مقاله پژوهشی

اندازه‌گیری تربیتی
دوره ۱۳، شماره ۵۰، زمستان ۱۴۰۱، ۱۳۳-۱۵۰
jem.atu.ac.ir
DOI: 10.22054/jem.2023.68624.3384

برآورد تعداد خوشه‌ها در تکنیک طبقه‌بندی ماتریس‌های نامنفی

ریحانه رحیمی | دانشجوی دکتری رشته سنجش و اندازه‌گیری، دانشگاه علامه طباطبائی، تهران، ایران

جلیل یونسی * | دانشیار گروه سنجش و اندازه‌گیری، دانشگاه علامه طباطبائی، تهران، ایران

علی مقدم‌زاده | دانشیار گروه سنجش و اندازه‌گیری، دانشگاه تهران، تهران، ایران

محمد عسگری | دانشیار گروه سنجش و اندازه‌گیری، دانشگاه علامه طباطبائی، تهران، ایران

تاریخ دریافت: ۱۴۰۱/۰۳/۳۰؛ تاریخ پذیرش: ۱۴۰۱/۰۸/۱۶

ISSN: 2252-004x
eISSN: 2470-6240

چکیده

استفاده از روش‌ها و تکنیک‌های مختلف در جهت رسیدن به بهترین نتایج و ایجاد هم‌افزایی بین آن‌ها می‌تواند در بسیاری از مسائل راه‌گشا و کمک کننده باشد. داده‌کاوی آموزشی یکی از حوزه‌های نسبتاجدیدی است که می‌توان به کمک آن در حل مسائل آموزشی و بویژه مسائل حوزه سنجش و اندازه‌گیری اقدام کرد. اما قبل از استفاده از این روش‌ها باید تا حد امکان با ان آشنا شده و مشکلات و معایب و مزایای آن مورد بررسی قرار گیرد. در این پژوهش نیز هدف بررسی تکنیک طبقه‌بندی ماتریس‌ها نامنفی و چگونگی تعیین تعداد خوشه‌ها قبل از اجرای مدل است. روش اجرای پژوهش از نوع توصیفی و جامعه مورد مطالعه تمامی حاضرین جلسه کنکور رشته علوم ریاضی و فنی در سال ۱۳۹۸ بود که از این تعداد ۵۰۰۰ نفر به صورت تصادفی توسط سازمان سنجش و آموزش کشور انتخاب و در اختیار پژوهشگر قرار گرفت. ابزار پژوهش سؤالات حسابان و هندسه کنکور بود. نتایج این تحلیل نشان داد که در برآورد تعداد خوشه‌های سؤالات حسابان اختلاف وجود دارد ولی در مورد سؤالات هندسه نتایج تمامی روش‌ها یکسان بود. با توجه به اختلاف مشاهده شده پیشنهاد می‌شود که در پژوهش‌های آتی با کمک شبیه‌سازی داده به بررسی دقیق‌تر این مسأله پرداخته شود.

کلیدواژه‌ها: حسابان، داده‌کاوی آموزشی، طبقه‌بندی ماتریس‌ها نامنفی، هندسه

مقاله حاضر برگرفته از رساله دکتری رشته سنجش و اندازه‌گیری دانشگاه علامه طباطبائی است.

