

## کاربرد تحلیل خوشه‌ای به منظور داده‌کاوی در پژوهش‌های روان‌شناسی

### Using cluster analysis for data mining in educational research

زهرا نقش<sup>۱</sup>، اعظم مقدم<sup>۲</sup>

#### Abstract

Data mining is the process of sorting and classification of large Data. This paper introduces cluster analysis as a data mining technique and provide examples of its application in clustering of procrastinators. So introduce the cluster analysis and a variety of approaches, how to choose each step of its implementation in a representative sample presented the sample consists of 200 students of Tehran university (100 boys and 100 girls) that get high score in procrastination on the scale of Solomon and Rothblum (1984) and this sample responded to scale of self regulation by Green and Miller (2004), self-efficacy of Midelton and Midgley (1997), the scale of the failure beliefs Harrington (2005) and irrational beliefs Koopmans et al. (1994). The results indicated that different clusters of procrastinators which with identify groups of procrastinators and the characteristics of each of the groups can be more effective ways to reduce procrastination each group presented.

**Keywords:** Data mining, Cluster analysis, Procrastination, self-regulation, self-efficacy, irrational beliefs, frustration belief.

#### چکیده

زمینه: داده‌کاوی فرایند مرتب‌سازی و طبقه‌بندی داده‌های حجیم است. هدف این مقاله معرفی تحلیل خوشه‌ای به عنوان یکی از تکنیک‌های داده‌کاوی و ارائه نمونه‌ای از کاربرد آن در خوشه‌بندی افراد تعلل‌ورز است. روش: بدین منظور ضمن معرفی تحلیل خوشه‌ای و انواع رویکردهای آن، نحوه انتخاب هر یک و گام‌های اجرای آن در یک نمونه معرف ارائه می‌شود این نمونه متشکل از ۲۰۰ دانشجوی دانشگاه تهران (۱۰۰ پسر و ۱۰۰ دختر) که در تعلل‌ورزی بر اساس مقیاس سولومون و روئیلوم (۱۹۸۴) نمرات بالایی به دست آورده بودند انتخاب و به مقیاس‌های پرسشنامه خودتنظیمی گرین و میلر (۲۰۰۴)، خودکارآمدی میدلتن و میگلی (۱۹۹۷)، مقیاس باورهای ناکامی هرینگتون (۲۰۰۵) و باورهای غیرمنطقی کوپمنز و همکاران (۱۹۹۴) پاسخ دادند. یافته‌ها: نتایج این پژوهش بیانگر وجود خوشه‌های مختلف از تعلل‌ورزان است که با شناسایی گروه‌های تعلل‌ورزان و ویژگی‌های هر یک از گروه‌ها می‌توان راهکارهای موثرتری برای کاهش تعلل‌ورزی هر گروه ارائه کرد.

**کلید واژه‌ها:** داده‌کاوی، تحلیل خوشه‌ای، تعلل‌ورزی، خودتنظیمی، خودکارآمدی، باورهای ناکامی و باورهای غیرمنطقی

## مقدمه

داده‌کاوی<sup>۱</sup>، فرایند کشف دانش جدید، معتبر و قابل پیگیری با استفاده از ابزارهای آماری در حجم بالایی از داده‌ها است. اصطلاح کشف دانش برای نخستین بار در دهه ۱۹۹۰ مطرح شد و توجه پژوهشگران را به سمت الگوریتم‌های داده‌کاوی معطوف کرد. به عبارت دیگر داده‌کاوی یعنی فرایند کشف و یا استنتاج الگوهای بالقوه سودمند، اطلاعات معتبر و بدیع، دانش پنهان و قابل فهم موجود در داده‌ها و یا بانک داده‌ها است (بانج<sup>۲</sup> و جادسون<sup>۳</sup>، ۲۰۰۵). دو هدف اصلی داده‌کاوی، پیشگویی و توصیف است. داده‌کاوی پیش‌گویی کننده، الگویی از سیستم ارایه می‌دهد که این الگو را مجموعه‌ای از داده‌های مشخص، پیش‌بینی می‌کند. هدف کلی آن ایجاد الگویی برای طبقه‌بندی، پیش‌بینی و تخمین داده‌ها است. داده‌کاوی توصیفی، اطلاعات جدید و غیر بدیهی را بر اساس مجموعه‌ای از داده‌های موجود ارایه می‌دهد و هدف کلی آن درک و شناخت سیستم‌های تجزیه و تحلیل شده با استفاده از الگوها و روابط موجود است (کیس<sup>۴</sup>، ۲۰۰۳).

تکنیک‌های داده‌کاوی بر اساس کارکردهایی که دارند به چند دسته تقسیم می‌شوند (تریولا<sup>۵</sup>، ۱۹۹۷): ۱. تشریح و تصویرسازی؛ این تکنیک کمک فراوانی در درک مجموعه‌ای از داده‌ها بخصوص داده‌های انبوه و تشخیص الگوهای مخفی در داده‌ها می‌کند به ویژه در داده‌های پیچیده که دارای تعاملات پیچیده و غیرخطی هستند. این تکنیک‌ها قبل از اینکه مدل سازی را شروع کنیم انجام شده و در حقیقت در مرحله شناخت داده قرار می‌گیرند. ۲. طبقه‌بندی و تخمین (پیش‌بینی)؛ تکنیک طبقه‌بندی مربوط به متغیرهایی است که ماهیتا قابل طبقه‌بندی هستند. ۳. جمع و خوشه‌بندی که در ادامه به تشریح بیشتر آن خواهیم پرداخت.

تحلیل خوشه‌ای روش آماری است که به طور گسترده برای داده‌کاوی در رشته‌های مختلف به کار می‌رود (اوریت<sup>۶</sup>، ۱۹۹۳). اصطلاح تحلیل خوشه‌ای اولین بار توسط تریون<sup>۷</sup> در سال ۱۹۳۹ برای گروه‌بندی اشیایی که شبیه بودند مورد استفاده قرار گرفت. تحلیل خوشه‌ای به گروه‌بندی موضوعات و داده‌ها بر اساس اطلاعات داده‌ها و روابط آنها می‌پردازد (تن و همکاران<sup>۸</sup>، ۲۰۰۶). هر موضوعی خیلی نزدیک یا شبیه به موضوعات دیگر در یک گروه است (شباهت بیشتر، بهتر)، اما از موضوعات گروه‌های دیگر متفاوت است (تفاوت بیشتر، بهتر). تحلیل خوشه‌ای با یک گروه شروع می‌شود و با تشکیل زیرگروه‌هایی که از نظر متغیرهای انتخاب شده متفاوت هستند ادامه می‌یابد. خوشه‌بندی برای کشف و پی‌بردن به ساختار داده‌ها و همچنین برای یافتن ویژگی‌های مشترک در بین زیر گروه‌هایی از داده‌ها به کار می‌رود. از آن جایی که تحلیل خوشه‌ای اهمیت زیادی در شناسایی گروه‌های هدف و کسب اطلاعات کامل از ویژگی‌ها، باورها و رفتار این گروه‌ها دارد کاربرد آن در روانشناسی می‌تواند بسیار مفید باشد (ابل<sup>۹</sup>، پلامبریج و گراهام<sup>۱۰</sup>، ۲۰۰۲؛ مارتین، وایلی و کسول<sup>۱۱</sup>، ۱۹۹۲؛ ماسچسکی - اسنیدر و گریسر<sup>۱۲</sup>، ۱۹۸۹؛ وایلی و کسول، ۱۹۹۷) هدف این مقاله معرفی این تکنیک به عنوان روشی سودمند، قوی و معتبر در روان‌شناسی است.

تحلیل خوشه‌بندی از مجموعه‌ای از الگوریتم‌ها یا قواعد برای تقسیم‌بندی ماتریس تقریب استفاده می‌کند. یکی از کاربردهای خوشه‌بندی در روان‌شناسی خوشه‌بندی افراد است. بیشتر روان‌شناسان در خوشه‌بندی به تشابه گروه‌های افراد در مجموعه متغیرها توجه دارند. شناسایی زیر

1. Data mining

4. Kiss

7. Tryon

10. Plumridge, &amp; Graham

12. Maschewsky-Schneider &amp; Greiser

2. Bunge

5. Trybula

8. Tan &amp; et al

11. Martin, Wyllie, &amp; Casswell

3. Judson

6. Everitt

9. Abel

گروه‌هایی از افراد مشابه در آزمون‌های روان‌شناختی اهمیت زیادی دارد. و این خوشه‌بندی با ماتریس اشخاص متغیرها صورت می‌گیرد به نحوی که اشخاص در ردیف و متغیرها در ستون‌ها قرار می‌گیرند. سپس خوشه‌بندی برای تشکیل گروه‌هایی با شبیه‌ترین ردیف‌ها (از اشخاص) در ماتریس صورت می‌گیرد. دومین کاربرد مهم خوشه‌بندی در روان‌شناسی خوشه‌بندی متغیرهاست که در نقطه مقابل رویکرد ماتریس اشخاص متغیرها است. در اینجا هدف ایجاد مجموعه مشابه از متغیرها است که نشان دهنده الگوهای مشابه از همبستگی در بین گروهی از افراد است. این رویکرد شبیه تحلیل عاملی است و توسط پژوهشگران برای خوشه‌بندی تعداد زیادی از متغیرها و برای تشکیل متغیرهای ترکیبی جدید به کار رود.

از خوشه‌بندی برای شناسایی گروه‌هایی از مردم در یک متغیر و یا شناسایی الگوهای تغییر نیز می‌توان استفاده کرد. برای مثال روانشناس می‌تواند از مراجعان خود بخواهد که مقیاس افسردگی را تکمیل و سپس با تحلیل خوشه‌ای تفاوت بین گروه‌های پاسخ‌دهنده به درمان به روش‌های مختلف را شناسایی کند. به طور کلی ممکن است محقق بر گروه‌بندی موضوعات یا متغیرها تاکید داشته باشند. در برخی موارد محقق ماتریس افراد متغیرها را ندارد اما قادر به تشکیل یک ماتریس تقریب و کاربرد یک الگوریتم خوشه‌بندی به طور مستقیم است. در همه موارد یک ماتریس تقریب و یک روش خوشه‌بندی وجود دارد (اوریت، ۱۹۸۰).

رامسبرگ<sup>۱</sup> (۱۹۸۴) اهداف خوشه‌بندی را کشف<sup>۲</sup>، تایید<sup>۳</sup> و ساده‌سازی<sup>۴</sup> بیان کرده است. در مواقعی که حوزه مورد مطالعه نسبتاً جدید است، خوشه‌بندی گام اولیه‌ای برای شناسایی و ساخت زیر گروه‌هایی با توجه به مساله پژوهش است. بعد از شناسایی ساختار به این طریق، فرایند پژوهش باید با آزمون تایید و اعتباریابی ساختار ادامه پیدا کند. تحلیل خوشه‌ای همچنین برای آزمودن طبقه‌بندی مبتنی بر رویکردها و نظریه‌های قبلی نیز به کار می‌رود که در پژوهش‌های روانشناسی کاربرد کمتری دارد. تحلیل خوشه‌ای همچنین تکنیکی مفید برای ساده‌سازی مجموعه داده‌هاست. لار<sup>۵</sup> (۱۹۸۳)، آلدندرفر و بلاشفیلد<sup>۶</sup> (۱۹۸۴) بلاشفیلد (۱۹۸۰) روانشناسان معروفی هستند که از تحلیل خوشه‌ای استفاده فراوانی کرده‌اند.

