرف کاربران هستند. مقایسه یافته‌ها با پژوهش‌های پیشین نشان‌دهندهٔ همخوانی مفهومی نسبی با برخی پرسونای شناسایی‌شده در مطالعات قبلی و درعین‌حال، ظهور الگوهای رفتاری جدید مبتنی بر مدل‌های اشتراک، تحرک‌پذیری و استفاده چنددستگاهی است. یافته‌های این پژوهش چارچوبی تجربی و قابل تکرار برای تحلیل پرسونای داده‌محور ارائه می‌دهد و می‌تواند در بخش‌بندی بازار، هدف‌گذاری مشتریان، طراحی تجربهٔ کاربری و تدوین استراتژی‌های بازاریابی دیجیتال در سکوهای داخلی و بین‌المللی مورداستفاده قرار گیرد.

کلیدواژه‌ها: استراتژی بازاریابی، پرسونا، خوشه‌بندی، سکوهای پخش آنلاین موسیقی، کسب‌وکار سکوئی.

مقدمه

در بازار متلاطم و به‌شدت رقابتی امروز، سازمان‌ها همواره به دنبال روش‌های جدید و البته بهینه برای کسب و حفظ مزیت رقابتی هستند و بدین منظور باید محصولی متمایز از رقبا را مبتنی بر الگوهای رفتاری مشتریان هدف ارائه دهند (آقازاده و همکاران، ۱۳۹۴). ازآنجاکه ویژه‌سازی محصول برای تمامی مشتریان ممکن نیست باید این کار برای گروه‌ها و دسته‌هایی از مشتریان صورت گرفته و اصطلاحاً بخش‌بندی بازار صورت گیرد (باقری و یزدانی، ۱۴۰۲). مفهوم پرسونا برای این منظور و دستیابی به یک به‌اصطلاح استراتژی بازاریابی موفق، مؤثر و البته بهینه کارکرد زیادی پیدا کرده است.

پرسونا[1] به‌عنوان یک مفهوم برآمده از حوزهٔ روان‌شناسی آکادمیک، در درجهٔ نخست توسط روان‌پزشک سوئیسی کارل گوستاو یونگ[2] توسعه داده‌شده است. این مفهوم به چهره یا نقاب اجتماعی اشاره دارد که یک فرد به دنیای بیرون ارائه می‌دهد و به‌عنوان نوعی نقش یا هویت برای هدایت انتظارات اجتماعی عمل می‌کند. در این دیدگاه پرسونا درواقع همان نقابی است که فرد برای مواجهه با دیگران به چهره می‌زند و خودِ واقعی خویش را پنهان می‌کند تا به منافع و اهدافش دست یابد (یعقوبی چوبری، ۱۳۹۶). اصطلاح *پرسونا* از لاتین گرفته‌شده است، به معنای نقاب (ماسک) نمایشی و بدین ترتیب پرسونا نقابی است که برای ایجاد تأثیر خاص بر دیگران طراحی‌شده است، درحالی‌که خود واقعی فرد را پنهان می‌کند (همان). این مفهوم از دههٔ ۹۰ میلادی به‌تدریج وارد حوزهٔ بازاریابی هم شده است. مفهوم پرسونا در بازاریابی ریشه در طراحی کاربر محور اولیه و شیوه‌های تقسیم‌بندی بازار دارد، اما به‌طور رسمی‌تر در اواخر قرن بیستم توسعه یافت. تاریخچهٔ پرسوناهای بازاریابی را می‌توان در نیاز کسب‌وکارها به درک بهتر و همدلی با مشتریان خود فراتر از داده‌های جمعیتی سنتی جست‌وجو کرد.

علیرغم محبوبیت روزافزون پرسونای خریدار به‌عنوان ابزاری در تولید محتوای بازاریابی، تحقیقات تجربی بسیار کمی در این زمینه انجام‌شده است (کلپک[3]، ۲۰۱۹). فقدان روش‌شناسی دقیق یا شواهد تجربی در مورد اثربخشی چنین روشی باید نگرانی‌هایی را ایجاد کند یا حداقل استفادهٔ محتاطانه از آن را تشویق کند (همان). عدم اجماع تحقیقاتی فعلی در مورداستفاده از پرسونای خریدار در بازاریابی، فضایی را برای وبلاگ‌ها و مقالات راهنمای انجام کار[4] بی‌دقت، انتزاعی و حدسی فراهم می‌کند. این موارد می‌توانند به‌طور بالقوه جامعه بازاریابان را گمراه کرده و نتایج کمپین‌های غیراستاندارد یا حتی فرضیات نادرست بدتری را برای توسعهٔ محصول ایجاد کنند. نتایج نشان‌دهندهٔ عدم ارتباط نهایی با کاربرد اصلی در توسعهٔ نرم‌افزار، راهنمایی روش‌شناختی بسیار کم و عدم استدلال برای استفاده از این مرحله احتمالاً غیرضروری در فرایند تولید محتوای بازاریابی است.

مفهوم پرسونا در زمینهٔ بازاریابی بیشتر در حوزهٔ بازاریابی دیجیتال و استراتژی بازاریابی دیجیتال کارکرد یافته چنان‌که وانگا[5] (۲۰۲۴) می‌گوید ظهور بازاریابی پرسونا، ارتباط نزدیکی با رشد سکوهای محبوب رسانه‌های اجتماعی مانند اینستاگرام، تیک‌تاک، یوتیوب و دویین[6] (یک شبکهٔ اجتماعی چینی) دارد. هر وب‌گاه کاربران را نه‌تنها به نمایش محصولات، بلکه به نمایش هویت نیز دعوت می‌کند. سلفی‌ها، ویدئوبلاگ‌ها، به‌روزرسانی‌های استوری (داستانک)، پخش زنده و پیشرفت‌های الگوریتمی، هم سازندگان مستقل و هم حساب‌های شرکتی را به متصدیان روایت تبدیل می‌کند که قطعات، ارزش‌ها و احساساتی را به اشتراک می‌گذارند که به یک طرح برندسازی گسترده‌تر کمک می‌کند (همان). کارکرد پرسونا در حوزه‌های مختلف نظیر برندسازی شخصی به‌طور مثال برای مدونا خوانندهٔ آمریکائی (کیل[7]، ۲۰۲۰) یا برای بخش‌بندی بازار نظیر تو[8] و

1. Persona
2. Carl Gustav Jung
3. Klepek
4. howtos
5. Wanga
6. Douyin
7. Qyll
8. Tu

همکاران (۲۰۱۰) و سالمینن[1] و همکاران (۲۰۲۱) متعدد بوده که بدان خواهیم پرداخت ولی پیش از آن لازم است به پیشینهٔ مفهوم پرسونا در دوران معاصر بپردازیم.

در حوزهٔ طراحی تعاملی[2] پرسوناها شخصیت‌هایی خیالی هستند که برای نشان دادن انواع کاربران مختلفی که از یک وب‌گاه یا محصول استفاده می‌کنند، تعریف می‌شوند. کاربرنماها عمدتاً در فرایند طراحی کاربر-محور مورد استناد قرار می‌گیرند (ژانگ[3]، ۲۰۱۹). بر همین اساس و به جهت آن‌که پرسونا از رفتار خریدار یا کاربر نمایندگی می‌کند، واژهٔ کاربرنما به‌عنوان معادل پرسونا نیز به کار می‌رود. به این تعبیر یک کاربرنما، نمایانگر و نمایندهٔ گروهی از کاربران است که در عمل از سیستم استفاده خواهند کرد (پروت و گرودین، ۲۰۰۳). با گذشت زمان، مفهوم پرسونا از طراحی تجربه کاربری به بازاریابی و توسعهٔ محصول مهاجرت کرد، جایی که به ابزاری ارزشمند برای بررسی پروفایل خریداران، گروه‌بندی آن‌ها و برنامه‌ریزی برای ارتباط هدفمند تبدیل شد. با پیشرفت بازاریابی دیجیتال و تجزیه‌وتحلیل داده‌ها، پرسوناها تکامل یافتند تا داده‌های روان‌شناختی و رفتاری عمیق‌تری را در بر بگیرند و استراتژی‌های بازاریابی دقیق‌تر و شخصی‌سازی‌شده‌تری را ممکن سازند. امروزه، پرسوناهای بازاریابی عنصری اساسی در رویکردهای بازاریابی مشتری‌محور هستند و به کسب‌وکارها در دوران تحول دیجیتال کمک می‌کنند تا محتوا، پیشنهادهای محصول و تعاملات مشتری را بر اساس تحقیقات کیفی و کمی ترکیبی، متناسب با گروه‌های مخاطب بخش‌بندی‌شده تنظیم کنند.

در تجارت الکترونیک پرسونای مشتری یا پرسونای کاربران به معنای ایجاد نمایه‌های خاص و قابل اتکا از بخش‌های کلیدی مخاطبان یک بازار است که بر اساس ویژگی‌های‌شان شکل‌گرفته و به طراحی بهینه تجربهٔ کاربری (UX) و استراتژی‌های بازاریابی هدفمند کمک می‌کند. در حوزهٔ سکوهای پخش آنلاین موسیقی، این تحلیل به دلیل تنوع سلایق، رفتارها و نیازهای کاربران اهمیت ویژه‌ای دارد. توسعه و کاربرد پرسونای مشتریان موجب افزایش تعامل کاربر، بهبود رضایت و درنهایت ارتقای نرخ تبدیل کاربران رایگان به کاربران پرداختی می‌شود. امروزه سکوهای استریم موسیقی مانند اسپاتیفای[4]، دیزر[5] و اپل میوزیک[6] با استفاده از تحلیل پرسونای مشتریان، خدمات و تجربه کاربری را شخصی‌سازی می‌کنند که منجر به افزایش وفاداری کاربران و درآمدزایی بالاتر می‌شود. بر اساس تحقیقات، شناخت دقیق از کاربران و درک نیازهای هر شخصیت کاربری کمک می‌کند تا قابلیت‌ها و ویژگی‌های متناسب هر گروه ارائه شود و تبلیغات، پیشنهادها و امکانات ویژه‌ای برای آن‌ها طراحی گردد.

با توجه به روبه رشد بودن اکوسیستم استارتاپی کشورمان و لزوم بررسی استراتژی‌های بازاریابی دیجیتال برای داخل و خارج از کشور، پرداختن به مدل‌های شناسایی پرسونای مشتریان حائز اهمیت بسیاری است. بر همین اساس، این پژوهش می‌کوشد ضمن بررسی ادبیات موضوع و استفاده از تکنیک‌های فرا ابتکاری، نسبت به شناسایی و صورت‌بندی پرسونای مشتریان سکوهای آنلاین اقدام نماید، به‌نحوی‌که قابل تکرار و پیاده‌سازی برای سکوهای داخلی نیز باشد. در این پژوهش، بر سکوهای پخش آنلاین موسیقی تمرکز شد و به جهت دو شاخص مقبولیت و در دسترس بودن اطلاعات به بررسی کاربران اسپاتیفای پرداخته و با الگوریتم جدیدی نسبت به شناسایی پرسونای آنان اقدام شده است. بر این اساس، مسئله اصلی پژوهش حاضر آن است که چگونه می‌توان با بهره‌گیری از الگوریتم‌های خوشه‌بندی مبتنی بر تراکم و بدون پیش‌فرض‌گذاری دربارهٔ تعداد خوشه‌ها، پرسونای کاربران سکوهای پخش آنلاین موسیقی را به‌صورت داده‌محور، تجربی و قابل تکرار شناسایی و تبیین کرد.