* نویسنده مسئول: jalilyounesi@gmail.com

مقدمه

داده‌کاوی[1] حوزه نسبتاً جدیدی است که هدف اصلی آن بیرون کشیدن اطلاعات از پایگاه داده‌های بسیار بزرگ است (Larose & Larose, 2014). داده‌کاوی و کشف دانش نامی است که برای اشاره به یک حوزه بین رشته‌ای شامل استفاده از روش‌های چندین حوزه پژوهشی (مستدل، عمدتاً یادگیری ماشین و آمار) برای بیرون کشیدن دانش از مجموعه داده‌های دنیای واقعی استفاده می‌شود (Freitas, 2002). در حقیقت داده‌کاوی با رویکرد استدلال استقرایی شواهدی را برای نتیجه‌گیری، پیش‌بینی و کشف دانش فراهم می‌کند. داده‌کاوی یک رویکرد تحلیل داده اکتشافی است که با فرضیه‌های از قبل تعیین‌شده هدایت نمی‌شود و به همین خاطر با تحلیل‌های آماری سنتی تفاوت برجسته‌ای دارد (Schumann, 2005). با توسعه سریع فناوری‌های اطلاعاتی، تعداد زیادی از تکنیک‌های داده‌کاوی در بسیاری از رشته‌ها از جمله علوم اجتماعی، فیزیک، مهندسی و پزشکی به کار گرفته شده است. مطالعاتی که از روش‌های داده‌کاوی در زمینه آموزش استفاده می‌کنند، عموماً به عنوان مطالعات داده‌کاوی آموزشی[2] (EDM) شناخته می‌شوند. از آنجایی که داده‌کاوی استخراج دانش از حجم وسیعی از داده است وحجم عظیمی از داده‌ها در پایگاه داده آموزشی وجود دارد، بنابراین تکنیک‌های مختلف داده‌کاوی به منظور بدست آوردن داده‌های مورد نیاز و یافتن رابطه‌ای که دیده نمی‌شود، استفاده می‌شوند.

داده‌کاوی آموزشی یک حوزه تحقیقاتی جدید و درحال رشد است که می‌تواند به‌عنوان کاربرد تکنیک‌های داده‌کاوی روی داده‌های خام سیستم آموزشی به‌منظور پاسخ به سؤالات و مشکلات آموزشی و کشف اطلاعات پنهان شده در پس این داده‌ها تعریف شود. در چندسال گذشته محبوبیت این رشته منجر به افزایش تعداد مطالعات پژوهشی در این حوزه شده است به‌طوری که احاطه و شناسایی سهم تکنیک‌های داده‌کاوی در سیستم آموزشی را دشوار کرده است. در حقیقت استفاده و درک داده‌های خام جمع‌آوری شده از سیستم آموزشی می‌تواند درحکم یک معدن طلا باشد که به طراحان و کاربران این سیستم در بهبود عملکرد و استخراج اطلاعات مفید از رفتار دانش‌آموزان در فرآیند یادگیری کمک می‌کند (PeñaAyala, 2013). داده‌کاوی آموزشی یک حوزه بین رشته‌ای است که شامل بازیابی

1. data mining
2. data mining education

اطلاعات[1]، سیستم‌های توصیه‌کننده[2]، تحلیل داده‌های بصری[3]، تحلیل شبکه‌های اجتماعی[4] (SNA)، روانشناسی تربیتی[5]، روانشناسی شناختی، روانسنجی و غیره می‌شود اما به این‌ها محدود نمی‌شود.