در مقایسه تحلیل خوشه‌ای و تحلیل عاملی می‌توان گفت هر دو روش‌هایی برای شناسایی ساختار زیر بنایی در مجموعه داده‌های چند متغیره هستند هر دو تحلیل برای کاهش داده‌ها به کار می‌روند و مبتنی بر قضاوت محقق و رویکرد مورد استفاده توسط او هستند. بورگن و ویس<sup>۷</sup> (۱۹۷۱) در تفاوت این دو روش معتقدند که در تحلیل عاملی معمولاً واریانس بین چندین منبع یا عامل در نظر گرفته می‌شود و مقدار اولیه واریانس موجود در داده‌های اصلی حفظ می‌شود در حالی که در تحلیل خوشه‌ای کل واریانس برای یک منبع در نظر گرفته می‌شود.

از طرفی در حالی که نتایج تحلیل عاملی نوعاً وضوح کمتری دارد، تحلیل خوشه‌ای نتایجی به دست می‌دهد که در آن متغیرها به مجموعه‌های جدا یا خوشه‌ها گروه‌بندی می‌شوند. تفاوت تحلیل خوشه‌ای و تحلیل تشخیصی در این است که در تحلیل تشخیصی گروه‌ها (خوشه‌ها) از قبل تشکیل شده‌اند و هدف از تحلیل آن است که ترکیب خطی آن دسته از متغیرهای مستقلی که گروه‌ها را به بهترین نحو از یکدیگر تفکیک می‌کنند تعیین شود، اما در تحلیل خوشه‌ای گروه‌ها از قبل تعیین نشده‌اند بلکه هدف این است که بهترین روشی را که بتوان از طریق آن متغیرها را در گروه‌هایی مشخص خوشه‌بندی کرد تعیین شود. برای مثال وقتی محقق نیاز به پیدا کردن افرادی با الگوهای مشابه خرید است یا وقتی هدف بررسی زیر گروه‌هایی از بیماران با تشخیص افسردگی براساس یک

1. Romesburg

2. Exploration

3. Confirmation

4. simplification

5. Lorr

6. Aldenderfer &amp; Blashfield

7. Borgen &amp; Weiss

چک لیست علائم و نتایج حاصل از آزمون‌های روانی است تحلیل خوشه‌ای مفید است. تحلیل تمیز نیازمند شناخت اعضای گروه براساس موضوع طبقه‌بندی است. ولی هدف تحلیل خوشه‌ای شناسایی گروه‌های واقعی است. بنابراین در تمایز بین چندین گروه مبتنی بر یک قاعده تحلیل تمیز مناسب است اما در تحلیل خوشه‌ای محقق نمی‌داند چه کسی یا چه چیزی به هر گروه تعلق دارد. در تحلیل خوشه‌ای گاهی شناخت به اعضای گروه‌ها وجود ندارد (آلدنדרفر و بلاشفیلد، ۱۹۸۴).

تحلیل خوشه‌ای با اندازه‌گیری مشاهدات مربوط به  $k$  متغیر آغاز می‌شود. سپس فاصله بین هر جفت مشاهده اندازه‌گیری می‌شود. برای تشکیل زیرگروه‌ها از الگوریتم‌هایی استفاده می‌شود. هدف از تحلیل خوشه‌ای ایجاد خوشه‌هایی با کمترین واریانس درون خوشه‌ای و بیشترین واریانس بین خوشه‌ای است. و در نهایت خوشه‌ها با هم مقایسه می‌شوند. تفاوت بین خوشه‌ها می‌تواند بیانگر اعتبار و ارزش خوشه‌های به دست آمده باشد. وقتی محقق تحلیل خوشه‌ای را راهبردی مناسب و مفید برای کاهش داده‌ها می‌داند، گام‌های زیر را طی می‌کند. این گام‌ها پیچیده هستند و محقق باید تصمیمات مهمی در طی این گام‌ها بگیرد که تصمیمات محقق نقش مهمی در تعیین نوع خوشه‌ها دارد.

۱. اندازه‌گیری تقریب یا مقیاس شباهت: مفاهیم فاصله و تشابه از مفاهیم اساسی تحلیل خوشه‌ای است. فاصله اندازه‌ای است که نشان می‌دهد دو مشاهده تا چه حد جدا از یکدیگرند. در حالی که تشابه شاخص نزدیکی آنها با یکدیگر است. تحلیل خوشه‌ای به گروه‌بندی افراد بر اساس شباهت آنها در متغیرهای انتخاب شده می‌پردازد. روش‌های زیادی برای ارزیابی شباهت وجود دارد از جمله مجذور فاصله اقلیدسی<sup>۱</sup> و همبستگی پیرسون<sup>۲</sup>. انتخاب مقیاس‌های شباهت تاثیر مهمی در نتایج تحلیل خوشه‌ای دارد. این مقیاس‌ها با توجه به الگوریتم تشکیل خوشه، ماهیت متغیرها (پیوسته، گسسته یا هر دو ارزشی) و مقیاس اندازه‌گیری انتخاب می‌شوند. برای مثال اگر هدف محقق گروه‌بندی افراد با توجه به سطح اضطراب آنها (افراد با اضطراب بالا در مقابل افراد با اضطراب پایین) باشد روش مجذور فاصله اقلیدسی مناسب و اگر هدف بررسی و طبقه‌بندی الگوی تغییر اضطراب (کاهش اضطراب، افزایش اضطراب یا بدون تغییر) باشد استفاده از روش همبستگی پیرسون مناسب است.

۲. انتخاب الگوریتم یا روش خوشه‌بندی: به طور کلی هدف از اجرای تحلیل خوشه‌ای پاسخ به این سوال است که آیا یک گروه خاص می‌تواند به زیر گروه‌هایی با ویژگی‌های متفاوت تقسیم شود؟ زیر گروه‌ها یا خوشه‌ها می‌توانند با توجه به ویژگی‌های رایج اعضای گروه از قبیل میانگین گروه یا بیشترین ارزش‌ها در گروه‌ها نامگذاری و تعریف شوند. روش‌های مختلفی برای اجرای تحلیل خوشه‌ای وجود دارد و انتخاب الگوریتم‌های دسته‌بندی مناسب جهت استفاده مفید از تجزیه و تحلیل خوشه‌ها بسیار مهم است (پانچ<sup>۳</sup> و استوارت<sup>۴</sup>، ۱۹۸۳).

تکنیک‌های تحلیل خوشه‌ای می‌تواند به صورت‌های زیر تقسیم شود (تن و همکاران، ۲۰۰۶):  
 (۱) سلسله مراتبی در مقابل غیر سلسله مراتبی<sup>۵</sup> یا روش جزیندی تکرار شونده (۲) قطعی در مقابل فازی<sup>۶</sup> و (۳) کامل در مقابل جزئی<sup>۷</sup>.

1. squared Euclidean distance  
 3. Punj  
 5. Hierarchical vs. Partitional  
 7. Complete vs. Partial

2. Pearson's correlation  
 4. Stewart  
 6. Exclusive vs. Overlapping vs. Fuzzy

اگرچه روش‌های مختلف خوشه‌بندی گسترش پیدا کرده‌اند، ادبیات موجود بیشتر به دو روش تراکمی<sup>۱</sup> و روش‌های افرازبندی تکراری<sup>۲</sup> تاکید دارند. روش تحلیل خوشه‌ای سلسله‌مراتبی از گام‌هایی تشکیل شده است. در این روش افراد گام به گام به خوشه‌های مشابه تقسیم می‌شوند. اگرچه بررسی خروجی برنامه توسط محققان نیاز به تصمیم‌گیری در مورد مناسب‌ترین تعداد تحلیل خوشه‌ای است (وارد، ۱۹۶۳).

الگوریتم‌های سلسله‌مراتبی طی مراحل پیشرفت، با اضافه نمودن یا کاستن اجزایی به خوشه‌ها یک ساختار درخت مانند را به وجود می‌آورد به این درخت سلسله‌مراتبی دندوگرام<sup>۳</sup> می‌گویند. (شکل ۲) که الگوریتم‌های افزایشی یا کاهش‌ی نامیده می‌شوند. در الگوریتم افزایشی یا پایین به بالا<sup>۴</sup> یا متراکم شونده<sup>۵</sup> ابتدا هر داده به عنوان خوشه‌ای مجزا در نظر گرفته می‌شود و در طی فرایندی تکراری در هر مرحله خوشه‌هایی که شباهت بیشتری با یکدیگر با یکدیگر دارند ترکیب می‌شوند تا در نهایت یک خوشه و یا تعداد مشخصی خوشه حاصل شود. پنج الگوریتم عمده افزایشی عبارتند از: اتصال منفرد<sup>۶</sup>، اتصال کامل<sup>۷</sup>، اتصال متوسط<sup>۸</sup>، روش مرکزی و روش وارد<sup>۹</sup> (هایر<sup>۱۰</sup>، آنرسون<sup>۱۱</sup>، تاتام<sup>۱۲</sup> و بلک<sup>۱۳</sup>، ۱۹۹۲) تفاوت بین این پنج الگوریتم در روش‌های ریاضی به کار رفته، برای محاسبه فاصله بین خوشه‌ها بکار می‌رود. هر الگوریتم دارای گرایش‌های سیستماتیک متفاوت جهت دسته‌بندی مشاهدات است. برای مثال روش مرکزی گرایش به ایجاد خوشه‌های نامنظم دارد. علاوه بر آن این روش فقط با داده‌های نسبی قابل استفاده است. روش وارد گرایش به ایجاد خوشه‌هایی با تعداد مشاهدات تقریباً یکسان دارد (میلیگان<sup>۱۴</sup>، ۱۹۸۰) با در نظر گرفتن این مسایل باید بین الگوریتم انتخابی با ساختار داده‌ها (از قبیل اندازه نمونه‌ها، توزیع، مشاهدات و نوع متغیرها) تناسبی وجود داشته باشد. در نتیجه برای مثال روش مرکزی در مواقعی مناسب است که اولاً تعداد مشاهدات در هر دسته تقریباً برابر باشد، ثانیاً متغیرهای غیر عادی نداشته باشیم.