باوجود گسترش سریع سکوهای پخش آنلاین موسیقی و افزایش حجم داده‌های رفتاری کاربران، بخش عمده‌ای از پژوهش‌های مرتبط با پرسونای کاربران این سکوها همچنان مبتنی بر روش‌های کیفی، پیمایشی یا گروه‌های کانونی محدود

1. Salminen
2. Interactive design
3. Zhang
4. Spotify
5. Deezer
6. Apple Music

بوده‌اند. این رویکردها اگرچه درک عمیقی از تجربهٔ کاربران فراهم می‌کنند، اما ازنظر مقیاس‌پذیری، تعمیم‌پذیری و کشف الگوهای پنهان رفتاری با محدودیت مواجه‌اند. در این میان، کاربران سکوی اسپاتیفای به‌عنوان یکی از بزرگ‌ترین جوامع کاربران پخش موسیقی در سطح جهانی، با تنوع بالای جمعیت‌شناختی، بافت‌های مصرف متفاوت و الگوهای شنیداری پیچیده، جامعهٔ آماری مناسبی برای تحلیل داده‌محور پرسونای مشتریان به شمار می‌روند. مسئله اصلی پژوهش حاضر آن است که چگونه می‌توان با بهره‌گیری از الگوریتم‌های خوشه‌بندی مبتنی بر تراکم و بدون پیش‌فرض‌گذاری دربارهٔ تعداد یا ساختار خوشه‌ها، پرسونای کاربران سکوهای پخش آنلاین موسیقی را به‌صورت تجربی، قابل تکرار و مبتنی بر داده‌های واقعی شناسایی و تبیین کرد. این خلأ روش‌شناختی، ضرورت انجام پژوهش حاضر را توجیه می‌کند.

مبانی نظری

مفهوم پرسونا در حوزه‌های مختلفی مورداستفاده قرار گرفته که از آن جمله می‌توان به توسعهٔ نرم‌افزار، سلامت، تحصیلات تکمیلی، بازاریابی، رباتیک، سیستم‌های پاسخ‌گو، بازی‌های ویدئویی، امنیت سیستم اشاره کرد (سالمینن و همکاران، ۲۰۲۲). چنان‌که ابتدا توسط آلن کوپر مطرح‌شده و سپس در مطالعاتی نظیر آدلین[1] و همکاران (۲۰۰۱)، بلوم کوئیس و آرولا[2] (۲۰۰۲) و آنتل[3] (۲۰۰۶) ادامه یافت، فرایند تولید و حفظ کیفیت برنامه‌های کاربردی، فریمورک‌ها و سایر اجزای نرم‌افزاری برای شناسایی و طرح پرسوناهایی برای گروه‌های مختلف کاربران به نحو بهتری قابل انجام است. بنابراین، ازجمله کارکردهای پرسونا برای توسعهٔ نرم‌افزار[4] است.

ارائهٔ رویه سامانمند خدمات درمانی و دارو به افراد یا گروه‌هایی از افراد با استفاده از مفهوم پرسونا و تعریف آن قابل انجام است. در این رویکرد، چنانچه توسط پژوهشگرانی نظیر هگبرگ[5] و همکاران (۲۰۰۸)، هندریکس[6] و همکاران (۲۰۱۳) و هنسلی-شینکینجر[7] (۲۰۱۵) به‌کارگیری و تعریف پرسونا در این موارد منجر به در نظر گرفتن شرایط متعدد مؤثر بر پیامدهای سلامت، ازجمله پایبندی به درمان، حمایت مراقب و عوامل فیزیکی خواهد شد. در حوزهٔ آموزش و تحصیلات تکمیلی نیز مطالعات مایر و تالمن[8] (۲۰۱۰)، فرانبرگ[9] و همکاران (۲۰۱۲) و بروکس و گرر[10] (۲۰۱۴) نشان می‌دهد که آموزش دانشگاهی به‌طور خاص در مقطع کارشناسی و بالاتر در سایهٔ استفاده از مفهوم پرسونا با شناسایی دانش‌آموزان غیرسنتی و همچنین کشف چگونگی تعامل و مواجه با ویژگی‌های گروه‌های مختلف دانش‌آموزان با یکدیگر به ترویج فراگیری در آموزش و شکل‌دهی به تجربیات یادگیری مؤثر منجر می‌شود. در مورد طراحی، ساخت و عملیاتی کردن و بهره‌برداری از ربات‌ها برای خدمت به مردم با استفاده از مفهوم پرسونا هم البته پژوهش‌هایی صورت گرفته که از آن جمله می‌توان به مطالعهٔ دوسانتوس[11] و همکاران (۲۰۱۴) در مورد پرسونای رفتاری برای تعامل انسان و ربات، رئا و یانگ[12] (۲۰۱۹) در مورد پرسونا و عملیات از راه دور و ترت[13] و همکاران (۲۰۱۷) در مورد کاربرد پرسونا در طراحی ربات‌های دستیار اشاره کرد.

کاربرد پرسونا در طراحی سیستم‌های گفت‌وگو و پیشنهاددهی آنلاین هم موضوع مطالعات متعددی قرار گرفته است.

1. Adlin
2. Blomquist & Arvola
3. Antle
4. Software development
5. Högberg
6. Hendriks
7. Hensely-Schinkinger
8. Maier & Thalmann
9. Frauenberger
10. Brooks & Greer
11. Dos Santos
12. Rea & Young
13. Threatt

سیستم‌های گفت‌وگو[1] یک برنامه رایانه‌ای است که می‌تواند از طریق متن، گفتار، لمس و سایر ابزارهای ارتباطی، با انسان‌ها در هر دو کانال ورودی و خروجی تعامل داشته باشد (اکنر-کوهلر و رنجبر[2]، ۲۰۱۶). علاوه بر این، در سیستم‌های پیشنهاددهنده[3] نیز مفهوم پرسونا برای فیلترینگ اطلاعات به کار می‌رود که هدف آن پیش‌بینی امتیاز یا ترجیح کاربر برای یک مورد خاص است. در این رابطه، پژوهش‌های محدودی نظیر داروادا[4] و همکاران (۲۰۰۷)، فریرا[5] و همکاران (۲۰۱۶) و بوناردل[6] و پیشات (۲۰۲۰) در هر دو مورد تحقیقات بیشتری برای شناسایی مزیت اصلی پرسوناها فراتر از درک عمومی کاربر مورد نیاز است.

طراحی بازی‌های رایانه‌ای و بازی‌های الکترونیکی از یک رابط کاربری یا دستگاه ورودی برای ایجاد بازخورد بصری برای اهداف سرگرمی استفاده می‌کند. مزیت اصلی پرسونا: بررسی شاخص‌های دینامیک بازی و آرکتایپ‌ها برای بهبود تجربیات بازی. علاوه بر مطالعات کانوسا و دراچن[7] (۲۰۰۹)، کاپینن[8] و همکاران (۲۰۱۶) و لین[9] و همکاران (۲۰۲۰) تحقیقات بیشتر نیز برای شناسایی مزایای اضافی پرسونا فراتر از درک عمومی کاربر مورد نیاز است. پیاده‌سازی اقدامات حفاظتی برای محافظت در برابر دسترسی ناخواسته به سیستم‌های دیجیتال و داده‌های محرمانه کاربر نیز با رجوع به پرسوناها و حتی بررسی پرسونا مهاجمین میسر خواهد شد (آزتنی[10] وهمکاران، ۲۰۱۱) (کیم[11] و همکاران، ۲۰۱۹). مزیت اصلی پرسوناها در این رابطه عبارت‌اند از آگاه‌سازی عوامل خطر برای حملات امنیت سایبری و حریم خصوصی برای طراحان نرم‌افزار به‌منظور ایجاد اقدامات حفاظتی است که شامل شناسایی ویژگی‌های مهاجم و همچنین ویژگی‌های کاربر می‌شود (استول[12] و همکاران، ۲۰۰۷). اما مهم‌ترین کارکردهای مفهوم بازاریابی ازجمله در حوزهٔ بازاریابی و به‌ویژه بازاریابی دیجیتال، فرایندی است که از طریق آن یک شرکت، خرید یا فروش یک محصول یا خدمات را به گروه یا گروه‌هایی از مشتریان با توجه به نوع و نیاز هر مشتری ترویج می‌دهد (لی[13] و همکاران، ۲۰۱۰) (جنسن[14] و همکاران، ۲۰۱۷). مزیت اصلی پرسونا در این فرایند ایجاد وحدت رویه در بین کارکنان بازاریابی نسبت به مشتریان خود برای اطمینان از هدف‌گیری مؤثرتر عملیات بازاریابی است (چن[15] و همکاران، ۲۰۱۱). مطالعات متعددی در کاربرد پرسونا برای بخش‌بندی بازارها صورت گرفته است.

در حوزهٔ بازاریابی پرسوناها ـ شخصیت‌های خیالی مبتنی بر مطالعهٔ بازار هستند که بخش‌های مختلف کاربران را نمایندگی و نمایان می‌کنند (جانسن[16] و همکاران، ۲۰۲۰). چنان‌که پیش‌تر اشاره شد ابتدا در حوزهٔ تعامل انسان و رایانه (HCI[17]) پدید آمدند و سپس بازاریابان آن را برای انسان‌محور کردن و عملیاتی‌سازی مخاطبان هدف به کار گرفتند. ادبیات اولیه، پرسونا را به‌عنوان ابزاری کیفی برای همدلی و تسهیل ارتباط در طراحی معرفی کرد؛ اما پژوهشگران بازاریابی به‌سرعت این مفهوم را برای نمایه‌سازی خریداران و طراحی پیام‌های بازاریابی اقتباس کردند، زیرا پرسونا نیازهای پنهان مشتری را به شکل صریح و قابل اقدام برای گروه‌های محتوا و تبلیغات آشکار می‌سازد.

از منظر روش‌شناسی، ساخت پرسونا در بازاریابی دیجیتال اکنون در طیفی از الگوهای دستی مبتنی بر مصاحبه تا

1. Dialogue systems
2. Akner-Koler & Ranjbar
3. Recommendation systems
4. Dharwada
5. Ferreira
6. Bonnardel & Pichot
7. Canossa & Drachen
8. Kauppinen
9. Laine
10. Atzeni
11. Kim
12. Stoll
13. Lee
14. Jensen
15. Chen
16. Jansen
17. Human-computer interaction (HCI)

پرسونـاهای داده‌محور و پیش‌بینانه مبتنی بر ردپای دیجیتال گسترده قرار دارد (پارک و کانگ[1]، ۲۰۲۲). مرورهـای نظـام‌منـد شامل مطالعات تو و همکاران (۲۰۱۰) و سالمینن و همکاران (۲۰۲۱) سه موج هم‌پوشان را در این حوزه نشـان مـی‌دهنـد: ۱. پرسونـاهای کیفی و پژوهش‌محور؛ ۲. رویکردهای کمی و خوشه‌بندی الگوریتمی و ۳. رویکردهای ترکیبی یا پـیش‌بینانـه کـه یادگیری ماشینی را با اعتبارسنجی کیفی تلفیق می‌کنند تا پرسونـاهایی مقیاس‌پذیر و ازنظر رفتاری پیش‌بین ایجاد کنند. بنابراین به نظر می‌رسد که با ظهور رویکردهای کمی مبتنی بر کلان داده، موضوع پرسوناهای داده محور[2] نیز اهمیت زیادی یافته‌اند. در این رویکرد با استفاده از تحلیل حجم زیادی از داده‌های مربوط به مشتریان تـلاش مـی‌شـود آن‌هـا را خوشـه‌بنـدی کـرده و پرسوناهای‌شان را شناسایی کنیم. حرکت به‌سوی پرسونـاهای داده‌محور، مسئله مقیاس‌پذیری در تبلیغات برنامه‌ریـزی‌شـده و شخصی‌سازی را حل می‌کند، اما چالش‌هایی مانند تفسیرپذیری و اعتبار چهره‌ای[3] را به همـراه دارد کـه همچنـان موردتوجـه پژوهشگران است (همان).