با گسترش کاربرد داده‌کاوی در زمینه‌های آموزشی به‌تدریج نه‌تنها این رشته به حوزه سنجش ورود کرد بلکه به دلیل مشکلاتی نظیر فقدان استقلال، جمع‌پذیری خطی و واریانس ثابت در روابط بین عوامل روانسنجی و عملکرد تحصیلی استفاده از این تکنیک‌ها رو به افزایش گذاشت (Carthy et al., 2014). در مجموع یک همپوشانی آشکاری بین مدل‌های داده‌کاوی آموزشی و مدل‌های روانسنجی وجود دارد از آن جمله می‌توان به مواردی همچون تحلیل عاملی، تحلیل طبقه مکنون، تحلیل خوشه‌بندی و شبکه‌های بیزی اشاره کرد که باعث چالش‌هایی نیز در این زمینه شده است (Mislevy et al., 2012). به همین جهت Carthy و همکاران (2014) در یک مطالعه مروری که با هدف کاربرد روش‌های مختلف تحلیل داده در مدلسازی پیشرفت‌های تحصیلی انجام شد؛ ضمن تأکید بر افزایش اهمیت تکنیک‌های داده‌کاوی در حوزه روانسجی با مطرح کردن این سؤال که " آیا روش مدل‌سازی تجربی داده‌کاوی می‌تواند ارزش‌افزوده‌ای در این حوزه ایجاد کند؟ " به‌صورت تلویحی نگه‌داشتن جانب احتیاط در کاربرد این تکنیک‌ها را گوشزد کرده‌اند. در راستای چنین مشکلات و ابهاماتی Reimann و همکاران (2011) بر اهمیت مقایسه نتایج تکنیک‌های مختلف تأکید کرده و آن را موجب دستیابی به بینشی عمیق‌تر از ساختار داده‌های چند متغیری دانسته‌اند. از همین روی Mislevy و همکاران (2012) در پژوهشی سعی کردند با بحث درباره ریشه‌های هستی‌شناسی و معرفت‌شناسی روانسنجی و داده‌کاوی تا حد امکان این ابهامات را رفع کرده و جایگاه هر یک از این دو رویکرد را در حوزه سنجش مشخص کنند. علاوه بر این مباحث نظری پژوهشی‌هایی در جهت مقایسه این روش‌ها نیز تاکنون انجام شده است. به‌عنوان‌مثال Man و همکاران (2019) به مقایسه توفیق مدل‌های نظریه سؤال پاسخ، زمان پاسخ و داده‌کاوی برای کشف افراد متقلب در آزمون پرداختند. در رشته‌های دیگر نیز چنین مقایسه‌هایی مرسوم بوده و استفاده می‌شود؛ برای مثال

1. information retrieval
2. recommender systems
3. visual data analytics
4. social network analysis
5. psychopedagogy

در پژوهشی در حوزه زمین‌شناسی به مقایسه تحلیل رگرسیون با تکنیک‌های داده‌کاوی (Nguyen et al., 2017) پرداخته شده است. خوشه‌بندی یکی از تکنیک‌های پرکاربرد در داده‌کاوی است که می‌تواند با مدل‌های روانسنجی همپوشانی داشته باشد. برای مثال کشف روابط بین داده‌های خام سطح پایین‌تر که یک کار بنیادی برای مدلسازی‌های روانسنجی است به کمک تکنیک‌های داده‌کاوی آموزشی نظیر خوشه‌بندی[1] امکان‌پذیر است (Mislevy et al., 2012).

به‌طورکلی تکنیک‌های داده‌کاوی بر اساس اهداف الگوریتم به دو دسته تحلیل همبستگی[2] و تحلیل طبقه‌بندی[3] (شامل طبقه‌بندی و خوشه‌بندی) تقسیم می‌شود (Yu, 2009) که برخی از این الگوریتم‌ها در داده‌کاوی آموزشی نیز کاربرد دارد. یکی از الگوریتم‌های مورداستفاده تکنیک‌های خوشه‌بندی است که برای شناسایی زیرگروه‌های متجانس در یک مجموعه داده به کار می‌رود (Baepler & Murdoch, 2010). خوشه‌بندی یک روش طبقه‌بندی بدون نظارت[4] است که به‌عنوان ابزار اکتشافی تحلیل داده برای روشن کردن رفتار چندمتغیری یک مجموعه داده به کار می‌رود (Reimann et al., 2011). از دیدگاه تاریخی خوشه‌بندی در ریاضیات، آمار و تحلیل‌های عددی ریشه دارد و از دیدگاه یادگیری ماشین[5] خوشه‌بندی با الگوهای پنهان مرتبط بوده و نتایج آن نشان‌دهنده مفهوم داده است (Berkhin, 2006). خوشه‌بندی با تحلیل عاملی و تحلیل مؤلفه‌های اصلی از منظر کاهش تعداد متغیرها شباهت دارد اما آنچه باعث تمایز بین این تکنیک‌های آماری می‌شود راه رسیدن به هدف نهایی است. درحالی‌که تحلیل عاملی و تحلیل مؤلفه‌های اصلی از ماتریس همبستگی بهره می‌برند، خوشه‌بندی مبتنی بر اندازه فاصله است (Reimann et al., 2011).