در الگوریتم کاهش‌ی یا بالا به پایین<sup>۱۵</sup> یا بخش پذیر<sup>۱۶</sup> ابتدا تمام داده‌ها به عنوان یک خوشه در نظر گرفته می‌شوند و سپس در طی یک فرایند تکراری در هر مرحله داده‌هایی که شباهت کمتری به هم دارند به خوشه‌های مجزایی شکسته می‌شوند و این روال تا رسیدن به خوشه‌هایی که دارای یک عضو هستند ادامه پیدا می‌کند. دو نوع روش تکنیک کاهش‌ی وجود دارد: روش یگانه‌ای و چند گانه‌ای. روش‌های یگانه‌ای با متغیرهای دودویی استفاده می‌شوند. یک نمونه به یک گروه که همگی واجد یا فاقد ویژگی مشترک باشند تقسیم می‌شود. این گروه‌ها سپس به گروه‌های کوچکتری بر اساس وجود یا عدم وجود ویژگی دیگر شکسته می‌شوند.

از آن جایی که در این روش مشاهدات به طور متوالی گروه‌بندی می‌شوند و نه همزمان مناسب تحقیقات ترکیبی نیست. روش‌های کاهش‌ی چندگانه نقطه مقابل روش تراکمی است. در روش تراکمی ابتدا هر مشاهده به عنوان یک خوشه یا دسته مجزا بررسی شده و سپس آنها بصورت گروه‌های کوچکتر جمع‌بندی شده و در نهایت تمام مشاهدات در یک گروه قرار داده می‌شوند. این بستگی به نظر محقق دارد که تعیین کند در چه مرحله‌ای تعداد گروه‌ها کافی است. روش‌های کاهش‌ی چند گانه‌ای مراحل متفاوتی را طی می‌کنند. در ابتدا تمام مشاهدات در یک گروه قرار می‌گیرند، سپس مشاهدات به گروه‌های کوچکتر تقسیم می‌شوند تا نهایتاً هر مشاهده به صورت

1. hierarchical agglomerative methods

2. iterative partitioning methods

5. Agglomerative

8. Average-Link

11. Anerson

14. Milligan

3. Dendogram

6. Single-Link

9. Wards method

12. Tatham

15. Top-Down

4. Bottom-Up

7. Complete-Link

10. Hair

13. Black

16. Divisive

یک خوشه جداگانه در می‌آید. باز هم این محقق است که تعیین می‌کند چه زمانی این تقسیمات کافی است. هر چند که این دو روش از دو انتهای مخالف خوشه‌ها شروع می‌کنند ولی تعداد خوشه‌های مشخص شده صرف نظر از نوع روش انتخابی باید یکسان باشد. در نتیجه هیچ کدام از این دو روش برتری عمده‌ای نسبت به دیگری ندارد (اوریت، ۱۹۸۰).

تمام الگوریتم‌های سلسله مراتبی مشکل دارند. اول اینکه محققین به دلیل اینکه از قبل ساختار زیر بنایی یک نمونه را نمی‌دانند، انتخاب الگوریتم صحیح برایشان مشکل است، دوم اینکه الگوریتم فقط یک بار داده‌ها را مورد بررسی قرار می‌دهد، لذا دسته‌های ضعیف نمی‌توانند تعیین شوند. و بالاخره راه حل‌ها معمولاً وقتی که اندازه یک نمونه کوچک باشد، ناپایدارند و لذا در نمونه‌های کوچک راه حلی که فقط با استفاده از روش سلسله مراتبی به دست آید خیلی معتبر نیست (جاردين و سیسون<sup>۱</sup>، ۱۹۷۱).

یکی از الگوریتم‌های خوشه‌بندی سلسله مراتبی، الگوریتم خوشه‌بندی دومرحله‌ای است. این روش، ابزاری اکتشافی است که برای آشکار نمودن خوشه‌های ذاتی و طبیعی موجود در مجموعه داده‌ها که به طور معمول دیده نمی‌شوند، طراحی شده است. این روش در مورد داده‌های بزرگ و استفاده توأم از داده‌های ترتیبی و کمی کاربرد دارد. به واقع تنها الگوریتمی است که می‌تواند استفاده توأم از داده‌های کمی و کیفی را فراهم سازد. در گام اول مشاهدات در قالب خوشه‌های اولیه قرار می‌گیرند و این پیش خوشه‌ها به عنوان یک هسته اولیه و به عنوان یک مشاهده قرار می‌گیرد. در گام دوم از روش سلسله مراتبی برای طبقه‌بندی این هسته‌ها که خروجی گام قبل هستند و مشاهدات مشابه را در یک هسته قرار داده‌اند، استفاده می‌شود.

الگوریتم تحلیل خوشه‌ای دومرحله‌ای بدین صورت است که با فرض مستقل بودن متغیرها در مدل خوشه‌بندی، اندازه درست‌نمایی<sup>۲</sup> را به عنوان فاصله موجودیت‌ها در نظر می‌گیرد. هم چنین، فرض بر این است که هر متغیر پیوسته دارای توزیع نرمال و هر متغیر گسسته دارای توزیع نرمال چندمتغیری است. البته آزمون‌های تجربی نشان می‌دهد که الگوریتم دو مرحله‌ای نسبت به تغییر فرض‌های استقلال و نرمال بودن، به اندازه کافی مقاوم است. الگوریتم خوشه‌بندی دومرحله‌ای به صورت زیر خلاصه می‌شود: مرحله اول. در این مرحله که مرحله پیش خوشه<sup>۳</sup> می‌باشد، بر اساس رویکرد خوشه‌بندی سلسله مراتبی تجمعی<sup>۴</sup>، هر رکورد به عنوان یک خوشه در نظر گرفته می‌شود. تمامی رکوردها یک به یک بررسی شده و بر اساس معیار در نظر گرفته شده برای فاصله موجودیت‌ها، در مورد ادغام آن با خوشه‌های قبل یا آغاز یک خوشه جدید تصمیم‌گیری می‌شود. بنابراین در این مرحله به زیرخوشه‌هایی دست می‌یابیم. در مرحله دوم زیرخوشه‌های مرحله اول به عنوان ورودی‌ها در نظر گرفته شده و بر اساس خوشه‌بندی سلسله مراتبی تجمعی به تعداد خوشه‌های مورد نظر تبدیل می‌شوند. در صورت عدم تعیین تعداد خوشه‌ها، این الگوریتم به صورت خودکار و با معیارهایی هم چون  $BIC^5$  و  $AIC^6$  زیرخوشه‌ها را به تعداد بهینه‌ای از خوشه‌ها تبدیل می‌کند.

روش‌های تکرار شونده یا غیر سلسله مراتبی از دو جهت با روش‌های سلسله مراتبی متفاوت هستند. اول، تعداد خوشه‌ها قبل از تحلیل مشخص هستند دوم، مشاهده‌ها می‌توانند در طی فرایند خوشه‌بندی برای بهینه‌شدن خوشه‌ها، از یک خوشه به خوشه دیگر جابه جا شوند. در این روش سری داده‌ها را به تعداد از قبل تعیین شده‌ای از خوشه‌ها تقسیم‌بندی می‌کند. روش‌های مختلف

1. Jardine and Sibson  
2. Likelihood Measure  
3. Pre cluster  
4. Agglomerative hierarchical clustering method  
5. Schwarz's Bayesian Criterion  
6. Akaike Information Criterion

غیر سلسله مراتبی باهم تفاوت دارند ولی همگی براساس یک روش عمل می‌کنند (هایر و همکاران، ۱۹۹۲). بعد از اینکه مرکز هر خوشه انتخاب گردید هر مشاهده‌ای به گروهی که به مرکز آن نزدیک‌تر است مربوط می‌شود. همچنان که یک مشاهده جدید اضافه می‌شود، مراکز خوشه‌ها مجدداً محاسبه می‌گردد. در نتیجه چندین بار سری داده‌ها بررسی می‌شود تا مشاهده‌ها، خوشه مربوطه‌شان را براساس فاصله‌شان از نقطه مرکزی محاسبه شده جدید، تغییر دهند. برای رسیدن به مطلوب‌ترین راه‌حل، بررسی سری داده‌ها همچنین ادامه می‌یابد تا جایی که دیگر هیچ مشاهده‌ای دسته مربوطه را عوض ننماید (آندربرگ<sup>۱</sup>، ۱۹۷۳).

یکی از معروف‌ترین و پرکاربردترین روش‌های غیر سلسله مراتبی روش  $k$  میانگین<sup>۲</sup> است. این روش اولین بار توسط مک کوئین (۱۹۶۷) ارائه شد. این روش برای خوشه‌بندی داده‌هایی طراحی شده است که به صورت کمی باشند. خوشه دارای مرکزی به نام میانگین است. در ابتدا داده‌ها به صورت تصادفی به  $k$  خوشه تقسیم می‌شوند. در مرحله بعد فاصله هر یک از داده‌ها از مرکز خوشه خود، که همان میانگین مشاهدات هر خوشه است، محاسبه می‌شود. در صورتی که فاصله هر مشاهده مورد نظر از میانگین خوشه خود زیاد و به خوشه دیگری نزدیک‌تر باشد، این مشاهده به خوشه‌ای که نزدیک‌تر است اختصاص می‌یابد. این کار تا حداقل شدن تابع خطا، که معمولاً مجموع فواصل مشاهدات از مرکز خوشه خودش است، و یا تغییر نیافتن اعضای خوشه‌ها ادامه می‌یابد (میلیگان، ۱۹۸۰).

روش‌های غیر سلسله مراتبی نسبت به روش‌های سلسله مراتبی دو مزیت بالقوه دارد. اول آنکه به علت داشتن امکان عوض کردن خوشه‌های مربوطه توسط مشاهدات، روش‌های غیر سلسله مراتبی کمتر تحت فشار اجزای بیرونی قرار می‌گیرند. اگر چه اجزای بیرونی می‌توانند خوشه‌ها را دچار ابهام نمایند، ولی این اشکال با بررسی‌های پی در پی داده‌ها که طی آن اعضای خوشه‌ها تغییر می‌یابند، تصحیح می‌گردد (آلدنر فر و بلشفیلد، ۱۹۸۴). دوم آنکه به علت بررسی‌های چند باره داده‌ها خوشه‌های همگن و ناهمگن، بهینه می‌شوند. رسیدن به این نتیجه مستلزم آن است که تعداد خوشه‌ها استقرایی باشد در بسیاری از زمینه‌ها این مساله ممکن نیست چون تجزیه و تحلیل خوشه‌ها عموماً به صورت اکتشافی هستند (میلیگان، ۱۹۸۰).