در عمل، از پرسونـاها در بازاریابی دیجیتال برای طراحی راهبردهای محتوایی بخش‌بنـدی‌شـده، بـومی‌سـازی خلاقیـت تبلیغاتی، ترسیم مسیر مشتری و به‌عنوان ورودی در سیستم‌های توصیه‌گر و هدف‌گیری استفاده می‌شود. چنـان‌کـه در مطالعـهٔ مروری آلوزگومز و میسـن[4] (۲۰۲۳) نشـان داده‌شـده اسـتفاده از پرسـونا و نمـایش پرسـونا بـرای بخـش‌بنـدی مشـتریان و سفارشی‌سازی عملیات بازاریابی الکترونیک، در تجارت الکترونیک کارکرد دارد. مرور پژوهش‌هـای بخـش‌بنـدی مشـتری و شخصی‌سازی در تجارت الکترونیکی نشان می‌دهد که نمایش‌های مبتنی بر پرسونـا ـ چه به‌صورت دستی و چه الگوریتمی ـ زمانی که به‌درستی اعتبارسنجی و در جریان تحلیل‌های بازاریابی ادغام شوند، اثربخشی پیام و ارتباط با مخاطـب را افـزایش می‌دهند. بااین‌حال، پژوهش‌های تجربی هشدار می‌دهند که استفاده سطحی یا غیرمستند از پرسونـا (مانند قالب‌های آماده[5] یـا الگوهای خریدار خیالی در وبلاگ‌ها[6]) می‌تواند فرضیات نادرست و نتایج هدف‌گیری ضعیف ایجاد کند (آلوزگومز و میسـن، ۲۰۲۳).

چالش‌ها و مسیرهای آینده در این حوزه عمدتاً بـه اعتبارسـنجی، حـریم خصوصـی و یکپارچـه‌سـازی عملیـاتی مربـوط می‌شوند. پژوهشگران خواستار تدوین پروتکل‌های قوی‌تر اعتبارسنجی (مانند آزمون‌های رفتاری نگه‌داشته شده[7] و سنجه‌های دقت پیش‌بینی[8]) و روش‌هایی هستند که فرایند ساخت پرسونـا را برای ذی‌نفعان شـفاف سـازند. رویکردهـای نـوینی ماننـد پرسونـاهای پیش‌بین[9] (هشو[10] و همکاران، ۲۰۲۵) و چارچوب‌های تبلیغات آنلاین مبتنـی بـر پرسونـا[11] (ملیـک[12]، ۲۰۱۹) می‌کوشند ساخت پرسونـا را مستقیماً به شاخص‌های کلیدی عملکرد (KPIs[13]) متصل کنند. درعین‌حال، ادبیات بر ملاحظات مربوط به حریم خصوصی، قوانین (مانند GDPR[14] و CCPA[15]) و چالش‌های اخلاقی در استفاده از داده‌های حساس دیجیتال تأکید می‌کند؛ توازن میان منافع شخصی‌سازی و الزامات کاهش داده و شفافیت، همچنان از محورهای پژوهشـی فعـال اسـت. درمجموع، شواهد علمی نشان می‌دهد که پرسونـاها همچنان ابزاری مفهومی و عملی مفید برای بازاریابـان دیجیتـال هسـتند،

1.Park & Kang
2.Data-Driven Persona
3.Face Validity
4.Alves Gomes & Meisen
5. ungrounded templates
6. unvalidated “buyer persona” blog heuristics
7 . holdout behavioral tests
8. predictive accuracy benchmarks
9. predictive personas
10. Hsu
11. persona-based online advertising frameworks
12. Malik
13. Key Performance Indicators
14.General Data Protection Regulation
15. California Consumer Privacy Act

مشروط بر آن‌که با دقت علمی ساخته، به‌طور مستمر اعتبارسنجی و در کنار تحلیل‌های مبتنی بر حفظ حریم خصوصی به کار گرفته شوند (هشو و همکاران، ۲۰۲۴).

جمع‌بندی ادبیات نظری نشان می‌دهد که حرکت از پرسونای کیفی و شهودی به سمت پرسونای داده‌محور، پاسخی مستقیم به چالش‌های مقیاس‌پذیری، تکرارپذیری و اتکاپذیری در تصمیم‌گیری‌های بازاریابی دیجیتال است. در این چارچوب، الگوریتم‌های خوشه‌بندی مبتنی بر تراکم، به‌ویژه HDBSCAN، به دلیل عدم نیاز به تعیین تعداد خوشه‌ها از پیش و توانایی شناسایی ساختارهای پیچیده رفتاری، بستر روش‌شناختی مناسبی برای استخراج پرسونای کاربران از داده‌های واقعی سکوهای دیجیتال فراهم می‌کنند. این پیوند مفهومی، مبنای انتخاب روش پژوهش حاضر را شکل می‌دهد.

پیشینهٔ پژوهش

مطالعات مختلفی در حوزهٔ تحلیل پرسونای کاربران سکوهای پخش موسیقی صورت گرفته است که به تشریح انواع مختلف پرسونای کاربران و ویژگی‌های رفتاری آن‌ها پرداخته‌اند. به‌عنوان‌مثال، در پژوهشی که توسط فولر[1] و همکاران (۲۰۱۶) انجام‌شده، با استفاده از یک روش غربالگری و ترکیبی از سؤالات و پاسخ‌ها (به شیوه‌ای عمدتاً کیفی) هفت پرسونای مختلف شناسایی شده‌اند که هر یک ویژگی‌ها و عادات شنیداری خاص خود را دارند؛ متصدی فعال[2]، معتاد[3] (به موسیقی)، گوش دهندهٔ هدایت‌شده[4]، شنوندهٔ با ذائقهٔ خاص[5]، کاربر منزوی[6]، بی‌اعتماد[7] و سرگردان یا کاوشگر موسیقی آزاد[8].

۱. متصدی فعال؛ فعالانه به دنبال موسیقی جدید است و از انتخاب لیست پخش لذت می‌برد.
۲. معتاد (به موسیقی)؛ جست‌وجوگر آیتم‌های شناخته‌شده است و موسیقی‌های سهل‌الوصول را ترجیح می‌دهد.
۳. گوش دهندهٔ هدایت‌شده؛ ترجیح می‌دهد کنترل را به سرویس پخش آنلاین بسپارد.
۴. شنوندهٔ با ذائقهٔ خاص؛ زمان قابل‌توجهی را صرف یادگیری و کشف موسیقی می‌کند.
۵. کاربر منزوی؛ معتقد است گوش دادن به موسیقی یک تجربهٔ شخصی است.
۶. بی‌اعتماد؛ نسبت به توصیه‌های سرویس‌های موسیقی بدبین است.
۷. سرگردان یا کاوشگر موسیقی آزاد؛ از کشف اتفاقی موسیقی و ژانرهای مختلف لذت می‌برد.

در مقالهٔ دیگری با رویکرد گروه کانونی[9] در همین سال نیز ایم و کیم[10] (۲۰۱۶) به شناسایی سه پرسونا در میان شوندگان تخصصی موسیقی پرداخته‌اند. آن‌ها سه شخصیت از شوندگان بالقوهٔ موسیقی با وضوح‌بالا (باکیفیت)[11] ساخته‌اند که عبارت‌اند از: ۱. شنونده‌ای که به‌صورت تفریحی به موسیقی با وضوح‌بالا به‌عنوان یک سرگرمی جدید گوش می‌دهد، ۲. جست‌وجوگر موسیقی که به موسیقی با وضوح‌بالا به‌عنوان راهی برای شناخت موسیقی گوش می‌دهد و ۳. عاشق هنرمند که وقتی هنرمند مورد علاقه‌اش نسخه‌ای با وضوح‌بالا از موسیقی منتشر می‌کند، به آن گوش می‌دهد.

این‌گونه دسته‌بندی‌ها نشان می‌دهد که کاربران موسیقی آنلاین ویژگی‌های بسیار متنوع و گاه متناقضی دارند که طراحی خدمات باید آن‌ها را پوشش دهد. به‌علاوه، مطالعات تأکید دارند که بسیاری از کاربران در شرایط و زمان‌های مختلف، ویژگی‌های چندین پرسونای متفاوت را از خود بروز می‌دهند که پیچیدگی تحلیل و طراحی تجربیات شخصی‌سازی‌شده را

1. Fuller
2. Active Curator
3. Addict
4. Guided Listener
5. Music Epicurean
6. Music Recluse
7. Non-Believer
8. Wanderer
9. Focus Group
10. Im & Kim
11. High-Resolution Music Users

افزایش می‌دهد (جانسن و همکاران، ۲۰۲۰).

مطالعهٔ صورت گرفته توسط باراتا و کوئلهو[1] (۲۰۲۱) با استفاده از مدل گسترش‌یافته UTAUT2[2] به بررسی عوامل مؤثر بر تصمیم کاربران برای استفاده از نسخه‌های پرداختی سکوهای موسیقی آنلاین پرداخته است. نتایج این تحقیق نشان داد که عوامل کلیدی مانند عادت مصرف، انتظارات عملکردی، ارزش قیمت، شخصی‌سازی محتوا و نگرش به سرقت موسیقی به‌طور معناداری بر قصد خرید و وفاداری به سکو تأثیر دارند. ژائو[3] (۲۰۲۴) در قالب یک رویکرد مبتنی بر پیمایش از طریق نظرسنجی از ۶۰۰ کاربر اسپاتیفای، سه بخش مجزا از پرسونای کاربران را شناسایی کرده که عبارت‌اند از: اعضای رایگان (۵۹ درصد)، اعضای رایگان که مایل به ارتقا به عضویت ویژه هستند (۲۴ درصد) و اعضای ویژه (۱۷ درصد). تجزیه‌وتحلیل رگرسیون لجستیک، عوامل کلیدی مؤثر بر تبدیل عضویت ویژه، ازجمله طرح فعلی، قیمت طرح ویژه و سن را نشان می‌دهد. درنهایت، این پژوهشگر برای درک عمیق‌تر از هر بخش کاربر، سه شخصیت یا پرسونا ایجاد کرد که به‌اصطلاح نیا[4] (عضو رایگان)، آریا (عضو رایگان با تمایل به ارتقاء) و سینا (عضو ویژه) برچسب‌گذاری شده‌اند. شخصی‌سازی به‌عنوان یکی از ابعاد مهم این مدل معرفی‌شده است که به کاربران امکان می‌دهد بر اساس سلیقه و نیازهای شخصی تنظیمات تجربهٔ کاربری خود را تغییر دهند و این امر نقش محوری در افزایش رضایت و وفاداری آن‌ها دارد. همچنین، ایجاد امکان تعامل اجتماعی و قابلیت‌های اشتراک‌گذاری، هرچند که در بعضی تحقیقات تمایل کاربران به استفاده از این قابلیت‌ها پایین گزارش‌شده است، می‌تواند به تقویت حس تعلق جمعی و تعامل بیشتر کاربری کمک کند.

بررسی ادبیات موضوع، به‌ویژه در حوزهٔ تحلیل پرسونا برای بهبود تجربهٔ کاربری و استراتژی بازاریابی نشان می‌دهد که تحقیقات صورت گرفته عمدتاً کیفی با استفاده از روش‌هایی نظیر مصاحبه یا گروه کانونی یا کمی از نوع پیمایش بوده‌اند و نیاز به انجام مطالعات مرتبط با استفاده از الگوریتم‌های فرا ابتکاری کماکان احساس می‌شود. بنابراین، تحقیق حاضر در ادامه با استفاده از الگوریتم خوشه‌بندی HDBSCAN پیشنهادی توسط کمپلو[5] و همکاران (۲۰۱۳) به بررسی و تحلیل پرسونای مشتریان در سکوهای موسیقی پرداخته است. با استفاده از تحلیل پرسونای مشتری، سکوهای موسیقی می‌توانند استراتژی‌های بازاریابی توسعه‌یافته، طراحی تجربه کاربری هدفمند و بهبود محتوای پیشنهادی را به‌کارگیرند. این تحلیل به توسعه‌دهندگان کمک می‌کند تا خدمات و رابط‌های متناسب با نیازهای کاربران مختلف را طراحی کرده و نرخ جذب و نگه‌داری مشتریان را افزایش دهند. همچنین، شناخت بهتر از پرسونای مخاطبان موجب بهبود تبلیغات هدفمند شده و بهینه‌سازی منابع مالی را برای شرکت‌ها به دنبال دارد. درنهایت، پرسونای مشتریان در شناخت بهتر رفتار مصرف، توسعهٔ ویژگی‌های جدید، بهبود کیفیت خدمات و افزایش رضایت و وفاداری کاربران نقشی راهبردی ایفا می‌کند که منجر به رشد پایدار سکوهای موسیقی آنلاین می‌گردد.