خوشه‌بندی در داده‌کاوی بیشتر در زمینه خوشه‌بندی افراد شناخته شده است ولی تکنیک‌هایی نیز در جهت خوشه‌بندی متغیرها ارائه شده است که هدف آن‌ها شناسایی فرایندها در الگوهای پاسخ از طریق الگوریتم‌های خوشه‌بندی است (Kerr & Chung, 2012). اینجا درست یکی از نقاط اتصال بین روانسنجی و داده‌کاوی است. در واقع یکی از

1. clustering
2. association
3. classification
4. Un supervised
5. machine learning

چالش‌های اصلی در نظریه‌های روانسنجی از جمله نظریه کلاسیک آزمون[1] و نظریه سؤال-پاسخ[2] یافتن ویژگی‌های پنهانی است که در ورای پاسخ مشاهده شده افراد وجود دارد و همان‌طور که در تعریف داده‌کاوی مطرح شد، وظیفه داده‌کاوی استخراج اطلاعات از حجم زیاد داده است. بنابراین یکی از اطلاعاتی که با کمک داده‌کاوی می‌توان کشف و استخراج کرد همین ویژگی‌های پنهان افراد است. Berkhin (2006) اعتقاد دارد که داده‌کاوی توانایی تسهیل شناسایی فرایندهای شناختی و رفتاری را دارد و می‌توان با تکنیک خوشه‌بندی سؤال‌ها از طریق داده‌کاوی به‌عنوان گزینه‌ای برای تعیین ابعاد شناختی لازم برای پاسخ به سؤال‌های آزمون نگاه کرد. یکی از تکنیک‌های داده‌کاوی که در جهت خوشه‌بندی متغیرها استفاده می‌شود، طبقه‌بندی ماتریس‌های نامنفی[3] است.

طبقه‌بندی ماتریس‌های نامنفی، اساس چندین تکنیک رایج برای تشخیص الگو و خوشه‌بندی بدون نظارت است. مشابه تجزیه و تحلیل مؤلفه‌های اصلی (PCA) یا تجزیه و تحلیل مؤلفه‌های مستقل (ICA)، هدف از طبقه‌بندی ماتریس نامنفی (NMF) توضیح داده‌های مشاهده شده با استفاده از تعداد محدودی مؤلفه پایه است که وقتی با هم ترکیب شوند، داده‌های اصلی را تا حد امکان به دقت تقریب می‌زنند. ویژگی‌های متمایزکننده تکنیک طبقه‌بندی ماتریس نامنفی این است که هم ماتریس که مؤلفه‌های پایه را نشان می دهد و هم ماتریس ضرایب مخلوط به ورودی‌های غیرمنفی محدود می‌شوند و هیچ محدودیت متعامد یا استقلال بر مؤلفه‌های پایه اعمال نمی‌شود. این ویژگی منجر به تفسیر ساده و شهودی عوامل در NMF می‌شود و به اجزای پایه اجازه می‌دهد با هم همپوشانی داشته باشند.

با اینکه می‌توان به کمک تکنیک طبقه‌بندی ماتریس‌های نامنفی در حوزه سنجش کمک گرفت ولی ابهامات و مشکلاتی نیز در کاربرد این مدل وجود دارد. یکی از مشکلاتی که در کاربرد این مدل وجود دارد این است که باید تعداد خوشه‌ها که از آن با عنوان برآورد رتبه[4] یاد می‌شود قبل از اجرای مدل مشخص شود. در واقع هیچ روشی به‌طور کامل و صد در صد با نظارت یا بدون نظارت نیست و نیاز است که مواردی قبل از اجرای تحلیل مشخص شوند. بنابراین با اینکه از خوشه‌بندی با عنوان روش بدون نظارت یاد می‌شود نیاز است که