با توجه به مشکلات هر یک از روش‌های سلسله مراتبی و غیر سلسله مراتبی، میلیگان (۱۹۸۰) پیشنهاد می‌کند که ابتدا روش سلسله مراتبی برای تعیین تعداد خوشه‌ها اجرا شود و سپس تحلیل  $k$  میانگین برای بهینه شدن نتایج اجرا شود. این راه حل توسط بسیاری از متخصصین پیشنهاد شده است. بدین ترتیب که ابتدا الگوریتم سلسله مراتبی استفاده شود تا تعداد دسته‌ها و مراکز آنها مشخص گردد. سپس این نتایج به عنوان نقطه شروع برای دسته‌بندی غیر سلسله مراتبی استفاده شود (هایر و همکاران، ۱۹۹۲؛ میلیگان، ۱۹۸۰). تحقیقات نشان داده است که استفاده از هر دو روش باعث افزایش اعتبار راه حل‌ها می‌گردد اگرچه استفاده از هر دو روش سلسله مراتبی و غیر سلسله مراتبی مستلزم وقت و انرژی بیشتری است ولی بهترین راه حل‌ها در استفاده از هر دو روش به دست می‌آیند. اس پی اس سه روش متفاوت از تحلیل خوشه‌ای را اجرا می‌کند: تحلیل خوشه‌ای سلسله مراتبی<sup>۳</sup>، تحلیل خوشه‌ای  $k$  میانگین<sup>۴</sup> و تحلیل خوشه‌ای دو مرحله‌ای<sup>۵</sup>. در صورتی که حجم نمونه زیاد باشد و یا ترکیبی از متغیرهای پیوسته و طبقه‌ای وجود داشته باشد تحلیل دو گامی، در صورتی که حجم نمونه کم و هدف یافتن راه حلی ساده برای افزایش تعداد خوشه‌ها باشد تحلیل سلسله مراتبی و در نهایت برای پی بردن به تعداد خوشه‌ها در یک حجم نمونه متوسط

1. Anderberg

2. k-means

3. Hierarchical cluster analysis

4. K-means cluster

5. Two-step cluster

تحلیل k میانگین مناسب است.

۳. تعیین تعداد خوشه‌ها: روش‌های مختلفی برای تعیین تعداد خوشه‌ها وجود دارد. همه این روش‌ها ذهنی هستند اوریت<sup>۱</sup> (۱۹۹۳) در مورد تاثیر انتظارات قبلی محقق بر تصمیم‌گیری هشدار داده است. محقق نیاز به قوانینی برای تصمیم‌گیری در مورد تعداد خوشه‌ها دارد این قوانین گاهی به عنوان قوانین توقف<sup>۲</sup> شناخته می‌شوند و نشان دهنده زمانی است که فرایند خوشه‌بندی N مشاهده به یک خوشه متوقف می‌شود. مستقیم‌ترین روش تصمیم‌گیری در مورد تعداد خوشه‌ها بررسی جدول الگوریتم<sup>۳</sup> و دندوگرام است. دندوگرام یک نمودار است که نشان دهنده ترتیب الحاق مشاهدات است و به طور چشمی مورد بررسی قرار می‌گیرد. یک محقق خوشه‌های طبیعی داده‌ها را که توسط شاخه‌های انبوه مشخص می‌شوند، جستجو می‌کند. این روش بستگی به نحوه تفسیر محقق دارد، در نتیجه باید با احتیاط استفاده شود (آلدندر فر و بلاشفیلد، ۱۹۸۴). یکی دیگر از روش تصمیم‌گیری درباره تعداد خوشه‌ها بر مبنای ملاک بیزی و ملاک آکاییک است. جوابی که در سطح حداقل ملاک بیزی و ملاک آکاییک که به دست می‌آید می‌تواند معرف بهترین تعادل موجود بین دقت و پیچیدگی باشد. همچنین راه دیگر برای تصمیم‌گیری در زمینه تعداد خوشه‌ها، استفاده از نسبت فاصله است. بهینه‌ترین تعداد خوشه‌ها هنگامی است که تغییری بزرگ در نسبت فاصله مشاهده می‌گردد. وقتی که روش سلسله مراتبی استفاده می‌شود. یک نظریه استقرایی نیز می‌تواند به عنوان یک ابزار غیر آماری جهت تعیین تعداد خوشه‌ها به کار رود (هایر و همکاران، ۱۹۹۲) اگرچه یک نظریه استقرایی در تحقیقات اکتشافی ضروری نیست، ولی می‌تواند در ارزیابی نتایج حاصله از بررسی‌ها نظری به کار رود (کتچن<sup>۴</sup> و هیوت<sup>۵</sup>، ۲۰۰۰). به طور خلاصه استفاده از یک روش خاص به منظور تعیین تعداد خوشه‌ها سوال برانگیز است زیرا هر روشی محدودیت‌هایی دارد (اوریت، ۱۹۸۰).

۴. ارزیابی اعتبار خوشه‌ها: روش‌های مختلفی برای تعیین ثبات خوشه‌ها وجود دارد. رایج‌ترین روش تقسیم تصادفی نمونه به دو گروه و تکرار تحلیل خوشه‌ای در مورد هر یک از آنها است. اگر تحلیل خوشه‌ای با ثبات باشد، ساختار خوشه‌ای مشابه در هر دو نیمه از نمونه به دست می‌آید. راه دیگر، تکرار تحلیل خوشه‌ای در نمونه‌ای دیگر از همان جامعه است. جولیف، جونز و مورگان<sup>۶</sup> (۱۹۸۲) فرایند حذف تعدادی متغیر و اجرای مجدد تحلیل برای بررسی ثبات نتایج و کوس و سنیکا<sup>۷</sup> (۲۰۰۰) و کتچن و هیوت (۲۰۰۰) تکرار تحلیل خوشه‌ای با روش خوشه‌بندی متفاوت را برای بررسی ثبات پیشنهاد کردند. ثبات خوشه‌ها لازم و ضروری است اما برای اعتبار کافی نیست (آلدندر فر و بلاشفیلد، ۱۹۸۴). اعتبار خوشه‌ها باید شامل شواهدی از ارزش آنها در زمینه مورد مطالعه باشد. لذا ارزیابی اعتبار خوشه‌های تعیین شده از گام‌های اساسی در تحلیل خوشه‌بندی است (آلدندر فر و بلاشفیلد، ۱۹۸۴؛ بلاشفیلد، ۱۹۸۰؛ گانجستاد<sup>۸</sup> و اسنیدر<sup>۹</sup>، ۱۹۸۵؛ اسکینر و بلاشفیلد، ۱۹۸۲). اهداف اعتبار بخشی عبارتند از حصول اطمینان از اینکه راه‌حل خوشه‌ای اعتبار خارجی دارد و نیز اعتبار ضابطه‌ای دارد و یا به عبارتی برای پیشگویی نتایج مهم مفید است (کرلینجر، ۱۹۸۶). دقت در اعتبار بخشی، موجب اطمینان است. زیرا علیرغم دقت در مراحل پیشین، بدون اعتبار نمی‌توان از دستیابی به یک سری خوشه مفید و معنی‌دار اطمینان حاصل نمود (پانچ و استوارت، ۱۹۸۳). قابلیت اطمینان امری لازم است ولی شرط کافی اعتبار نیست (کرلینجر، ۱۹۸۶) دو روش اولیه برای ارزیابی قابلیت اطمینان وجود دارد. در روش اول، محقق می‌تواند یک

1. Everitt

4. Ketchen

7. Kos and Psenicka

2. 'stopping rules

5. Huit

8. Gangestad

3. agglomeration schedule

6. Jolliffe, Jones, and Morgan

9. Snyder



تحلیل خوشه‌ای چند مرتبه با تغییر الگوریتم‌ها و روش‌ها، برای مرتب‌کردن چندین هم خطی انجام دهد، درجه و میزان سازگاری در راه حل‌ها نمایانگر قابلیت اطمینان است (هایر و همکاران، ۱۹۹۲).

در روش دوم محقق می‌تواند یک نمونه را تقسیم نموده و دو نمونه را جداگانه تحلیل کنند (همبریک، ۱۹۸۳). مدل اصلاح شده این روش، به دست آوردن مراکز ثقل خوشه‌ها از یک نیمه نمونه و استفاده از آنها به منظور تعریف خوشه‌ها در نیمه دیگر است. در هر صورت سازگاری در میان نمونه‌های دو نیم شده، نشانگر قابلیت اطمینان است (هایر و همکاران، ۱۹۹۲). لذا استانداردی جهت بیان سطح و میزان قابل قبول سازگاری وجود ندارد و این مساله، تشخیص را تا حد زیادی به قضاوت محققین واگذار می‌کند. چنانچه قابلیت اطمینان ثابت شده باشد، می‌توان توجه را به اعتبار خارجی معطوف نمود. این امر می‌تواند هم از طریق تحلیل خوشه نمونه مورد علاقه و یک نمونه مشابه ثانوی و سپس از طریق بیان تشابه نتایج انجام شود (هایر و همکاران، ۱۹۹۲). از نمونه‌های چند گانه برای اعتبار بخشی، تنها هنگامی که متناسب و سازگار با فرضیات یک بررسی هستند می‌توان بهره‌گیری کرد.

اعتبار ضابطه‌ای را نیز می‌توان از طریق آزمایش‌های معنی‌دار با متغیرهای خارجی مورد بررسی قرار داد (غالباً آنالیز واریانس درجه ناپایداری) (آلندر فر و براشفیلد، ۱۹۸۴؛ آندربرگ، ۱۹۷۳). در مجموع ارزش روش‌هایی که به منظور بررسی قابلیت اطمینان و اعتبار خارجی وجود دارند محدود است، چرا که آنها وابسته به قضاوت محقق هستند. در کل وقتی یک طرح تحقیقاتی از روش‌های خوشه‌ای و جداسازی بهره می‌گیرد. قابلیت اطمینان و اعتبار آن پرسش برانگیز است. تنها وقتی تحلیل خوشه‌ای به تکنیک‌های دیگری افزوده می‌شود می‌توان مطمئن بود که نتایج به دست آمده قطعی است.

در ادامه با کاربرد تجزیه و تحلیل خوشه‌ای به شناسایی زیر گروه‌هایی معتبر از تعلل‌ورزان خواهیم پرداخت. تعلل‌ورزی<sup>۱</sup> صفتی روان‌شناختی است که در زندگی و در میان قشرها و رده‌های سنی مختلف دیده می‌شود. به این پدیده در محیط‌های آموزشی میان یادگیرندگان، بسیار شایع است و با پیامدهای زیان بار از جمله گرفتن نمره پایین کلاسی تا انصراف‌های دوره‌ای همراه است. تعلل‌ورزی پدیده رایجی است که ایجاد کننده یک رفتار ثابت است (گاستاوسون<sup>۲</sup>، میاک<sup>۳</sup>، هویت<sup>۴</sup> و فریدمن<sup>۵</sup>، ۲۰۱۴) و با عوامل روانشناختی زیادی ارتباط دارد (استیل<sup>۶</sup>، ۲۰۰۷). علیرغم مطالعات و شواهد مرتبط با عوامل کلیدی مرتبط با تعلل‌ورزی، هنوز فهم منسجمی در مورد آن وجود ندارد (ویلسون<sup>۷</sup> و نگاین<sup>۸</sup>، ۲۰۱۲) در واقع اگرچه تعلل‌ورزی به عنوان عاملی موثر از عوامل شناختی، هیجانی و انگیزشی در نظر گرفته شده است (پیچل<sup>۹</sup> و فلت<sup>۱۰</sup>، ۲۰۱۲) ولی این عوامل به طور مستقل از هم مورد بررسی قرار گرفته‌اند.