روش‌شناسی پژوهش

این پژوهش با رویکرد مشاهده‌ای-تحلیلی و مبتنی بر داده انجام‌شده است و از الگوریتم خوشه‌بندی مبتنی بر تراکم HDBSCAN برای کشف الگوهای رفتاری و استخراج پرسونای مشتریان در میان ۸۰۰۰ کاربر فعال اسپاتیفای استفاده می‌کند. داده‌های مورد استفاده شامل ده شاخص رفتاری و جمعیت‌شناختی است: جنسیت، سن، کشور، نوع اشتراک، زمان گوش دادن روزانه، تعداد آهنگ‌های پخش‌شده در روز، نرخ رد کردن آهنگ، نوع دستگاه، تعداد تبلیغات شنیده‌شده در هفته و میزان گوش دادن آفلاین. در ادامه، اجزای داده‌ها، مراحل پیش‌پردازش، مهندسی ویژگی، کاهش بُعد، نحوهٔ راه‌اندازی و تنظیم HDBSCAN، ارزیابی پایداری و اعتبار خوشه‌ها، فرایند تبدیل خوشه‌ها به پرسوناهای مشتری و همچنین ملاحظات مربوط به

1. Barata & Coelho
2. Unified theory of acceptance and use of technology
3. Zhao
4. Nia (free member), Aria (free member with willingness to upgrade), and Sienna (premium member)
5. Campello

بازتولیدپذیری و اخلاق پژوهش به‌تفصیل تشریح می‌شود.

مرحلهٔ اول؛ توصیف داده‌ها

داده‌های مورداستفاده در این پژوهش، داده‌های مربوط به ۸۰۰۰ کاربر فعال در سکوی اسپاتیفای، به‌عنوان یکی از بزرگ‌ترین سکوهای پخش آنلاین موسیقی است. توزیع کاربران برحسب نوع دستگاه، شامل دسکتاپ ۲۷۷۸ نفر (۳۴٫۷ درصد)، تلفن همراه ۲۵۹۹ نفر (۳۲٫۵ درصد) و وب‌گاه ۲۶۲۳ نفر (۳۲٫۸ درصد) است. توزیع نوع اشتراک در کل نمونه (از جمع سه دستگاه) شامل اشتراک خانواده ۱۹۰۸ نفر (۲۳٫۹ درصد)، رایگان ۲۰۱۸ نفر (۲۵٫۲ درصد)، پریمیوم ۲۱۱۵ نفر (۲۶٫۴ درصد) و دانشجویی ۱۹۵۹ نفر (۲۴٫۵ درصد) است. توزیع جنسیت در کل نمونه نیز شامل زن ۲۶۵۹ نفر (۳۳٫۲ درصد)، مرد ۲۶۹۱ نفر (۳۳٫۶ درصد) و سایر ۲۶۵۰ نفر (۳۳٫۱ درصد) است. این دسته‌بندی داده‌ها نشان می‌دهد نمونه از منظر نوع دستگاه، نوع اشتراک و جنسیت متوازن و نزدیک به هم است و برای تحلیل‌های خوشه‌بندی، نیازی به وزن‌دهی تصحیحی ضروری مشاهده نشده است. کشور، سن، زمان گوش دادن روزانه، تعداد آهنگ‌های روزانه، نرخ رد کردن آهنگ، تعداد تبلیغات شنیده‌شده در هفته و میزان گوش دادن آفلاین به‌صورت مقیاس‌های عددی و پیوسته/نسبی نیز برای ۸۰۰۰ کاربر مورداستفاده قرار گرفته است. در صورت وجود مقادیر پرت یا خارج از دامنهٔ منطقی (برای مثال نرخ رد کردن خارج از [۰،۱] یا زمان‌های منفی)، قواعد پاک‌سازی که در بند پیش‌پردازش ذکر می‌شود، اعمال گردید. کاربران دارای شناسه تکراری، رکوردهای ناقص در چند متغیر کلیدی (مانند نوع اشتراک، نوع دستگاه یا نرخ رد کردن) که با روش‌های استاندارد قابل‌ترمیم نبود و همچنین رکوردهایی با تناقض منطقی (مثلاً زمان گوش دادن صفر، همراه با تعداد آهنگ بالا) از تحلیل کنار گذاشته شدند و در موارد قابل‌ترمیم، با قواعد توصیف‌شده در بخش پیش‌پردازش عمل شد.

مرحلهٔ دوم؛ پیش‌پردازش داده‌ها

به‌منظور پیش‌پردازش داده‌ها، مراحل زیر انجام‌شده است:

- یکپارچه‌سازی و پاک‌سازی: کنترل یکتایی کاربر، حذف رکوردهای تکراری، یکسان‌سازی قالب مقادیر طبقه‌ای (حساسیت به حروف بزرگ/کوچک و فاصله)، اعتبارسنجی دامنه‌ها (سن در بازهٔ معقول بزرگسالی، نرخ رد در [۰،۱]، عدم منفی بودن زمان‌ها و شمارش‌ها).
- مدیریت مقادیر گمشده: برای متغیرهای عددی از میانه (به‌جای میانگین، به دلیل تاب‌آوری در برابر چولگی و پرت‌ها) و برای طبقه‌ای‌ها از «رایج‌ترین» یا برچسب نامشخص استفاده شد.
- مهار داده‌های پرت: برای متغیرهایی با چولگی بالا (مانند زمان گوش دادن روزانه و تبلیغات شنیده‌شده) از وینسورایز ۱ و ۹۹ درصد یا تبدیل یا–جانسون/لگاریتمی استفاده شد تا اثر نقاط دورافتاده بر فاصله‌ها کاهش یابد.
- نرمال‌سازی/استانداردسازی: متغیرهای عددی با استانداردساز نرمال (میانگین صفر، واریانس یک) مقیاس‌بندی شدند و در صورت حضور پرت‌های شدید، از مقیاس بند مقاوم[1] استفاده شد.
- کدگذاری متغیرهای طبقه‌ای: جنسیت، نوع اشتراک، نوع دستگاه و کشور با کدگذاری وان–هات[2] تبدیل شدند. برای کشورها، گروه‌بندی به «۱۰ کشور/بازار پرتکرار + سایر» به‌منظور کنترل ابعاد/چولگی اعمال شد.

کاهش بُعد و کنترل هم‌خطی برای کاهش همبستگی ویژگی‌ها و تثبیت هندسه فضا پیش از HDBSCAN، از تحلیل مؤلفه‌های اصلی[3] استفاده شد. مؤلفه‌ها تا جایی نگه‌داشته شدند که حداقل ۸۵–۹۰ درصد واریانس کل را پوشش دهند. این کار ضمن تسریع محاسبات، اثر ویژگی‌های کم اطلاعات و هم‌خط بر معیار فاصله را کاهش داد. برای تصویرسازی دوبُعدی خوشه‌ها در نتایج (شکل ۳)، از نگاشتی مجزا UMAP بر فضای استاندارد/کاهش‌یافته استفاده شد تا مرزبندی بصری و تراکم هسته‌ها بهتر دیده شود؛ اما خوشه‌بندی اصلی در فضای PCA انجام شد.

1 . RobustScaler
2. One-Hot
3. PCA

مرحله سوم؛ خوشه‌بندی HDBSCAN

الگوریتم خوشه‌بندی (HDBSCAN) و تنظیمات الگوریتم HDBSCAN به دلیل توانایی شناسایی خوشه‌های با تراکم متفاوت، عدم نیاز به تعیین تعداد خوشه‌ها از پیش و قابلیت برچسب‌گذاری نقاط نویز انتخاب شد. تنظیمات کلیدی به‌صورت زیر تعیین و داده‌محور بهینه‌سازی شد:

- اندازه کمینه خوشه[1]: با جست‌وجو در بازه‌ای متناسب با اندازه نمونه (مثلاً ۵۰ تا ۲۰۰) و انتخاب مقداری که توازن مناسبی میان «تفکیک‌پذیری» و «پایداری» ایجاد کند.
- حداقل نمونه[2]: در اغلب سناریوها برابر یا نزدیک به‌اندازهٔ کمینهٔ خوشه قرار داده شد تا کنترل محافظه‌کارانه‌تری بر نویز اعمال شود.
- معیار فاصله: در فضای مؤلفه‌های اصلی، فاصلهٔ اقلیدسی به‌عنوان پیش‌فرض به کار رفت؛ در صورت مشاهده ساختارهای خطی-ریز یا داده‌های با بردارهای پراکنده، فاصلهٔ کسینوسی نیز آزمون شد.

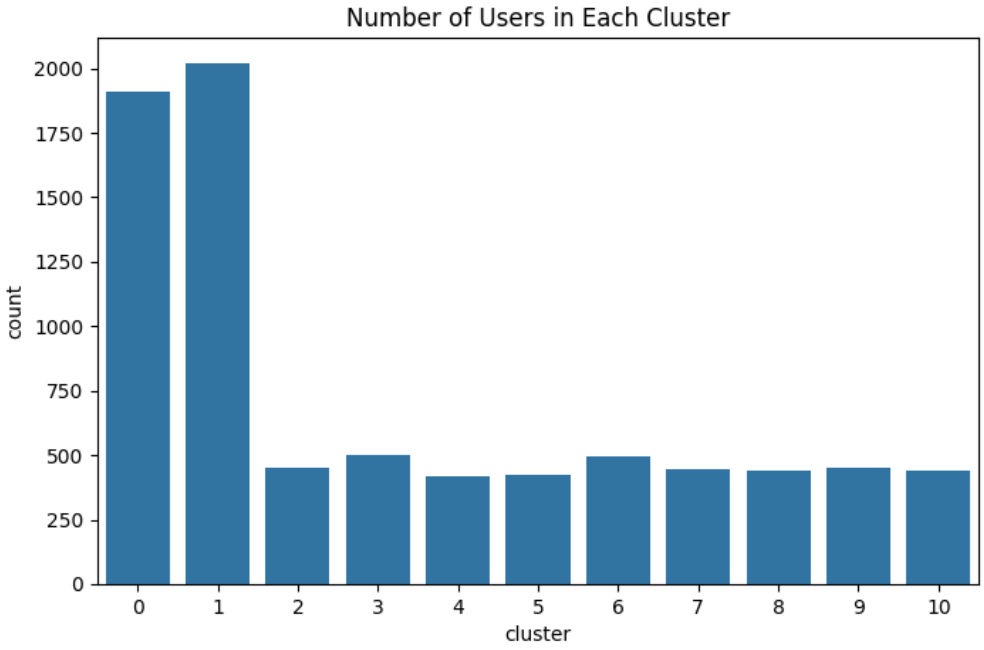
به‌منظور ارزیابی پایداری و اعتبار خوشه‌ها برای اعتبارسنجی نتایج خوشه‌بندی (در غیاب برچسب‌های درست/نادرست)، مجموعه‌ای از سنجه‌های درونی و روش‌های پایداری به‌کاررفته است:

۱. شاخص اعتبار مبتنی بر تراکم DBCV[3] برای انتخاب ترکیب مناسب از پارامترهای HDBSCAN و تعداد مؤلفه‌های PCA.
۲. پایداری خوشه‌ها با بوت‌استرپ طبقه‌ای: نمونه‌گیری مجدد با حفظ نسبت‌های «نوع دستگاه × نوع اشتراک × جنسیت» بر اساس جدول توصیفی و اجرای مجدد خوشه‌بندی؛ سپس سنجش شباهت خوشه‌ها با جاکارد/آراند تعدیل‌شده است.
۳. پایداری در برابر اغتشاش: افزودن نویز کوچک گاوسی به ویژگی‌های عددی (در حد ۱–۲ درصد مقیاس) و بررسی ثبات برچسب‌ها.
۴. تمایز آماری خوشه‌ها: آزمون کروسکال-والیس برای تفاوت میانگین/میانه در متغیرهای عددی بین خوشه‌ها و آزمون کای-دو برای تفاوت توزیع‌ها در متغیرهای طبقه‌ای (جنسیت، نوع اشتراک، نوع دستگاه، کشور).
۵. تبدیل خوشه‌ها به پرسونا: پس از تثبیت خوشه‌ها، برای هر خوشه پروفایل توصیفی ساخته شد: میانه، چارک‌ها و درصدهای افراطی برای متغیرهای عددی (زمان گوش دادن روزانه، تعداد آهنگ‌های روزانه، نرخ رد کردن، تبلیغات هفتگی، آفلاین) و توزیع درصدی برای متغیرهای طبقه‌ای (جنسیت، نوع اشتراک، نوع دستگاه، کشور). سپس «امضای رفتاری» هر خوشه با اتکا به نسبت‌های کلیدی (مثلاً نسبت زمان گوش دادن به تعداد آهنگ در حضور نرخ رد) استخراج شد. نام‌گذاری پرسوناها بر پایهٔ برجسته‌ترین ویژگی‌های تفکیک‌گر (سطح خدمت: رایگان/پریمیوم؛ عمق توجه: تمرکز/اکتشاف؛ بافت مصرف: موبایل/چنددستگاهی و زمینه جمعیت‌شناختی) انجام گرفت. برای جلوگیری از برچسب‌گذاری شتاب‌زده، از ماتریس شباهت معنایی میان خوشه‌ها و مرور متقاطع شاخص‌های کمی استفاده شد تا هر نام با شواهد داده‌ای پشتیبانی شود.

یافته‌های پژوهش

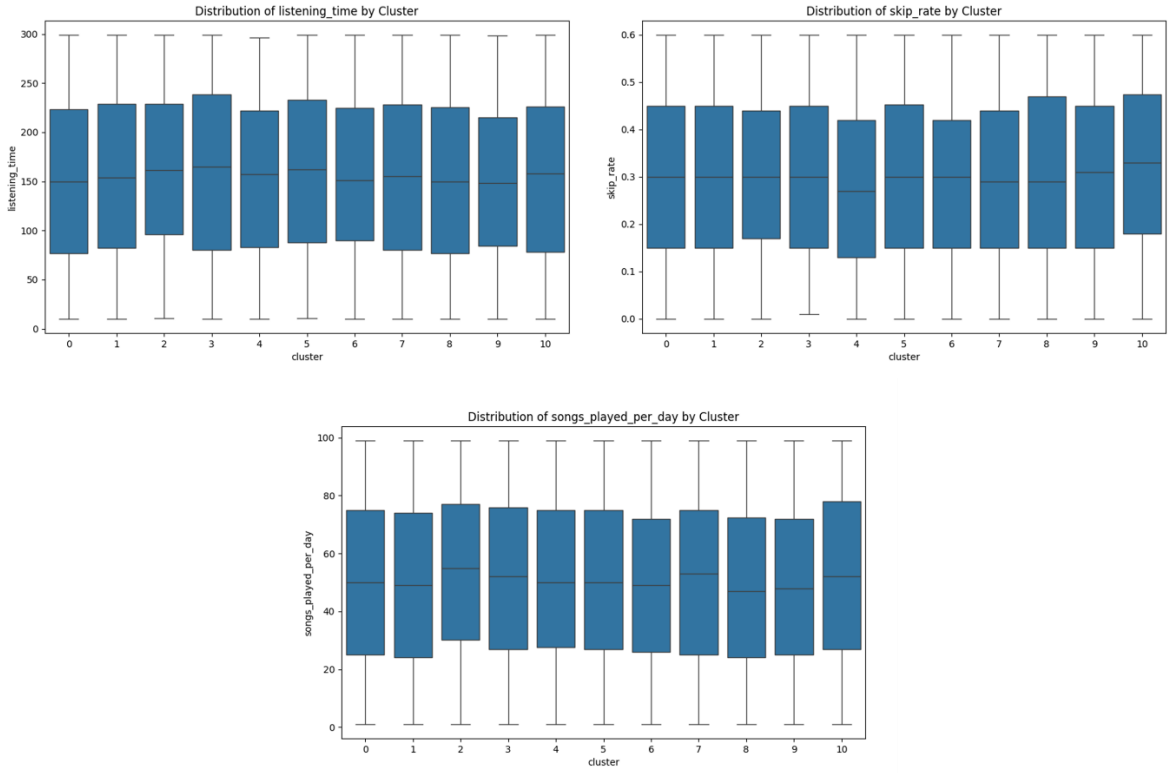
این بخش به تبیین جامع نتایج حاصل از خوشه‌بندی ۸۰۰۰ کاربر فعال اسپاتیفای با استفاده از الگوریتم HDBSCAN می‌پردازد. داده‌ها شامل ده شاخص رفتاری و جمعیت‌شناختی است که عبارت‌اند از: جنسیت، سن، کشور، نوع اشتراک، زمان گوش دادن روزانه، تعداد آهنگ‌های پخش‌شده در روز، نرخ رد کردن آهنگ، نوع دستگاه، تعداد تبلیغات شنیده‌شده در هفته و میزان گوش دادن آفلاین. هدف این تحلیل، تفسیر اندازه و اهمیت نسبی خوشه‌ها، واکاوی دقیق تفاوت‌های درون‌خوشه‌ای و بین‌خوشه‌ای در شاخص‌های یادشده، ارزیابی مرزبندی و هم‌پوشانی خوشه‌ها در فضای ویژگی‌ها و درنهایت، استخراج و تشریح پرسوناهای مشتریان در سکوهای آنلاین پخش موسیقی است.

1. min_cluster_size
2. min_samples
3. Density-based clustering validation



شکل ۱. فراوانی جامعهٔ موردبررسی در هر خوشه

شکل ۱، فراوانی جامعهٔ موردبررسی در هر خوشه را نشان می‌دهد. همان‌گونه که مشاهده می‌شود توزیع کاربران در میان خوشه‌ها یکنواخت نیست و معمولاً یک یا دو خوشه سهم چشمگیرتری از کل جامعه را تشکیل می‌دهند، درحالی‌که خوشه‌های کوچک‌تر نمایندهٔ الگوهای رفتاری خاص‌تری هستند. این ناهم‌ترازی نشان می‌دهد که خوشه‌های پرجمعیت، اهرم اصلی برای بهبود شاخص‌های کلیدی عملکرد محسوب می‌شوند، زیرا تغییرات کوچک در آن‌ها اثر تجمعی بزرگی بر کل جامعه می‌گذارد. در نقطه مقابل، خوشه‌های کم‌جمعیت، با وجود اندازهٔ محدود، به دلیل ترجیحات متمایز، بهترین میدان برای آزمایش محصولات شخصی‌سازی‌شده و طراحی بسته‌های خدمت اختصاصی هستند. در مسیر شناسایی پرسوناهای مشتریان، بخشی از کاربران به‌عنوان «نویز» تشخیص داده‌شده می‌شوند که می‌توان این گروه را کاربرانی دانست که الگوی شنیداری‌شان هنوز پایدار نشده یا در آستانهٔ جابه‌جایی میان پرسوناها قرار دارند که توجه به این دسته از کاربران نیز با توجه به اهمیت نگه‌داشت مشتریان، حائز اهمیت است. ازنظر توالی اقدامات، توصیه می‌شود تحلیل‌های بهینه‌سازی تجربه ابتدا بر خوشه‌های پرجمعیت متمرکز شود و هم‌زمان، برای خوشه‌های کوچک‌تر، طرح‌های آزمایشی (A/B) طراحی گردد تا ارزش‌افزوده اختصاصی هر گروه سنجیده شود.



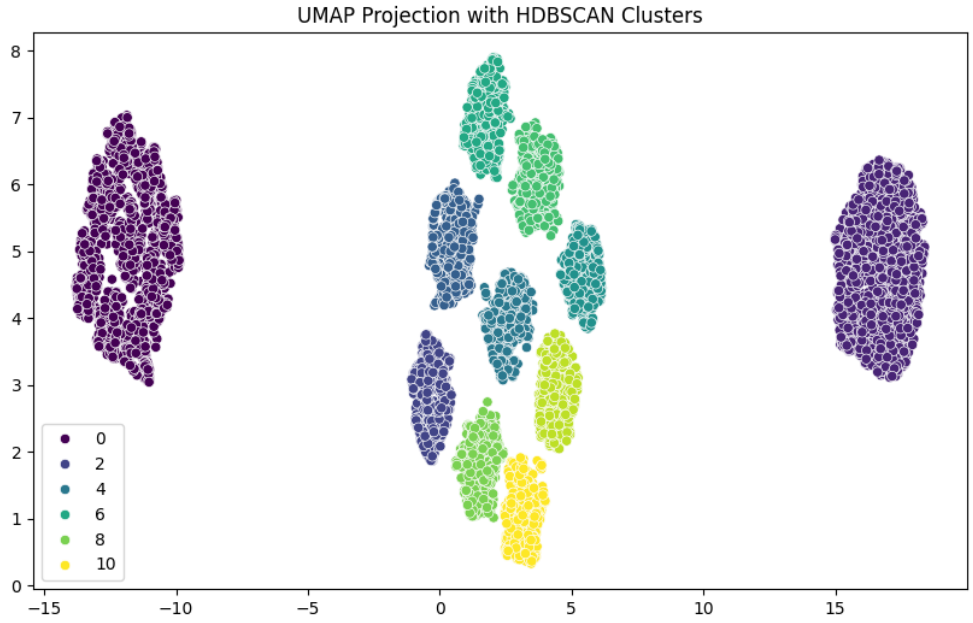
شکل ۲. توزیع خوشه‌ها در ویژگی‌های مختلف مشتریان

برای درک عمیق تفاوت‌های رفتاری و جمعیت‌شناختی میان خوشه‌ها، مقایسه توزیع شاخص‌ها ضروری است. تحلیل شکل ۲ به‌طورمعمول دو رژیم رفتاری متمایز را نمایان می‌کند: در یک‌سو، کاربران جوان‌تر قرار دارند که موسیقی را با رویکرد اکتشافی مصرف می‌کنند (نرخ رد کردن آهنگ بالاتر، تعداد آهنگ‌های پخش‌شده در روز بیشتر، تمایل به تغییر سریع قطعه و تکیه بر تلفن همراه) و در سوی دیگر، کاربران مسن‌تر که شنیداری باثبات‌تری دارند (زمان گوش دادن روزانه بیشتر، نرخ رد کردن آهنگ پایین‌تر، گرایش به تجربه بدون تبلیغ (نوع اشتراک پریمیوم) و استفاده معنادار از حالت آفلاین). نوع اشتراک به‌عنوان تمایزگر اقتصادی-تجربی کاربران، با دو شاخص «تعداد تبلیغات شنیده‌شده در هفته» و «میزان گوش دادن آفلاین» ارتباط معناداری دارد.