1. Classic Test Theory (CTT)
2. Item Response Theory (IRT)
3. nonnegative matrix factorizations
4. estimate rank

پژوهشگر برخی از موارد مثل الگوریتم یا تعداد خوشه‌ها را قبل از اجرای مدل مشخص کند. البته روش‌هایی برای تعیین تعداد خوشه‌ها معرفی شده است ولی این روش‌ها تاحدی شهودی بوده و بستگی به قضاوت پژوهشگر دارد. همچنین گاها مشاهده می‌شود که بین نتایج روش‌های مختلف تناقض وجود دارد. لذا در این پژوهش بررسی دقیق‌تر این روش‌ها و مقایسه آن‌ها با روش‌های مورد استفاده در تحلیل مؤلفه‌های اصلی که یکی دیگر از تکنیک‌های داده‌کاوی است، مد نظر قرار گرفته است. بنابراین مسأله اصلی یافتن روش دقیقی در جهت تعیین تعداد خوشه‌ها است.

روش

روش پژوهش حاضر توصیفی است؛ زیرا هدف اکتشاف تعداد خوشه‌ها در تکنیک طبقه‌بندی ماتریس‌های نامنفی است. در واقع هر پژوهشی که هدف آن توصیف یک پدیده یا موقعیت با هدف شناخت بهتر است در زمره پژوهش‌های توصیفی قرار می‌گیرد (سرمد و همکاران، ۱۳۹۱).

جامعه آماری شامل کلیه داوطلبان حاضر در جلسه آزمون سراسری کارشناسی سال ۱۳۹۸ در رشته علوم ریاضی و فنی است. در مورد حجم نمونه براساس ماهیت پژوهش که از نوع تحلیل داده‌های بزرگ مقیاس است به حجم بالای دیتا نیاز است و هرچه این حجم بیشتر باشد، مسیر رسیدن به نتایج مطلوب هموارتر خواهد شد. اما در این مورد پژوهشگر به تنهایی تصمیم‌گیرنده نبوده و رضایت و موافقت سازمان سنجش و آموزش کشور به عنوان مجری و متولی کنکور، نیز شرط بود. لذا بعد از رایزنی‌ها و ارائه درخواست کتبی، سازمان سنجش و آموزش کشور با ارائه نمونه‌ای ۵۰۰۰ نفری به پژوهشگر موافقت کرد. لازم به ذکر است که همه ۵۰۰۰ داده دریافتی قابل تحلیل نبودند؛ زیرا در صورت وجود حتی یک مورد داوطلبی که به همه سؤالات پاسخ نداده یا پاسخ اشتباه داده است، اجرای مدل داده‌کاوی مورد استفاده در این پژوهش را غیرممکن ساخته و دستورات اجرا نمی‌شود. لذا بعد از سرند داده‌ها حجم نمونه برای تحلیل سؤالات حسابان برابر با ۲۹۶۹ نفر و برای سؤالات هندسه برابر با ۲۷۱۷ نفر است. با توجه به محدودیت‌های سازمان سنجش و آموزش کشور و محرمانه بودن اطلاعات، محقق در انتخاب آزمودنی نقشی نداشت و فرآیند انتخاب نمونه توسط سازمان مربوطه به روش تصادفی انجام گرفت. بنابراین روش نمونه‌گیری به صورت تصادفی و توسط سیستم انجام شد.

ابزار پژوهش سؤالات حسابان و هندسه کنکور رشته علوم ریاضی و فنی در سال ۱۳۹۸ است. ابتدا سرالات دفترچه مطالعه شد و با مشورت سه متخصص حوزه آموزش ریاضی که تجربه تدوین کتب درسی را نیز داشتند، سؤالات مربوط به هر حوزه محتوایی مشخص گردید. بر این اساس ۱۹ سؤال در حوزه حسابان و ۱۸ سؤال در حوزه هندسه قرار گرفتند. تمامی مراحل کار در محیط نرم‌افزاری RStudio انجام شد. پکیج " NMF"، پکیج اصلی مورد استفاده در این تحلیل بود، منتها برای کار با این پکیج به پکیج‌های دیگری از جمله پکیج "pkgmaker، registry، cluster و rngtools" نیاز بود که همه آن‌ها نصب و فراخوانی شدند. همچنین از پکیج "ClustOfVar " برای تحلیل‌های مربوط به روش تحلیل مؤلفه‌های اصلی استفاده شد.