در بررسی عوامل مؤثر بر تعلل‌ورزی تحصیلی مؤلفه‌های زیادی در پژوهش‌های مختلف شناسایی شده‌اند (برای مثال براون<sup>۱۱</sup>، ۱۹۸۳؛ شونبرگ<sup>۱۲</sup>، ۲۰۰۴). با وجود شناسایی عوامل متعدد مؤثر بر تعلل‌ورزی و پیامدهای آن و مدل‌های مختلفی که توسط محققان مطرح شده است، ولی همه محققان در مورد تأثیر عوامل شناختی، هیجانی و رفتاری بر تعلل مؤثر توافق دارند. لذا در مطالعه حاضر بر چندین پیش‌بینی کننده تعلل‌ورزی از بعد شناختی، عاطفی و رفتاری تأکید شده است.

1. Procrastination  
4. Hewitt  
7. Wilson  
10. Flett

2. Gustavson  
5. Friedman  
8. Nguyen  
11. Brown

3. Miyake  
6. Steel  
9. Pychyl  
12. Schouwenburg

چارچوب نظری پژوهش حاضر رویکرد عقلانی، عاطفی، رفتاری (REBT) است. این نظریه، نظریه جامعی از نظریه رفتار انسانی (فراگات<sup>۱</sup>، ۲۰۰۵)، و یک روان‌درمانی انسان محور (الیس، ۱۹۷۳) است. ایده اصلی در این نظریه این است که مردم به طور شناختی، هیجانی و رفتاری زندگی می‌کنند. لذا تفکر آنها در هم تنیده با هیجان و رفتار آنهاست، و در نتیجه به ندرت به طور خالص عمل می‌کنند (الیس، ۱۹۷۹). به عبارتی، شناخت‌ها، هیجان‌ها و رفتارها همگی بر یکدیگر اثر می‌گذارند (الیس، ۱۹۶۲). علی‌رغم این درهم تنیدگی، اندیشه‌های افراد اثر قوی بر هیجان‌ها آنها دارد (الیس و دریدن، ۱۹۹۷). الیس (۱۹۶۲) مدلی به نام مدل A-B-C را مطرح نمود. در این مدل رویداد در حال فعالیت B باور افراد در مورد آن رویداد، و C نشان‌دهنده پاسخ‌های هیجانی یا رفتاری است. همسو با الیس (۱۹۶۲)، میچنابام و باتلر (۱۹۸۰) نیز بر روابط بین ابعاد شناختی، عاطفی و رفتاری تأکید دارند. آنها معتقدند شناخت افراد اثر مهمی بر هیجان و رفتار آنها دارد. در نظریه منطقی الیس، اولین گام، شناخت و درک رویدادها، عقاید و پیامدها یا همان (A-B-C) است. الیس معتقد است راه درست تغییر دادن پیامدهای ناراحت کننده رویدادها، نه بررسی کردن رویدادها (A) و نه تجزیه و تحلیل پیامدهای عاطفی و رفتاری (C) هاست، بلکه راه درست، تغییر دادن مستقیم عقاید (B) غیرمنطقی فرد است.

همان‌طور که گفته شد تعلل‌ورزی شامل ابعاد شناختی، عاطفی و رفتاری است به ویژه جنبه عاطفی تعلل‌ورزی با روحیه و هیجان‌ها در ارتباط است، جنبه شناختی آن بر عوامل غیرمنطقی و باورهای ناامیدی تأکید دارد و منجر به نتایج منفی می‌شود و مؤلفه رفتاری تعلل‌ورزی به رفتارهای تأخیری در مطالعه اشاره دارد (روتبلاد<sup>۲</sup>، سالومن<sup>۳</sup> و ماراکامی<sup>۴</sup>، ۱۹۸۶).

باورهای ناامیدی مبتنی بر امتناع از پذیرش تفاوت بین آرزو (خواسته) و واقعیت است (هارینگتون<sup>۵</sup>، ۲۰۰۶). نیکولسن<sup>۶</sup> و شارف<sup>۷</sup> (۲۰۰۷) نشان دادند که تحمل ناامیدی جز مهمی برای احساس شادی است. در حالی که برخی محققان بیان می‌کنند که ناامیدی دلیلی برای تعلل‌ورزی است (هارینگتون، ۲۰۰۵). دریدن (۲۰۰۳) نشان داد که تحمل ناامیدی نوعی از باورهای منطقی است. افراد با تحمل بالا ممکن است تصور کنند که ممکن است تحمل یک رویداد منفی یا موقعیت سخت باشد ولی قابل تحمل است. این باور شامل دو بخش است. در بخش اول، تحمل رویداد یا موقعیت سخت به نظر می‌رسد ولی در بخش دوم رویداد یا موقعیت قابل تحمل به نظر می‌رسد. لذا رویکرد عقلانی، عاطفی، رفتاری بر اهمیت تحمل ناامیدی تأکید دارد چرا که انسان‌ها هیجان‌ها منفی را تجربه می‌کنند و می‌توانند مشکلات را حل کنند. ناامیدی به نظر می‌رسد مانعی در مسیر رفتار هدفمند است. به طور مشابه، الیس و کناس (۱۹۷۷) بیان می‌کنند که عدم تحمل ناامیدی ایجاد کننده تعلل‌ورزی است و تعلل‌ورزانی که معتقدند زمان کافی برای تکمیل تکلیف خود دارند و تحمل ناامیدی کمتری دارند، تمایل دارند خود را تنبل یا بی‌انگیزه معرفی کنند (فروهلچ<sup>۸</sup>، ۱۹۸۷).

باورهای غیرمنطقی یکی دیگر از متغیرهای انگیزشی مورد مطالعه است. از مشخصات تفکر غیرمنطقی، انعطاف ناپذیری، همسوس نبودن با قوانین طبیعی، جزمی بودن و ناکارآمد بودن است. مانند: «من حتماً باید در این امتحان موفق شوم وگرنه نمی‌توانم زندگی کنم. باورها مهم‌ترین عامل در رویکرد عقلانی عاطفی رفتاری هستند (الیس، ۱۹۶۲). الیس باورهای غیرمنطقی را به صورت موانع شادی و انجام آرزوها تعریف می‌کند. مبتنی بر این نظریه، مطالعات زیادی رابطه باورهای

1. Froggat  
4. Murakami  
7. Scharff

2. Rothblum  
5. Harrington  
8. Froehlich

3. Solomon  
6. Nicholson

غیرعقلانی و تعلل‌ورزی را نشان داده‌اند (از جمله بریج و روج، ۱۹۸۸؛ بسویک و همکاران، ۱۹۸۸) ایس بیان می‌کند که باورهای غیرعقلانی اساساً از دو جنبه با تعلل‌ورزی ارتباط دارند. باور افراد به عدم لیاقت و باور به اینکه جهان خیلی سخت است (الیس و کناس، ۱۹۷۷).

باور خودکارآمدی یکی از عوامل شناختی مؤثر بر تعلل‌ورزی است که در این پژوهش مورد بررسی قرار گرفته است. شناخت به مشاهده یا ادراک فرد در مورد جهان اطرافش اشاره دارد. شناخت‌ها اجزای مهمی در رفتارهای انسان هستند. ایس<sup>۱</sup> و دریدن<sup>۲</sup> (۱۹۸۷) بیان می‌کنند که رویدادها ناامیدکننده انسان‌ها نیستند بلکه اندیشه آنها در مورد رویداد می‌تواند ناامیدی را در آنها ایجاد کند لذا ایس معتقد است که انسان‌ها با تفکر و اندیشه بهتر تمایل به بهتر عمل کردن را دارند. مطالعات مربوط به تعلل‌ورزی مبتنی بر چشم انداز شناختی به بررسی این سؤال می‌پردازند که چرا دانش‌جویان اگرچه با پیامدهای منفی تعویق آشنا هستند ولی باز به طور هوشیارانه تکالیف خود را به تعویق می‌اندازند.

خودتنظیمی عامل رفتاری مؤثر بر تعلل‌ورزی در این پژوهش است. واژه رفتار در رویکرد عقلانی عاطفی رفتاری به رفتار و عمل آشکار اشاره دارد، و یکی از بهترین روشها برای نظارت بر اصلاح رفتار است (فراگات، ۲۰۰۵). محققان زیادی (از جمله آندرسون<sup>۳</sup>، ۲۰۰۱ فراری<sup>۴</sup> و تیس<sup>۵</sup>، ۲۰۰۰؛ وولترز<sup>۶</sup>، ۲۰۰۳) به بررسی رابطه خودتنظیمی تعلل‌ورزی پرداخته‌اند و بیان می‌کنند که خودتنظیمی یکی از قوی‌ترین پیش‌بینی‌کننده‌های رفتاری تعلل‌ورزی است. تعلل‌ورزان در تنظیم عملکردشان در موقعیت‌های استرس‌زا و در موقعیت‌های با بار شناختی بالا ناتوان هستند. محققان زیادی (از جمله کلاسن و همکاران، ۲۰۰۸؛ ون ارد، ۲۰۰۳؛ هاول و واتسون، ۲۰۰۷) معتقدند که افراد تعلل‌ورز در تنظیم و نظم بخشی به رفتارهای خود مشکل دارند. تعلل‌ورزی به عنوان شکست در خودتنظیمی، بیانگر وجدان کم، و رفتار تکانشی بالا تعریف می‌شود (استیل، ۲۰۰۷).

از آن جایی که تعلل‌ورزی نتیجه مجموعه پیچیده‌ای از عوامل شناختی، هیجانی و انگیزشی است به نظر می‌رسد کاربرد تحلیل خوشه‌ای برای شناسایی گروه‌های مختلف تعلل‌ورز مناسب است. در واقع تحلیل خوشه‌ای یک رویکرد مبتنی بر شخص است که به جستجوی این مطلب می‌پردازد که چگونه متغیرهای متفاوت به خوشه‌های مختلف ترکیب می‌شوند که امکان بررسی چنین متغیرهایی را به طور مستقل از هم فراهم می‌کند.

### روش پژوهش

جامعه آماری پژوهش کلیه دانشجویان دانشگاه تهران در دانشکده‌های مختلف بوده است. این تعداد بنابه گزارش دانشگاه تهران در سال تحصیلی (۹۲-۹۳) بدون در نظر گرفتن پردیس‌های پاکدشت-ورامین، کرج، جلفا، کیش، قم و کرج ۳۱۷۰۷ است که از این تعداد ۱۷۲۸۴ نفر مرد و ۱۴۴۲۳ نفر زن هستند. نمونه‌ای برابر با ۹۰۰ نفر از دانشجویان به روش نمونه‌گیری تصادفی طبقه‌ای با انتساب متناسب انتخاب و به پرسشنامه‌ها پاسخ دادند. که از این تعداد ۴۴۷ نفر زن و ۴۵۳ نفر مرد هستند که با توجه به حجم طبقات (تعداد دانشجو در هر دانشکده) انتخاب شده‌اند. به دلیل متفاوت بودن حجم طبقات (تعداد دانشجو در هر دانشکده) با استفاده از روش نمونه‌گیری تصادفی طبقه‌ای متناسب، حجم نمونه مورد نظر تقسیم شد. به این ترتیب که با توجه به حجم طبقات از هر دانشکده تعدادی به صورت تصادفی انتخاب شده و سپس به طور تصادفی یک یا دو کلاس به عنوان نمونه، مورد بررسی قرار گرفتند و در هر کلاس، پرسشنامه‌ها به تعدادی از

دانشجویان حاضر در کلاس داده شد. از بین این ۹۰۰ نفر نمونه، ۲۰۰ دانشجو (۱۰۰ پسر و ۱۰۰ دختر) که در تعلل‌ورزی نمرات بالای به دست آورده بودند انتخاب شدند. میانگین سن نمونه مورد مطالعه ۲۱.۸۵ سال است.