در خوشه‌هایی که خرید اشتراک پریمیوم غالب است، مواجهه با تبلیغات، به‌صورت حداقلی است و استفاده از آفلاین به سقف محلی خود می‌رسد. این در حالی است که در خوشه‌های دربرگیرندهٔ کاربران رایگان، الگو معکوس است و پارت‌های شنیداری به‌واسطه تبلیغات کوتاه‌تر و گسسته‌تر می‌شود. نسبت «زمان گوش دادن روزانه» به «تعداد آهنگ‌های پخش‌شده در روز» در حضور «نرخ رد کردن آهنگ»، کیفیت توجه کاربران را رمزگشایی می‌کند: دستهٔ اول، زمان زیاد همراه با تعداد آهنگ بالا و نرخ رد پایین، نشانهٔ تمرکز عمیق و انتخاب آگاهانه است. دستهٔ دوم، زمان متوسط با تعداد آهنگ بالا و نرخ رد زیاد، نشانه اکتشاف سریع کاربران و دستهٔ سوم با زمان پایین با تعداد آهنگ کم و نرخ رد پایین، معمولاً با کاربرانی درگیر است که آهنگ را در پس‌زمینهٔ فعالیت‌های عادی روزانه خود در برخی زمان‌های خاص استفاده می‌کنند.

«نوع دستگاه» نقش مهمی در تحلیل پرسونای مخاطبان دارد که به‌طورکلی سه دستهٔ اصلی را شامل می‌شود: غلبهٔ تلفن همراه معمولاً با پارت‌های کوتاه اما پرتواتر و رفتار سیال‌تر همراه است. غلبهٔ دسکتاپ با پارت‌های کم‌تواتر اما طولانی‌تر و رفتار متمرکزتر و الگوی چنددستگاهی با جابه‌جایی میان موقعیت‌های کاری/خانگی/سفر همبستگی دارد. تفاوت‌های «جنسیت» و «کشور» نیز می‌توانند تعدیل‌گر رفتار باشند. در برخی خوشه‌ها، نسبت‌های جنسی متفاوت با تفاوت‌های ظریفی

در نرخ رد یا تمایل به آفلاین همراه می‌شود و در بُعد جغرافیا، الگوهای زمانی استفاده و شدت مواجهه تبلیغاتی می‌تواند برحسب بازار (کشور) تغییر کند. کنار هم نهادن این الگوها در شکل ۲، عملاً «امضای رفتاری» هر خوشه را می‌سازد و نشان می‌دهد کدام ترکیب از شاخص‌ها، خوشه‌ها را از یکدیگر متمایز ساخته است.



شکل ۳. نمایش خوشه‌های پرسونای مشتریان

برای ارزیابی تمایز، هم‌پوشانی و تراکم درونی خوشه‌ها، شکل ۳، نمایش دوبُعدی (بر پایه نگاشت کاهش‌بُعد) را نشان می‌دهد. این شکل بیانگر آن است که کدام خوشه‌ها هسته‌های متراکم و مرزهای مشخص دارند و بنابراین پایدارترین پرسوناها را نمایندگی می‌کنند و کجا هم‌پوشانی‌های لبه‌ای وجود دارد که بیانگر کاربران با نوع رفتاری غیریکنواخت است؛ یعنی کسانی که بخشی از رفتارشان با یک پرسونا و بخشی دیگر با پرسونای مجاور هم‌راستاست. چنین هم‌پوشانی‌هایی معمولاً مسیرهای جابه‌جایی میان پرسوناها را درگذر زمان نشان می‌دهد و می‌تواند شامل ارتقای نوع اشتراک (از رایگان به پریمیوم)، تغییر دستگاه غالب (از موبایل به دسکتاپ یا برعکس)، یا تغییر ریتم زندگی (افزایش رفت‌وآمد یا سفر) می‌تواند کاربر را از یک خوشه به خوشهٔ دیگر منتقل کند. فاصله‌های هندسی میان هسته‌های خوشه‌ها نیز سرنخ‌هایی از شاخص‌های تفکیک‌گر اصلی می‌دهد؛ در اغلب موارد، سن، نوع اشتراک و نرخ رد کردن آهنگ، محورهای برجسته تمایزند. وجود زیر توده‌ها در پیرامون هستهٔ هر خوشه نشان‌دهندهٔ ریز تیپ‌های درون‌پرسونایی است. برای مثال، در خوشه‌ای با غالبیت استفاده از تلفن همراه، ممکن است زیرگروهی از کاربران بخش معناداری از زمان خود را روی دسکتاپ صرف کنند و الگویی چنددستگاهی از خود نشان دهند. در ادامه و بر پایهٔ مشاهدات سه شکل پیش‌گفته و نیز خلاصهٔ عددی جدول ۱ که بیانگر ۱۱ خوشهٔ شناسایی‌شده با روش HDBSCAN است، پرسوناهای مشتریان استخراج‌شده و به‌صورت روشن و عملیاتی تشریح می‌شوند. به‌منظور پشتیبانی کمی این پرسوناها، جدول ۱، اطلاعات دقیقی از جزئیات مربوط به ۱۱ خوشهٔ اصلیِ مستخرج از میان ۸۰۰۰ دادهٔ بررسی‌شده را ارائه می‌دهد. در این جدول، متغیرهایی مانند سن، نسبت نوع اشتراک، میانگین زمان گوش دادن روزانه، میانگین تعداد آهنگ‌های روزانه، میانگین نرخ رد کردن آهنگ، میانگین تعداد تبلیغات شنیده‌شده در هفته، میانگین میزان گوش دادن آفلاین، نوع دستگاه غالب و توزیع کشورهای کاربران به‌طور کامل گزارش‌شده است. گفتنی است که خوشه‌های خام استخراج‌شده (۱۱ خوشه) در مرحلهٔ تفسیر و تجمیع مفهومی، بر اساس شباهت‌های رفتاری و کاربردی، در قالب شش پرسونای عملیاتی ادغام و بازتعریف شدند.

جدول ۱. خلاصهٔ نتایج تحلیل خوشه‌های پرسونای مشتریان

device type	subscription type	country	gender	offline listening	ads listened per week		skip rate		Songs played per day		listening time		age		
<lambda>	<lambda>	<lambda>	<lambda>	mean	std	mean	std	mean	std	mean	std	mean	std	mean	cluster
Desktop	Family	AU	Male	1	0	0	0.17	0.3	28.31	50.38	83.86	151.04	12.55	38.08	0
Web	Free	PK	Female	0	12.98	27.53	0.18	0.3	28.71	49.2	84.47	155	12.74	37.29	1
Desktop	Student	IN	Other	1	0	0	0.17	0.3	28.46	52.61	81.83	160.72	12.65	37.91	2
Desktop	Premium	US	Male	1	0	0	0.17	0.3	28.79	50.91	86.91	161.17	13.02	37.58	3
Mobile	Premium	DE	Male	1	0	0	0.17	0.28	28.28	50.85	81.41	153.02	12.89	37.84	4
Web	Premium	AU	Female	1	0	0	0.18	0.3	27.7	50	82.39	158.82	12.73	36.49	5
Desktop	Premium	AU	Female	1	0	0	0.17	0.29	28.05	49.27	83.32	154.26	13.02	38.83	6
Mobile	Student	AU	Female	1	0	0	0.17	0.3	28.45	51.18	84.94	152.72	12.73	38.65	7
Mobile	Premium	IN	Other	1	0	0	0.18	0.3	28.2	48.3	86.12	150.26	12.99	37.7	8
Web	Premium	US	Male	1	0	0	0.17	0.3	28.29	49.54	81.12	150.78	12.68	36.56	9
Web	Premium	IN	Other	1	0	0	0.17	0.32	29.17	51.56	85.89	152.78	12.58	37.14	10

به‌طورکلی، از تجمیع خوشه‌های شناسایی‌شده و تحلیل رفتار کاربران هر خوشه، می‌توان شش نوع پرسونای مخاطب در سکوهای آنلاین پخش موسیقی را به شرح زیر شناسایی کرد:

۱. مشتریان پُرمصرفِ پریمیوم: شنوندگانی با عمق توجه بالا؛ زمان گوش دادن روزانه زیاد، تعداد آهنگ‌های پخش‌شده بالا و استاندارد انتخاب سخت‌گیرانه‌تر که می‌تواند به نرخ رد کردن آهنگ متوسط تا کمی بالا بینجامد، اما این دسته از مشتریان، با بازگشت‌های مکرر به پلی‌لیست‌های منتخب همراه هستند. تجربهٔ بدون تبلیغ برای این گروه حیاتی است (نوع اشتراک پریمیوم)، میزان گوش دادن آفلاین بالاست و در موقعیت‌های متنوع (رفت‌وآمد، سفر و ورزش) شنیدن بی‌وقفه را ممکن می‌کند. غالباً مبتنی بر تلفن همراه است، اما نشانه‌هایی از الگوی چنددستگاهی در آن‌ها دیده می‌شود (مثلاً دسکتاپ در ساعات کاری). ازنظر محتوایی، نسبت به پیشنهادهای شخصی‌سازی‌شده دقیق واکنش مثبت نشان می‌دهند و در صورت تکرار بیش‌ازحد الگوها ممکن است دچار اشباع شوند. توصیه اجرایی در مورد این دسته از مشتریان، تقویت تنوع پیشنهادها با کنترل تازگی، دسترسی آفلاین پایدار و ارائه گزینه‌های افزایش کیفیت آهنگ‌ها برای آن‌ها است. نرخ رضایت از پیشنهادها، نرخ مصرف آفلاین و سهم پارت‌های بدون وقفه تبلیغ را می‌توان به‌عنوان شاخص‌های کلیدی نظارت و ارزیابی این دسته از کاربران، پیشنهاد داد.

۲. مشتریان رایگان پس‌زمینه‌ای: موسیقی برای این گروه «پس‌زمینه» فعالیت‌های دیگر است. زمان گوش دادن روزانه و تعداد آهنگ‌های پخش‌شده پایین تا متوسط، نرخ رد کردن آهنگ پایین (به سبب گوش‌سپاری نیمه‌فعال)، مواجهه تبلیغاتی زیاد (نوع اشتراک رایگان) و میزان گوش دادن آفلاین پایین. رفتار شنیداری آن‌ها به بافت روزمره وابسته است (کار، مطالعه، امور خانه). پیشنهاد می‌شود، پلی‌لیست‌هایی با محوریت تمرکز و مطالعه و آرامش به این دسته از مشتریان پیشنهاد شود و میزان تبلیغات ارائه‌شده به آن‌ها کاهش پیدا کند تا بتوان از آن‌ها مشتریان وفادار غیر رایگان ایجاد کرد.

۳. شکارچیان ترندِ موبایلی: شاخصه‌های اصلی این دسته از مشتریان، جوان، پرتحرک و حساس به ترند؛ تعداد آهنگ‌های پخش شده زیاد، نرخ رد کردن آهنگ بالا (اکتشاف سریع)، زمان گوش دادن روزانه در حد متوسط است. تلفن همراه دستگاه غالب است و الگوی استفاده کوتاه و پرتواتر در میان آن‌ها دیده می‌شود. مواجهه با تبلیغات در این دسته، در صورت رایگان بودن قابل‌توجه است، اما به‌واسطه حساسیت به وقفه‌ها، کشش ارتقا به پریمیوم وجود دارد. ازنظر محتوایی، به پلی‌لیست‌های به‌روز و گونه‌های نوظهور واکنش مثبت نشان می‌دهند و چرخه‌های کشف–ذخیره‌سازی کوتاه دارند. پیشنهاد می‌شود برای این دسته از مشتریان، پلی‌لیست‌هایی با محتوای آهنگ‌های ترند شبکه‌های اجتماعی ارائه شود و یا ریمیکس‌های با کیفیت از ترکیبی از آن‌های ترند موجود در شبکه‌های اجتماعی همچون اینستاگرام ارائه شود تا به

مشتریان پولی تبدیل شوند.