یافته‌ها

در این بخش نتایج حاصل از روش‌های تعیین تعداد خوشه‌ها با استفاده از روش‌های پیشنهادی در تکنیک طبقه‌بندی ماتریس‌های نامنفی (NMF) که یک تکنیک داده‌کاوی است، گزارش شده است . همچنین برای بررسی روایی نتایج حاصل از این تکنیک تعداد خوشه‌ها از طریق روش‌های تحلیل مؤلفه‌های اصلی نیز برآورد شده و در اینجا گزارش شده است.

معادله اصلی تکنیک طبقه‌بندی ماتریس‌های نامنفی از تخمین ماتریس R توسط دو ماتریس W و H بدست می‌آید. این معادله به صورت زیر است:

$$\boldsymbol{R} \approx \boldsymbol{W} \times \boldsymbol{H}$$

ماتریس R در واقع همان ماتریسی است که شامل افراد و سؤالات است که در این تکنیک سعی می‌شود با بهترین تقریب به دو ماتریس تجزیه شود که یکی از ماتریس‌ها شامل مؤلفه‌ها و آزمودنی‌های و ماتریس دیگر شامل سؤالات و مؤلفه‌هاست و هدف این تکنیک نیز رسیدن به همین مؤلفه‌هاست که در روانسنجی از آن با عنوان متغیر مکنون یاد می‌شود. هدف اصلی این پژوهش نیز یافتن تعداد این مؤلفه‌ها یا متغیرهای مکنون است. برای این هدف نیز دو مجموعه سؤال یکی در حوزه محتوایی حسابان و دیگری در حوزه هندسه مورد بررسی قرار گرفت که نتایج هر کدام به صورت جداگانه گزارش شده است. تمامی

روش‌های برآورد تعداد خوشه‌ها مانند بیشتر روش‌های داده‌کاوی براساس نمودارها و ابزارهای بصری است.

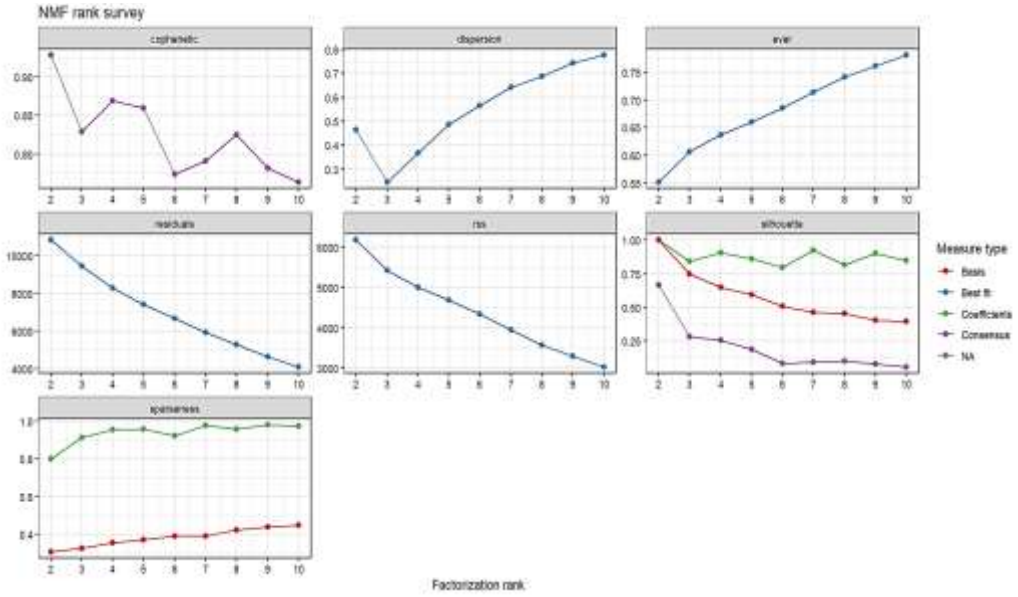
حسابان: در ابتدا ۱۹ سؤال حوزه محتوایی حسابان مورد بررسی قرار گرفت. در شکل زیر یک مجموعه نمودار نمایش داده شده است که از اجرای دستور زیر به دست آمده‌اند:

```
estim.r<- nmf(Hesaban, 2:10)
```