### مقیاس‌ها

در پژوهش حاضر برای گردآوری داده‌ها از پنج ابزار به ترتیب زیر استفاده شد:

تعلل‌ورزی: خود تنظیمی در این پژوهش از طریق پرسشنامه سولومون و روثلوم<sup>۱</sup> (۱۹۸۴) اندازه گیری شد. این مقیاس دارای ۲۷ گویه بوده و در طیف لیکرت از "همیشه" تا "هرگز" تهیه شده است. ضریب قابلیت اعتماد این پرسشنامه با روش الفای کرنباخ محاسبه و برابر با ۰/۸۲ بدست آمد. به منظور شناسایی و تأیید عوامل اندازه‌گیری سازه تعلل‌ورزی، اعتبار سازه این پرسشنامه با انجام تحلیل عامل تأییدی مورد آزمون قرار گرفت. شاخص‌های نسبت مجذور کای به درجه آزادی (۲/۲۷)، نیکویی برازندگی (GFI=۰/۹۸)، شاخص تعدیل شده نیکویی برازش (AGFI=۰/۹۷) و ریشه خطای میانگین مجذورات تقریب (RMSEA=۰/۰۴) حاکی از برازش خوب مدل بود.

خودتنظیمی<sup>۲</sup>: خود تنظیمی در این پژوهش از طریق پرسشنامه گرین و میلر (۲۰۰۴) اندازه‌گیری شد. این مقیاس شامل برنامه‌ریزی و سازماندهی تلاش‌های مطالعه و خود نظارتی دانشجویان است و دارای ۱۲ گویه بوده و در طیف لیکرت از "همیشه" تا "هرگز" تهیه شده است. میلر و همکاران (۱۹۹۶) ضریب قابلیت اعتماد این مقیاس را با استفاده از روش آلفای کرونباخ ۰/۸۱ بدست آورده‌اند. ضریب قابلیت اعتماد این پرسشنامه با روش آلفای کرنباخ محاسبه و برابر با ۰/۹۶ بدست آمد. اعتبار سازه این پرسشنامه با انجام تحلیل عامل تأییدی مورد آزمون قرار گرفت. شاخص‌های نسبت مجذور کای به درجه آزادی (۲/۴۷)، نیکویی برازندگی (GFI=۰/۹۷)، شاخص تعدیل شده نیکویی برازش (AGFI=۰/۹۶) و ریشه خطای میانگین مجذورات تقریب (RMSEA=۰/۰۶) حاکی از برازش خوب مدل بود.

خود کارآمدی<sup>۳</sup>: در این پژوهش خودکارآمدی دانشجویان از طریق پرسشنامه تدوین شده توسط میدلتن و میگلی (۱۹۹۷) اندازه‌گیری شد. این مقیاس دارای ۱۰ گویه بوده و در طیف لیکرت از "همیشه" تا "هرگز" تهیه شده است. ضریب قابلیت اعتماد این پرسشنامه با روش الفای کرنباخ محاسبه و برابر با ۰/۹۸ بدست آمد. به منظور شناسایی و تأیید عوامل اندازه‌گیری سازه خودکارآمدی، اعتبار سازه این پرسشنامه با انجام تحلیل عامل تأییدی مورد آزمون قرار گرفت. شاخص‌های نسبت مجذور کای به درجه آزادی (۱/۴۶)، نیکویی برازندگی (GFI=۰/۹۹)، شاخص تعدیل شده نیکویی برازش (AGFI=۰/۹۷) و ریشه خطای میانگین مجذورات تقریب (RMSEA=۰/۰۳) حاکی از برازش خوب مدل بود.

باورهای ناامیدی: این متغیر از طریق "مقیاس باورهای ناکامی هرینگتون"<sup>۴</sup> (۲۰۰۵) اندازه‌گیری شده است. این مقیاس دارای ۴۴ گویه بوده و در طیف لیکرت از "همیشه" تا "هرگز" تهیه شده است. ضریب قابلیت اعتماد این پرسشنامه با روش الفای کرنباخ محاسبه و برابر با ۰/۷۵ بدست آمد. اعتبار سازه این پرسشنامه با انجام تحلیل عامل تأییدی مورد آزمون قرار گرفت. شاخص‌های نسبت مجذور کای به درجه آزادی (۲/۰۸)، نیکویی برازندگی (GFI=۰/۹۶)، شاخص

1. Solomon and Rothblum  
3. Self-efficacy

2. Self-regulated learning  
4. Harrington's Frustration Discomfort Scale (FDS)

تعدیل شده نیکویی برازش ( $AGFI=0/95$ ) و ریشه خطای میانگین مجذورات تقریب  $RMSEA=0/05$  حاکی از برازش خوب مدل بود.

باورهای غیر منطقی: باورهای غیر منطقی از طریق سیاهه باورهای غیرمنطقی<sup>۱</sup> کوپمنز و همکاران (۱۹۹۴) اندازه‌گیری شد و دارای ۵۸ گویه بوده و در طیف لیکرت از "همیشه" تا "هرگز" تهیه شده است. ضریب قابلیت اعتماد با روش الفای کرنباخ محاسبه و برابر با  $0/85$  بدست آمد. اعتبار سازه این پرسشنامه با انجام تحلیل عامل تأییدی مورد آزمون قرار گرفت. شاخص‌های نسبت مجذور کای به درجه آزادی ( $2/52$ )، نیکویی برازندگی ( $GFI=0/95$ )، شاخص تعدیل شده نیکویی برازش ( $AGFI=0/93$ ) و ریشه خطای میانگین مجذورات تقریب  $RMSEA=0/05$  حاکی از برازش خوب مدل بود.

### یافته‌های پژوهش

در ابتدا همبستگی پیرسون برای بررسی رابطه بین تعلل‌ورزی و دیگر متغیرها (خودتنظیمی، خودکارآمدی، باورهای ناامیدی و باورهای غیرمعقول) اجرا شد. بررسی همبستگی بین متغیرهای مرتبط با تعلل‌ورزی به منظور اطمینان از عدم هم خطی نیز اهمیت زیادی دارد. چرا که هم خطی تاثیر مهمی بر تحلیل خوشه‌ای با دادن وزن بیشتر به متغیرهای خطی دارد (هیر، بلک، بابین و آندرسون، ۲۰۱۰). سپس، از تحلیل خوشه‌ای برای شناسایی زیر گروه‌های شرکت‌کنندگان با توجه به متغیرهای مرتبط با تعلل‌ورزی استفاده شد. در ابتدا از روش سلسله مراتبی تراکمی با مجذور فاصله اقلیدسی برای تعیین و کشف تعداد خوشه‌ها استفاده شد. بعد از مشخص شدن تعداد خوشه‌ها با استفاده از تحلیل دندوگرام و ضرایب متراکم، روش خوشه‌بندی  $k$  میانگین که یک روش تکرار شونده است با توجه به حداکثر شباهت در درون خوشه‌ها و بیشتریت تفاوت در بین خوشه‌ها اجرا شد. در نهایت ثبات ساختار راه حل خوشه‌ها با تعیین توافق بین دو روش راه حل، کاربرد آزمون  $V$  کرامر مورد بررسی قرار گرفت. تحلیل واریانس یک راهه با آزمون تعقیبی بونفرونی به طور جداگانه برای خوشه‌ها نیز اجرا شد و تفاوت بین خوشه‌ها در متغیرهای مرتبط با تعلل‌ورزی بررسی شد.

غربال‌گری داده‌ها و تحلیل توصیفی: کجی و کشیدگی داده‌ها برای همه مقیاس‌ها مورد بررسی قرار گرفت مقدار کجی از  $-0.65$  تا  $0.86$  و مقدار کشیدگی از  $0.87$  تا  $0.97$  است از آن جایی که مقادیر بیشتر از ۳ برای کجی و بزرگتر از ۲۰ برای کشیدگی نشان دهنده غیر نرمال بودن توزیع هستند (وستون و گور، ۲۰۰۶) لذا نتایج این پژوهش حاکی از نرمال بودن توزیع متغیرها است ۴۰ تا از داده‌های بدون جواب (با توجه به تعداد نمونه که ۴۰۰ نفر بود ۱۰ درصد داده‌ها بدون جواب بودند) با میانگین متغیر مربوطه جایگزین شدند. میانگین و انحراف معیار متغیرها در جدول ۱ گزارش شده است.

تحلیل همبستگی: همبستگی بین تعلل‌ورزی و سایر متغیرهای پژوهش (خودتنظیمی، خودکارآمدی، باورهای ناامیدی و باورهای نامعقول) در جدول ۱ آورده شده است. تعلل‌ورزی به طور مثبت با باورهای ناامیدی و باورهای نامعقول رابطه دارد و رابطه آن با خودتنظیمی و خودکارآمدی منفی و معنی‌دار است. همبستگی پیرسون بین متغیرهای مرتبط با تعلل‌ورزی نشان‌دهنده فقدان همخطی بود (دامنه همبستگی  $0.1$  تا  $0.44$ ).