۴. سنت‌گراهای باکیفیت و آفلاین: شاخصه‌های اصلی این دسته از مشتریان، سن بالا، وفاداری بالا به فهرست‌های تثبیت‌شده و ژانرهای آشنا، نرخ رد کردن آهنگ پایین، زمان گوش دادن روزانه بالا، میزان گوش دادن آفلاین زیاد و مواجهه ناچیز با تبلیغات به دلیل گرایش به پریمیوم است. تجربه باثبات و بی‌وقفه برای آن‌ها خط قرمز است؛ تغییرات ناگهانی در رابط یا منطق پیشنهادها می‌تواند برای آن‌ها ایجاد نارضایتی کند. ازنظر محتوایی، پلی‌لیست‌های هنرمندمحور و نسخه‌های بی‌کلام با کیفیت صوتی بالا مطلوب است. حفظ ثبات تجربه، گزینه‌های با کیفیت بالا و بسته‌های خانوادگی، از پیشنهادهای قابل‌ارائه به این دسته از مشتریان است.

۵. جوانانِ سریع‌گذرِ تبلیغ‌پذیر: این دسته از مشتریان، رایگان‌محور با الگوی «قطعه‌قطعه» هستند. به این معنا که تعداد آهنگ‌های روزانه زیاد اما زمان گوش دادن روزانه نسبتاً پایین، نرخ رد کردن آهنگ بالا، مواجهه تبلیغاتی زیاد و میزان گوش دادن آفلاین پایینی دارند. این دسته از مشتریان معمولاً در کمربند هم‌پوشان میان «مشتریان رایگان پس‌زمینه‌ای» و «شکارچیان ترند» جای می‌گیرند و ازنظر راهبردی «نقطهٔ کانونی تبدیل» هستند. ازنظر محتوایی، ارائه مینی‌پلی‌لیست‌های کوتاه، هایلایت‌های ژانری و پیشنهادهای آهنگ به‌صورت روزانه، برای این دسته از مشتریان کارآمدتر است.

۶. حرفه‌ای‌هایِ چنددستگاهیِ جهانی: ویژگی اصلی این دسته از کاربران، مصرف متعادل اما عمیق با جابه‌جایی معنادار میان تلفن همراه و دسکتاپ است. در این دسته از کاربران، زمان گوش دادن روزانه نسبتاً بالا، نرخ رد کردن آهنگ متوسط، میزان گوش دادن آفلاین متوسط تا زیاد و تنوع جغرافیایی بالا است. الگوی شنیداری‌شان به نقش/زمینه بستگی دارد (کار/خانه/سفر) و به همگام‌سازی بی‌نقص میان دستگاه‌ها حساس‌اند. ازنظر محتوایی، پلی‌لیست‌های بافت‌محور (محیط کار، رانندگی و عصرانه) و پیشنهادهای مربوط به رویدادهای مربوط به تاریخ روز، کارآمد است. تقویت یکپارچگی میان دستگاه‌ها، ازسرگیری یک‌کلیک در دستگاه دوم و پیشنهادهای زمینه‌محور، از راهکارهایی است که می‌تواند کاربران رایگان این دسته را به کاربران پولی تبدیل کند.

بحث، پیشنهاد و نتیجه‌گیری

یافته‌های این پژوهش نشان داد که الگوهای مصرف موسیقی در سکوهای پخش آنلاین را نمی‌توان صرفاً بر اساس متغیرهای جمعیت‌شناختی یا نوع اشتراک کاربران تبیین کرد، بلکه رفتار شنیداری کاربران حاصل برهم‌کنش هم‌زمان چند بُعد کلیدی شامل عمق توجه، سطح خدمت و بافت مصرف است. خوشه‌بندی داده‌های رفتاری ۸۰۰۰ کاربر فعال اسپاتیفای با استفاده از الگوریتم HDBSCAN منجر به شناسایی شش پرسونای متمایز شد که هر یک بیانگر الگوی خاصی از مصرف موسیقی، نحوهٔ تعامل با سکوی پخش و انتظارات متفاوت از تجربهٔ کاربری هستند. این یافته‌ها نشان می‌دهد که حتی در یک سکوی واحد، کاربران در طیفی از مصرف عمیق و متمرکز تا مصرف سریع، اکتشافی و قطعه‌قطعه قرار می‌گیرند و همین ناهمگنی رفتاری، ضرورت رویکردهای شخصی‌سازی‌شده در طراحی خدمات و استراتژی‌های بازاریابی را برجسته می‌سازد.

تحلیل دقیق ویژگی‌های هر پرسونا نشان داد که تمایز میان خوشه‌ها عمدتاً حول سه بُردار زیربنایی شکل می‌گیرد؛ عمق توجه که ترکیبی از زمان گوش دادن روزانه، تعداد آهنگ‌های پخش‌شده و نرخ رد کردن آهنگ است، سطح درگیری شناختی کاربران با محتوا را بازنمایی می‌کند. سطح خدمت که از طریق نوع اشتراک، میزان مواجهه با تبلیغات و استفاده از قابلیت آفلاین نمود می‌یابد، به‌طور مستقیم با کیفیت تجربه شنیداری و تحمل کاربران نسبت به وقفه‌ها مرتبط است. درنهایت، بافت مصرف شامل نوع دستگاه و پراکندگی جغرافیایی کاربران، زمینه‌های موقعیتی استفاده از موسیقی را مشخص می‌کند. ترکیب این سه بُعد، چارچوبی تبیینی فراهم می‌آورد که فراتر از خوشه‌بندی صرف آماری، امکان تفسیر رفتاری پرسوناها را مهیا می‌سازد.

مقایسه نتایج این پژوهش با مطالعات پیشین، به‌ویژه پژوهش فولر و همکاران (۲۰۱۶) و ایم و کیم (۲۰۱۶)، نشان‌دهندهٔ همخوانی مفهومی نسبی میان برخی پرسونای شناسایی‌شده است. به‌طور مشخص، پرسونای مشتریان پُرمصرف پریمیوم با «شنونده با ذائقه خاص»، شکارچیان ترند موبایلی با «کاوشگر موسیقی آزاد» و حرفه‌ای‌های چنددستگاهی جهانی با «متصدی

فعال» در مطالعات قبلی تطابق مفهومی دارند. بااین‌حال، این پژوهش توانست با اتکا به داده‌های رفتاری واقعی و در مقیاس بزرگ، الگوهایی را شناسایی کند که در پژوهش‌های کیفی پیشین کمتر برجسته شده بودند. ظهور پرسونای «سنت‌گراهای باکیفیت و آفلاین» بیانگر افزایش اهمیت تجربه پایدار، بدون وقفه و با کیفیت صوتی بالا در میان بخشی از کاربران است؛ الگویی که احتمالاً بازتاب تغییرات سبک زندگی و افزایش مصرف موسیقی در موقعیت‌های غیرآنلاین یا نیمه‌آنلاین است. همچنین، عدم مشاهدهٔ پرسونای «بی‌اعتماد» و «کاربر منزوی» که در برخی مطالعات پیشین گزارش شده‌اند، می‌تواند ناشی از بلوغ الگوریتم‌های پیشنهاددهی و افزایش اعتماد کاربران به سکوهای جریان اصلی موسیقی باشد.

از منظر کاربردی، نتایج این پژوهش پیامدهای مهمی برای طراحی تجربهٔ کاربری، توسعهٔ خدمات و تدوین استراتژی‌های بازاریابی دیجیتال به همراه دارد. شناسایی خوشه‌های پرجمعیت این امکان را فراهم می‌آورد که بهبودهای کوچک در پیشنهاد محتوا، طراحی رابط کاربری یا سیاست‌های اشتراک، اثر تجمعی قابل‌توجهی بر کل جامعه کاربران داشته باشد. در مقابل، خوشه‌های کوچک‌تر اما متمایز، بستر مناسبی برای آزمون خدمات نوآورانه، بسته‌های اشتراکی خاص یا ویژگی‌های شخصی‌سازی‌شده فراهم می‌کنند. همچنین، تمرکز بر پرسونای «جوانان سریع‌گذرِ تبلیغ‌پذیر» به‌عنوان یک ناحیه هم‌پوشان میان مصرف رایگان و اکتشافی، می تواند راهبردی مؤثر برای افزایش نرخ تبدیل کاربران رایگان به کاربران پرداختی باشد. به‌طورکلی، چارچوب پرسونای استخراج‌شده در این پژوهش می‌تواند به‌عنوان یک ابزار تصمیم‌یار برای مدیران سکوهای دیجیتال در حوزه‌های فراتر از موسیقی نیز مورداستفاده قرار گیرد.

باوجود یافته‌های معنادار، این پژوهش با محدودیت‌هایی نیز همراه است. نخست آن‌که داده‌های مورداستفاده به یک سکوی خاص (اسپاتیفای) محدود بوده و تعمیم نتایج به سایر سکوهای پخش موسیقی یا سکوهای دیجیتال نیازمند احتیاط است. دوم، داده‌ها ماهیت مقطعی دارند و پویایی زمانی رفتار کاربران و جابه‌جایی آن‌ها میان پرسوناها در طول زمان بررسی نشده است. سوم، متغیرهای روان‌شناختی، انگیزشی و زمینه‌ای که می‌توانند به تفسیر عمیق‌تر رفتار کاربران کمک کنند، به دلیل محدودیت داده‌های در دسترس در مدل لحاظ نشده‌اند. افزون بر این، هرچند الگوریتم HDBSCAN قابلیت بالایی در شناسایی خوشه‌های با تراکم متفاوت دارد، اما نتایج آن همچنان به انتخاب شاخص‌ها و پیش‌پردازش داده‌ها حساس است.

بر این اساس، پیشنهاد می‌شود پژوهش‌های آتی با استفاده از داده‌های طولی، تغییرات پرسونای کاربران و مسیرهای انتقال میان خوشه‌ها را درگذر زمان بررسی کنند. همچنین، ترکیب داده‌های رفتاری با داده‌های کیفی یا روان‌شناختی می‌تواند به غنای تفسیر پرسونای استخراج‌شده کمک کند. به‌کارگیری این رویکرد در سکوهای بومی و مقایسه نتایج با نمونه‌های بین‌المللی نیز می‌تواند به توسعه مدل‌های بومی‌سازی‌شده پرسونای مشتریان منجر شود. درنهایت، استفاده از چارچوب‌های یادگیری ماشین پیش‌بینانه برای اتصال مستقیم پرسونای کاربران به شاخص‌های کلیدی عملکرد، می‌تواند مسیر پژوهش‌های آینده را به‌سوی کاربردهای عملی‌تر و تصمیم‌محورتر هدایت کند.

منابع

آقازاده، هاشم، جعفری، عماد، اسمعیلی، حسین. (۱۳۹۹). بررسی رابطه همخوانی شخصیت مصرف‌کننده با برند، هویت پذیری مصرف‌کننده از برند و نیات رفتاری پس از خرید. *نشریه علمی راهبردهای بازرگانی*. دورهٔ ۱۲، شمارهٔ ۶، ۱۷-۴۰. https://cs.shahed.ac.ir/article_2293.html

یعقوبی چوبری، علی. (۱۳۹۶). زبان استعاری و بینش جامعه‌شناختی اِروینگ گافمن. مطالعات جامعه‌شناختی (نامه علوم اجتماعی سابق), ۲۴(۱), ۱۴۹-۱۷۰ https://doi.org/10.22059/jsr.2017.63433

باقری، جواد، و یزدانی، ناصر. (۱۴۰۲). تأثیر خلق و به‌کارگیری هوش بازار صادراتی بر عملکرد از طریق استراتژی بخش‌بندی و اولویت‌بندی مشتریان و مدل‌های فروش. نشریه علمی راهبردهای بازرگانی، ۲۰(۲۲), ۶۵-۸۲. https://doi.org/10.22070/cs.2024.19239.1386

Adlin, T., Jamesen, H., & Krebs, T. (2001). Fake people and sticky notes: Fostering communication for human-centered software design. Whitepaper, published under: http://www. jamesen. com/publications/FakePeople-G. pdf.

Aghazade, H., Jafari, E., & Esmaeily, H. (2020). Investigate the Relationship between Consumer Personality Congruence with the Brand, Brand Identity and Post-Purchase Behavioral Intentions. Commercial Strategies, 12(6), 17-40. https://cs.shahed.ac.ir/article_2293.html?lang=en **[in Persian]**

Akner-Koler, C., & Ranjbar, P. (2016). Integrating sensitizing labs in an educational design process for haptic interaction. FormAkademisk, 9(2). https://doi.org/10.7577/formakademisk.1269

Alves Gomes, M., & Meisen, T. (2023). A review on customer segmentation methods for personalized customer targeting in e-commerce use cases. Information Systems and e-Business Management, 21(3), 527-570. https://doi.org/10.1007/s12652-020-02015-w

Antle, Alissa Nicole. "Child-personas: fact or fiction?." Proceedings of the 6th conference on Designing Interactive systems. 2006. https://doi.org/10.1145/1142405.1142411

Atzeni, A., Cameroni, C., Faily, S., Lyle, J., & Fléchais, I. (2011, August). Here's Johnny: a methodology for developing attacker personas. In 2011 Sixth International Conference on Availability, Reliability and Security (pp. 722-727). IEEE . https://doi.org/10.1109/ARES.2011.115

Bagheri, J., & Yazdani, N. (2024). The Impact of the Creation and Application of Market Intelligence on Performance Through Strategy of Segmentation, Prioritization, and Sales Models. Commercial Strategies, 20(22), 65-82. https://doi.org/10.22070/cs.2024.19239.1386 **[in Persian]**

Barata, M. L., & Coelho, P. S. (2021). Music streaming services: understanding the drivers of customer purchase and intention to recommend. Heliyon, 7(8), e07783. [PMC8379453] https://doi.org/ 10.1016/j.heliyon.2021.e07783

Blomquist, Å., & Arvola, M. (2002, October). Personas in action: ethnography in an interaction design team. In Proceedings of the second Nordic conference on Human-computer interaction (pp. 197-200). https://doi.org/10.1145/572020.572044

Bonnardel, N., & Pichot, N. (2020). Enhancing collaborative creativity with virtual dynamic personas. Applied ergonomics, 82, 102949. https://doi.org/10.1016/j.apergo.2019.102949

Brooks, C., & Greer, J. (2014, March). Explaining predictive models to learning specialists using personas. In Proceedings of the fourth international conference on learning analytics and knowledge (pp. 26-30). https://doi.org/10.1145/2567574.2567612

Campello, R. J., Moulavi, D., & Sander, J. (2013, April). Density-based clustering based on hierarchical density estimates. In *Pacific-Asia conference on knowledge discovery and data mining* (pp. 160-172). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-37456-2_14

Canossa, A., & Drachen, A. (2009, August). Play-personas: behaviours and belief systems in user-centred game design. In IFIP Conference on Human-Computer Interaction (pp. 510-523). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-03658-3_55

Chen, R. C., Nivala, W. C. Y., & Chen, C. B. (2011, July). Modeling the role of empathic design engaged personas: an emotional design approach. In International Conference on Universal Access in Human-Computer Interaction (pp. 22-31). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-21663-3_3

Dharwada, P., Greenstein, J. S., Gramopadhye, A. K., & Davis, S. J. (2007, October). A case study on use of personas in design and development of an audit management system. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 51, No. 5, pp. 469-473). Sage CA: Los Angeles, CA: SAGE Publications. https://doi.org/10.1177/154193120705100509

Dos Santos, T. F., de Castro, D. G., Masiero, A. A., & Aquino Junior, P. T. (2014, June). Behavioral persona for human-robot interaction: a study based on pet robot. In International Conference on Human-Computer Interaction (pp. 687-696). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-07230-2_65

Ferreira, B. M., Barbosa, S. D., & Conte, T. (2016, June). PATHY: using empathy with personas to design applications that meet the users' needs. In International Conference on Human-Computer Interaction (pp. 153-165). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-39510-4_15

Frauenberger, C., Good, J., Keay-Bright, W., & Pain, H. (2012, May). Interpreting input from children: a designerly approach. In Proceedings of the SIGCHI conference on human factors in computing systems (pp. 2377-2386). https://doi.org/10.1145/2207676.2208399

Fuller, J., Hubener, L., Kim, Y. S., & Lee, J. H. (2016, August). Elucidating User Behavior in Music Services Through Persona and Gender. In ISMIR (pp. 626-632). https://doi.org/10.5281/zenodo.1415928

Hendriks, N., Truyen, F., & Duval, E. (2013, September). Designing with dementia: Guidelines for participatory design together with persons with dementia. In IFIP Conference on Human-Computer Interaction (pp. 649-666). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-40483-2_46

Hensely-Schinkinger, Susanne,. Pinatti de Carvalho, Aparecido Fabiano,. Glanznig, Michael,. and Tellioğlu, Hilda,. 2015. The Definition and Use of Personas in the Design of Technologies for Informal Caregivers. In Human-Computer Interaction: Design and Evaluation, Masaaki Kurosu (ed.). Springer International Publishing,Cham, 202-213. https://doi.org/978-3-319-20901-2/10.1007_19

Högberg, D., Hanson, L., Lundström, D., Jönsson, M., & Lämkull, D. (2008, August). Representing the elderly in digital human modeling. In Proceedings of the 40th Annual Nordic Ergonomic Society Conference, Reykjavik, Iceland.
https://www.researchgate.net/publication/247778784_Representing_the_elderly_in_digital_human_modelling

Hsu, P. F., Lu, Y. H., Chen, S. C., & Kuo, P. P. Y. (2024). Creating and validating predictive personas for target marketing. International Journal of Human-Computer Studies, 181, 103147. https://doi.org/10.1016/j.ijhcs.2023.103147

Im, H., & Kim, N. W. (2016, May). Three Personas of Potential High-Resolution Music Users. In Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (pp. 2851-2856). https://doi.org/10.1007/978-3-031-63223-5

Jansen, B. J., Salminen, J. O., & Jung, S. G. (2020). Data-driven personas for enhanced user understanding: Combining empathy with rationality for better insights to analytics. Data and Information Management, 4(1), 1-17. https://doi.org/10.2478/dim-2020-0005

Jensen, M. B., Elverum, C. W., & Steinert, M. (2017). Eliciting unknown unknowns with prototypes: Introducing prototrials and prototrial-driven cultures. Design Studies, 49, 1-31. https://doi.org/10.1016/j.destud.2016.12.002

Kauppinen, S., Luojus, S., & Lahti, J. (2016, October). Involving citizens in open innovation process by means of gamification: The case of WeLive. In Proceedings of the 9th Nordic Conference on Human-Computer Interaction (pp. 1-4). https://doi.org/10.1145/2971485.2971526

Kim, E., Yoon, J., Kwon, J., Liaw, T., & Agogino, A. M. (2019, July). From innocent irene to parental patrick: Framing user characteristics and personas to design for cybersecurity. In Proceedings of the Design Society: International Conference on Engineering Design (Vol. 1, No. 1, pp. 1773-1782). Cambridge University Press. https://doi.org/10.1017/dsi.2019.183

Klepek, M. (2019). Buyer personas: its use and limitations in online marketing. Marketing Identity, 7(1), 886-896. https://www.ceeol.com/search/article-detail?id=871161

Laine, T. H., Normark, J., Lindvall, H., Lindqvist, A. K., & Rutberg, S. (2020). A distributed multiplayer game to promote active transport at workplaces: user-centered design, implementation, and lessons learned. IEEE Transactions on Games, 12(4), 386-397. https://doi.org/10.1109/TG.2020.3021728

Lee, B., Srivastava, S., Kumar, R., Brafman, R., & Klemmer, S. R. (2010, April). Designing with interactive example galleries. In Proceedings of the SIGCHI conference on human factors in computing systems (pp. 2257-2266). https://doi.org/10.1145/1753326.1753667

Lee, J. H., Price, R., et al. (2015). Understanding users of music streaming services through persona identification. https://visualizationdesign.wordpress.com/2015/06/05/understanding-users-of-music-streaming-services-through-persona-identification/

Maier, R., & Thalmann, S. (2010). Using personas for designing knowledge and learning services: results of an ethnographically informed study. International Journal of Technology Enhanced Learning, 2(1-2), 58-74. https://doi.org/10.1504/IJTEL.2010.031260

Malik, A. (2019). Persona based marketing strategies: creation of personas through data analytics (Master's thesis, A. Malik). https://urn.fi/URN:NBN:fi:oulu-201906052360

Park, D., & Kang, J. (2022). Constructing data-driven personas through an analysis of mobile application store data. Applied Sciences, 12(6), 2869. https://doi.org/10.3390/app12062869

Pruitt, J., & Grudin, J. (2003, June). Personas: practice and theory. In Proceedings of the 2003 conference on Designing for user experiences (pp. 1-15). https://doi.org/10.1145/997078.997089

Qyll, N. (2020). Persona as key component in (cultural) person branding. Persona studies, 6(1), 56-71. https://doi.org/10.21153/psj2020vol6no1art941

Rea, D. J., & Young, J. E. (2019, March). Backseat teleoperator: affective feedback with on-screen agents to influence teleoperation. In 14 2019th ACM/IEEE International Conference on Human-Robot Interaction (HRI) (pp. 19-28). IEEE. https://dl.acm.org/doi/10.1145/3488242

Salminen, J., Guan, K., Jung, S. G., & Jansen, B. J. (2021). A survey of 15 years of data-driven persona development. International Journal of Human–Computer Interaction, 37(18), 1685-1708. https://doi.org/10.1080/10447318.2021.1908670

Salminen, J., Wenyun Guan, K., Jung, S. G., & Jansen, B. (2022, April). Use cases for design personas: A systematic review and new frontiers. In Proceedings of the 2022 CHI Conference on human factors in computing systems (pp. 1-21). https://doi.org/10.1145/3491102.3517589

Stoll, J., McColgin, D., Gregory, M., Crow, V., & Edwards, W. K. (2008, May). Adapting personas for use in security visualization design. In VizSEC 2007: Proceedings of the Workshop on Visualization for Computer Security (pp. 39-52). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-78243-8_3

Threatt, A. L., Merino, J., Brooks, J. O., Healy, S., Truesdail, C., Manganelli, J., ... & Green, K. E. (2017). The design, prototyping, and formative evaluation of an assistive robotic table (art) for stroke patients. HERD: Health Environments Research & Design Journal, 10(3), 152-169. https://doi.org/10.1177/1937586716687802

Tu, N., He, Q., Zhang, T., Zhang, H., Li, Y., Xu, H., & Xiang, Y. (2010, November). Combine qualitative and quantitative methods to create persona. In 2010 3rd International Conference on Information Management, Innovation Management and Industrial Engineering (Vol. 3, pp. 597-603). IEEE. https://doi.org/10.1109/ICIII.2010.463

Wagner, C., Hess, T., & Wlömert, N. (2014). Freemium to premium conversion: The role of perceived feature-fit. https://doi.org/10.1007/s12525-014-0168-4

Wanga, R. (2024). Persona Marketing on Social Media: The Evolution from Constructed Authenticity to Consumer Trust. Journal of Advanced Academic Research and Studies 高等学术研究期刊, 1(12), 1-16. http://www.nlbaeai.org/static/upload/file/20250804/1754277658828753.pdf#page=6

Yaghoobi, A. (2017). Metaphorical Language and Erving Goffman's Sociological Imagination. Sociological Review, 24(1), 149-170. https://doi.org/10.22059/jsr.2017.63433 **[in Persian]**

Zhang, G. (2019, July). Creating Enhanced User Experience Through Persona and Interactive Design: A Case of Designing a Motion Sensing Game. In International Conference on Human-Computer Interaction (pp. 382-394). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-23570-3_28

Zhao, Z. (2024). From Free to Premium: A Persona-Based Approach to Understanding Spotify Subscription Upgrades. MA thesis at the University of Chicago. https://doi.org/10.6082/uchicago.11903