1. The Irrational Beliefs Inventory (IBI)

2 Weston & Gore

## جدول ۱. آمار توصیفی، آلفا کرونباخ و همبستگی پیرسون با تعلق‌ورزی

تعلق‌ورزی	آلفا کرونباخ	انحراف معیار	میانگین	
تعلق‌ورزی	۰.۸۷	2.84629	15.0375	
خودتنظیمی	۰.۹۴	2.61045	16.7600	***-۰.۴۱۵
خودکارآمدی	۰.۹۶	2.81677	15.9750	***-۰.۲۱۸
باور ناامیدی	۰.۷۶	3.52853	13.1750	***.۲۶۰
باور نامعقول	۰.۷۴	3.30954	13.0425	***.۲۴۳

از آنجایی که هدف اصلی این مقاله معرفی تحلیل خوشه‌ای است و امکان قرار دادن همه خروجی‌ها برای ۲۰۰ نمونه مورد مطالعه نیست در ابتدا برخی از مفاهیم اصلی این تحلیل را با ۲۰ نفر از دانشجویان نمونه و براساس دو متغیر خود تنظیمی و خود کارآمدی توضیح داده و سپس به اجرای تحلیل خوشه‌ای بر کل نمونه مورد مطالعه می‌پردازیم. در گروه‌بندی ۲۰ دانشجو بر اساس این دو متغیر در گام اول فاصله بین هر جفت از دانشجویان محاسبه می‌شود بدین منظور از مجذور فاصله اقلیدسی یا d2 استفاده می‌کنیم. لذا فاصله بین دانشجویان در خود تنظیمی و خود کارآمدی برای هر جفت از ۲۰ دانشجو در ماتریس تقریب ۱ آورده شده‌اند (جدول ۱). وقتی ماتریس تقریب داریم، می‌توانیم از روش خوشه‌ای برای تقسیم گروه‌ها به زیر گروه‌های همگن استفاده کنیم. در این جا از روش خوشه‌بندی واردی (۱۹۶۳) استفاده می‌کنیم. روش وارد ماتریس تقریب را به کار می‌برد. در این مثال فرد ۱۳ و ۱۸ با هم هیچ فاصله‌ای ندارند در واقع هر دو در این دو متغیر نمره یکسانی کسب کرده‌اند. فاصله دانشجو ۱۴ و ۱۹ نیز که ۰.۰۷۱ است و به هم نزدیک هستند در یک گروه گروه‌بندی می‌شوند دانشجو ۳، ۵ و ۱۵ با فاصله ۰.۱۴۵ در یک گروه قرار می‌گیرند و (شکل ۱). این روش ادامه می‌یابد تا گروه‌هایی با کمترین واریانس درون گروهی تشکیل شود. این فرایند تا زمانی که گروه‌ها شکل بگیرد ادامه می‌یابد (شکل ۲)

## جدول ۲. ماتریس تقریب (مجذور فاصله اقلیدسی) برای ۲۰ دانشجو در تحلیل خوشه‌ای

دانش آموز	1	2	3	4	5	6	7	8	9	10
2	9.545									
3	4.15	15.562								
4	0.947	4.781	4.428							
5	4.15	15.562	0	4.428						
6	2.649	2.48	9.448	0.947	9.448					
7	2.636	11.237	0.366	2.316	0.366	6.168				
8	2.636	7.999	1.249	1.433	1.249	4.402	0.294			
9	2.857	6.601	1.911	1.212	1.911	3.74	0.662	0.074		
10	0.294	6.601	5.327	0.358	5.327	1.177	3.225	2.636	2.562	
11	2.857	13.077	0.145	2.978	0.145	7.272	0.074	1.177	3.74	
12	1.779	9.189	0.947	1.303	0.947	4.428	0.145	0.366	0.366	2.074
13	0.145	8.497	5.733	0.935	5.733	1.911	3.782	3.488	3.561	0.145
14	0.714	7.359	2.316	0.366	2.316	2.48	0.935	0.641	0.714	0.714

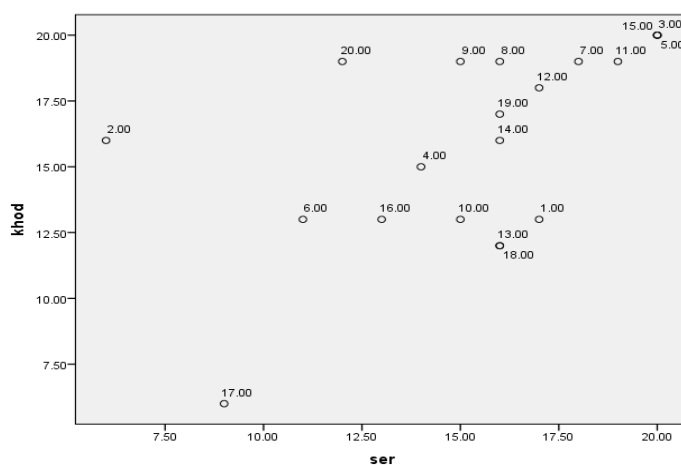
1 Proximity matrix

2 Ward

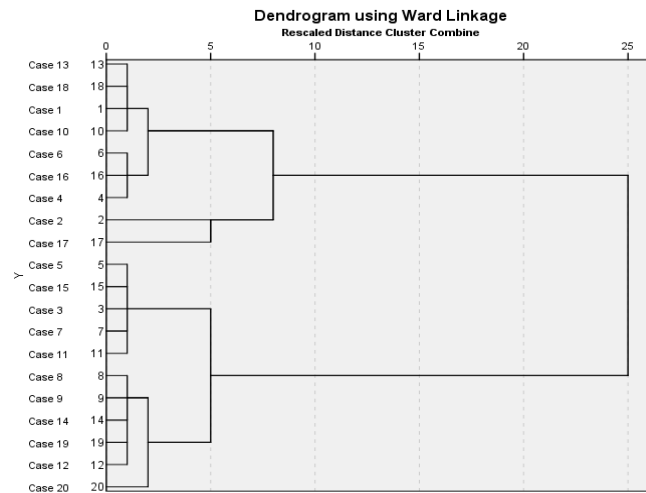
5.327	1.911	1.249	0.366	9.448	0	4.428	0	15.562	4.15	15
0.294	2.857	3.225	4.402	0.294	7.093	0.358	7.093	4.246	1.177	16
6.137	14.68	15.63	17.99	3.782	22.85	7.605	22.85	7.78	8.197	17
0.145	3.561	3.488	3.782	1.911	5.733	0.935	5.733	8.497	0.145	18
1.212	0.358	0.285	0.579	2.978	1.818	0.579	1.818	7.43	1.212	19
3.225	0.662	1.177	2.649	2.636	4.781	1.433	4.781	3.29	4.402	20

ادامه جدول ۲. ماتریس تقریب (مجذور فاصله اقلیدسی) برای ۲۰ دانشجو در تحلیل خوشه‌ای

دانش آموز	1	2	3	4	5	6	7	8	9
2	9.545								
3	4.15	15.562							
4	0.947	4.781	4.428						
5	4.15	15.562	0	4.428					
6	2.649	2.48	9.448	0.947	9.448				
7	2.636	11.237	0.366	2.316	0.366	6.168			
8	2.636	7.999	1.249	1.433	1.249	4.402	0.294		
9	2.857	6.601	1.911	1.212	1.911	3.74	0.662	0.074	
10	0.294	6.601	5.327	0.358	5.327	1.177	3.225	2.636	2.562
11	2.857	13.077	0.145	2.978	0.145	7.272	0.074	0.662	1.177
12	1.779	9.189	0.947	1.303	0.947	4.428	0.145	0.145	0.366
13	0.145	8.497	5.733	0.935	5.733	1.911	3.782	3.488	3.561
14	0.714	7.359	2.316	0.366	2.316	2.48	0.935	0.641	0.714
15	4.15	15.562	0	4.428	0	9.448	0.366	1.249	1.911
16	1.177	4.246	7.093	0.358	7.093	0.294	4.402	3.225	2.857
17	8.197	7.78	22.85	7.605	22.85	3.782	17.99	15.63	14.68
18	0.145	8.497	5.733	0.935	5.733	1.911	3.782	3.488	3.561
19	1.212	7.43	1.818	0.579	1.818	2.978	0.579	0.285	0.358
20	4.402	3.29	4.781	1.433	4.781	2.636	2.649	1.177	0.662



شکل ۱. نمودار ۲۰ دانشجو در خودتنظیمی و خودکارآمدی



شکل ۲. نمودار دندوگرام در خوشه بندی ۲۰ دانشجو با روش وارد

از آن جایی که این مثال بسیار ساده است، و مبتنی بر گروه‌هایی با تنها ۲ متغیر است در شکل ۱ می‌توان فقط فضای دو بعدی را بررسی کرد.

در ادامه برای اجرای تحلیل خوشه‌ای بر کل نمونه، متغیرهای مرتبط با تعلل‌ورزی (خودتنظیمی، خودکارآمدی، باورهای ناامیدی و باورهای نامعقول) انتخاب شدند. بررسی عینی از ضرایب دندوگرام و متراکم به دست آمده با روش وارد نشان دهنده راه حل ۲ خوشه‌ای بود. سپس تحلیل خوشه‌ای با  $k$  میانگین با کاربرد روش تکرار شونده و مشخص کردن ۲ خوشه اجرا شد. راه حل ۲ خوشه‌ای با توافق خوب بین روش وارد و خوشه بندی  $k$  میانگین حمایت شد (کرامر  $p < .07$ ). تحلیل عملکرد تفکیک نشان داد که دو خوشه به اندازه کافی در تمایز فضای عملکرد از هم جدا شده اند و ۹۷.۸ درصد موارد به طور صحیح طبقه‌بندی شده است. دو خوشه شناسایی شده منعکس کننده دو خوشه از تعلل‌ورزان به صورت زیر است:

خودتنظیمی و خودکارآمدی متوسط و باورهای ناامیدی و نامعقول بالا (خوشه ۱، تعلل‌ورزان با باورهای نادرست)

خودتنظیمی و خودکارآمدی پایین و باورهای ناامیدی و نامعقول پایین (خوشه ۲، تعلل‌ورزان با انگیزش پایین)

جدول ۳. نتایج آمار توصیفی خوشه‌ها و آنوا

متغیر	خوشه ۱: فراوانی: ۱۱۶	خوشه ۲: فراوانی: ۸۴.۶	F	P
	درصد: ۵۷.۸	درصد: ۴۲.۳		
	میانگین	انحراف معیار	میانگین	انحراف معیار
تعلل‌ورزی	16.0996	2.23676	13.5858	2.95301
خودتنظیمی	19.7662	2.01664	13.3846	2.70581
خودکارآمدی	18.8485	2.31976	12.7811	2.99891
باور ناامیدی	15.4069	2.07023	10.1243	2.73686
باور نامعقول	14.9221	2.32622	10.4734	2.66607

با توجه به تفاوت بین خوشه‌ها در تعلل‌ورزی نتایج نشان داد که دو خوشه به طور معنی‌داری از هم متفاوت هستند و میانگین نمرات تعلل‌ورزی به طور معنی‌داری برای گروه تعلل‌ورز با باور غیر



منطقی و ناامیدی بالا (خوشه ۱) بیشتر از گروه تعلل‌ورز با باور غیر منطقی و ناامیدی پایین (خوشه ۲) است. آزمون تحلیل واریانس به منظور بررسی اعتبار خوشه‌ها انجام شده است.

### بحث و نتیجه‌گیری

امروزه در اطراف ما داده‌های فراوانی وجود دارند که به علت عدم شناخت از نحوه استفاده از آنها، ارزش چندانی برای روانشناسان ندارند و مورد غفلت قرار می‌گیرند. در حالی که اگر همین داده‌های به ظاهر کم ارزش بصورت هدفمند حفظ، ذخیره‌سازی و در نهایت مورد داده‌کاوی قرار بگیرند، دانش فراوانی تولید کرده و در تصمیم‌گیری‌های روانشناسی ما را یاری خواهند کرد. لذا می‌توان از تکنیک‌های داده‌کاوی مانند خوشه‌بندی استفاده نموده و الگوهای پنهان را از این داده‌های انبوه استخراج نمود تا بتوان تصمیمات بهینه و به هنگام در مسایل روانشناسی اتخاذ نمود. در این مقاله به اختصار تحلیل خوشه‌بندی به عنوان یکی از تکنیک‌های داده‌کاوی مورد بررسی قرار گرفت. تحلیل خوشه‌ای اطلاعات و دانشی را در اختیار ما قرار می‌دهد که افراد برای تصمیم‌گیری هوشمندانه درباره مشکلات و مسایل پیچیده و چند لایه پیش روی خود به آنها نیاز دارند. تحلیل خوشه‌ای راه حلی اساسی برای عصر انفجار داده‌هاست عصری که از سویی در داده‌ها غرق هستیم و از سوی دیگر تشنه دانشی سودمند و کاربردی هستیم. در این مقاله ضمن تشریح منطقی و ضرورت توسعه و کاربرد مضاعف روش‌های تحلیل خوشه‌ای به عنوان یکی از تکنیک‌های رایج داده‌کاوی و کاربرد آن در روانشناسی به شناسایی زیر گروه‌های از تعلل‌ورزان براساس عوامل موثر بر آن و مبتنی بر مدل رویکرد عقلانی، عاطفی، رفتاری الیس با کاربرد تحلیل خوشه‌ای پرداختیم نتایج استفاده از تحلیل خوشه‌ای در بین دانشجویان تعلل‌ورز نشان دهنده دو خوشه یا دو زیر گروه بوده است. اولین زیر گروه تعلل‌ورزی را به عنوان یک شکست در خود تنظیمی و خودکارآمدی پایین در نظر می‌گیرد. ۵۷.۸ درصد نمونه مورد مطالعه در این زیر گروه قرار گرفتند. و این گروه دارای باور نادرست نامیده شدند. دومین زیر گروه باورها را مرکز تعلل‌ورزی در نظر می‌گیرد. ۴۲.۳ درصد نمونه مورد مطالعه در این زیر گروه قرار گرفتند. و این گروه تعلل‌ورزان با انگیزش پایین نامیده شدند. نتایج این پژوهش بیانگر وجود خوشه‌های مختلف از تعلل‌ورزان است که با شناسایی گروه‌های تعلل‌ورزان و ویژگی‌های هر یک از گروه‌ها می‌توان راهکارهای درست‌تر و بهتری برای هر گروه داشت. این مطالعه به متغیرهای مدل الیس محدود است و بررسی متغیرهای دیگر در خوشه‌بندی افراد تعلل‌ورز به منظور شناسایی دقیق زیر گروه‌ها بسیار ضروری است. لذا روانشناسان و پژوهشگران باید با آشنایی هرچه بیشتر این تکنیک از آن در تحلیل‌های خود استفاده کنند. و بعد از آشنایی با آن و با داشتن دقت و حساسیت در انتخاب روش، انتخاب تعداد خوشه‌ها و نحوه اجرای تحلیل از آن استفاده کنند.

### References

- Abel, G., Plumridge, L., & Graham, P. (2002). Peers, networks or relationships: Strategies for understanding social dynamics as determinants of smoking behaviour. *Drugs-Education Prevention and Policy*, 9, 325-338.
- Aldenderfer, M. S., & Blashfield, R. K. (1984). *Cluster analysis*. Newbury Park, CA: Sage Publishing.
- Anderberg, M. (1973). *Cluster Analysis for Application*. Academic Press.

- Anderson, E. M. (2001). The relationships among task characteristics, selfregulationand procrastination.Unpublished Doctoral Dissertation,Loyola University, Chicago
- Blashfield, R. K. (1980). The growth of cluster analysis: Tryon, Ward and Johnson. *Multivariate Behavioral Research*, 15, 439-458.
- Borgen, F. H., & Weiss, D.J. (1971). Cluster analysis and counseling research. *Journal of Counseling Psychology*, 18, 583-591.
- Brown, C. J. (1983). The joys of Fs: A student guide for getting lousy grades.*Teaching of Psychology*, 10(3), 176-177.
- Bunge J.A., D.H.Judson (2005). Data Mining, In Bunge, J.A., Judson, D. H. (eds.) *Encyclopedia of Social Measurement*, Elsevier Academic Press, :617-624.
- Ellis, A., & Dryden, W. (1987).*The practice of rational emotive therapy*. NY:Springer.
- Everitt, B. (1980). *Cluster analysis* (2nd ed.). New York: Halsted.
- Everitt, B. S. (1993). *Cluster analysis* (3rd éd.). London: Arnold.
- Ferrari, J. R., & Tice, D. M. (2000). Procrastination as a self-handicap for menand women: A task-avoidance strategy in a laboratory setting. *Journal of Research in Personality*, 34(1), 73-83.
- Froehlich, R. A. (1987). A desriptive study of general procrastination in a grouporiented treatment setting.Unpublished Doctoral Dissertation, United States International University, Kenya.
- Froggat, W. (2005).A brief introduction to racional emotive behaviour therapyRetrieved October, 09, 2010, fromHarrington, N. (2005b). It's too difficult! Frustration intolerance beliefs andprocrastination.*Cognitive Therapy Research*, 30, 699–709
- Gangestad, S., & Snyder, M.(1985).”To carve nature at its joints”:On the existence of discrete classes in personality. *Psychological Review*, 92, 317-349.
- Hair, J.F., R.E. Anerson, R.L. Tatham and W.C. Black (1992). *Multivariate Data A nalysis* (3rd ed) Macmillan. New York.
- Harrington, N. (2006). Frustration intolerance beliefs: Their relationship withdepression, anxiety, and anger, in a clinical population. *Cognitive Therapy& Research*, 30, 699–709
- Jardine, N., and Sibson, R. (1971), *Mathematical Taxonomy*, 131-153.
- Jolliffe, I. T., Jones, B., & Morgan, B. J. T. (1982). Utilizing clusters: A case-study involving the elderly. *Journal of the Royal Statistical Society*, 145, 224-236.
- Kerlinger, F.N. (1986). *Foundations of Behavioral Research*, Holt, Rinchart & Winston, Fort Worth, TX.
- Ketchen, D. J., & Huit, T. M. (2000). Validating cluster assignments. *Psychological Reports*, 87, 1057-1058.

- Kiss, F. (2003). Credit scoring process from a knowledge management prospective. *Periodica Polytechnica Ser.SOC. MAN. SCI*, 11 (1): 95-110.
- Kos, A. J., & Psenicka, C. (2000). Measuring cluster similarity across methods. *Psychological Reports*, 86, 858-862.
- Lorr, M. (1983). *Cluster analysis for the social science*. San Francisco: Jossey-Bass.
- Martin, C, Wyllie, A., & Casswell, S. (1992). Types of New Zealand drinkers and their associated alcohol-related problems. *Journal of Drug Issues*, 22, 773-796.
- Maschewsky-Schneider, U., & Greiser, E. (1989). Primary prevention of coronary heart disease versus health promotion - a contradiction? *Annals of Medicine*, 21, 215-218.
- Milligan, G. W (1980). An examination of the effect of six types of error perturbation on fifteen clustering algorithms. *Psychometrika*, 45, 325-342.
- Nicholson, L., & Scharff, L. F. V. (2007). The effects of procrastination and self-awareness on emotional responses. *Psi Chi Journal of Undergraduate Research*, 12(4), 139-145.
- Punj, G., & Stewart, D. (1983). Cluster analysis in marketing research. Review and suggestions for application. *Journal of Marketing Research*, 20, 134-148.
- Pychyl, T. A., & Flett, G. L. (2012). Procrastination and self-regulatory failure: An introduction to the special issue. *Journal of Rational-Emotive and Cognitive Behavior Therapy*, 30, 203-212
- Romesburg, H. C. (1984). *Cluster analysis for researcher*. Belmont, CA: Lifetime Learning Publication.
- Rothblum, E. D., Solomon, L. J., & Murakami, J. (1986). Affective, cognitive and behavioral differences between high and low procrastinators. *Journal of Counseling Psychology*, 33(4), 387-394.
- Skinner, H, A., & Blashfield, R. K. (1982). Increasing the impact of cluster analysis research: The case of psychiatric classification. *Journal of Consulting and Clinical Psychology*, 50, 727-735.
- Steel, P. (2007). The nature of procrastination: A meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychological Bulletin*, 133, 65-94
- Trybula WJ, (1997). Data mining and knowledge discovery. *Annual Review of Information Science and Technology*, 32:197-229.
- Ward, J. H. J. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58, 236-244.
- Wilson, B. A., & Nguyen, T. D. (2012). Belonging to tomorrow: An overview of procrastination. *International Journal of Psychological Studies*, 4, 211-217
- Wolters, C. A. (2003). Understanding procrastination from a self-regulated learning perspective. *Journal of Educational Psychology*, 95(1), 179-187

- Wyllie, A., & Casswell, S. (1997). Gender focus of target groups for alcohol health promotion strategies in New Zealand. *Health Promotion International*, 12, 141-149.
- Andreou, C. (2007). Understanding procrastination. *Journal for the Theory of Social Behaviour*, 37(2), 183-193.
- Beswick, G., Rothblum, E. D., & Mann, L. (1988). Psychological antecedents of student procrastination. *Australian Psychologist*, 23(2), 207-217.
- Brown, C. J. (1983). The joys of Fs: A student guide for getting lousy grades. *Teaching of Psychology*, 10(3), 176-177.
- Dryden, W. (2003). *Rational emotive behaviour therapy: Theoretical developments*. NY: runner-Routledge.
- Ellis, A. (1973). *Humanistic psychotherapy: The rational-emotive approach*. NY: The Julian Press, Inc.
- Ellis, A., & Dryden, W. (1987). *The practice of rational emotive therapy*. NY: Springer.
- Ellis, A., & Dryden, W. (1997). *The practice of rational emotive behavior therapy*. NY: Springer Publishing Company.
- Ellis, A., & Knaus, W. (1977). *Overcoming procrastination*. NY: Signet.
- Froehlich, R. A. (1987). *A descriptive study of general procrastination in a group-oriented treatment setting*. Unpublished Doctoral Dissertation, United States International University, Kenya.
- Gustavson, D. E., Miyake, A., Hewitt, J. K., & Friedman, N. P. (2014). Genetic relations among procrastination, impulsivity, and goal-management ability: Implications for the evolutionary origin of procrastination. *Psychological Science*, 25, 1178–1188.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (۲۰۱۰). *Multivariate data analysis: A global perspective*. London, England: Pearson Education
- Klassen, R. M., Krawchuk, L. L., & Rajani, S. (2008). Academic procrastination of undergraduates: Low self-efficacy to self-regulate predicts higher level of procrastination. *Contemporary Educational Psychology*, 33(4), 915-931.
- Schouwenburg, H., C. (2004). Procrastinators and fear of failure: An exploration of reasons for procrastination. *European Journal of Personality*, 6, 225-236.
- Van Eerde, W. (2003). Procrastination at work and time management training. *The Journal of Psychology*, 137(5), 421-434.