براساس دستور بالا تابع nmf برای داده‌های حسابان که شامل همان پاسخ آزمودنی‌ها است اجرا شده و تعداد خوشه‌های پیشنهادی ۲ تا ۱۰ خوشه است. خروجی این دستور در شکل زیر نمایش داده شده است.

Brunet و همکاران (2004) پیشنهاد می‌کنند که براساس ضریب کوهنتیک[1] در مورد تعداد خوشه‌ها تصمیم‌گیری شود؛ آن‌ها معتقدند در نقطه‌ای که بعد از آن نمودار سیر نزولی دارد برابر با تعداد خوشه‌ها است. همچنین Hutchins (2008) معتقد است که تعداد خوشه‌ها برابر است با نقطه عطف نمودار rss.

شکل ۱. نمودار برآورد تعداد خوشه‌های سؤالات حسابان

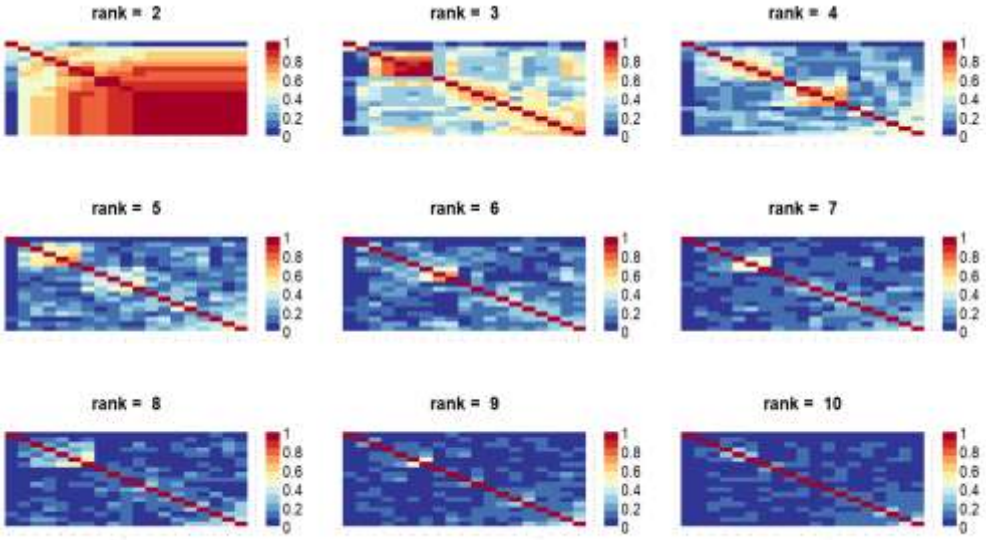


1. cophenetic

براساس نمودارهای بالا و با توجه به رهنمودهای مطرح شده تعداد خوشه‌های مناسب برای این داده‌ها برابر با ۲ یا ۴ خوشه است. در روش دیگر به ماتریس‌های اجماع[1] رجوع می‌شود. ماتریس‌های اجماع از طریق دستور زیر به دست می‌آیند.

```
consensusmap(estim.r2, tracks=NA)
```

شکل ۲. ماتریس‌های اجماع برآورد تعداد خوشه‌های سؤالات حسابان



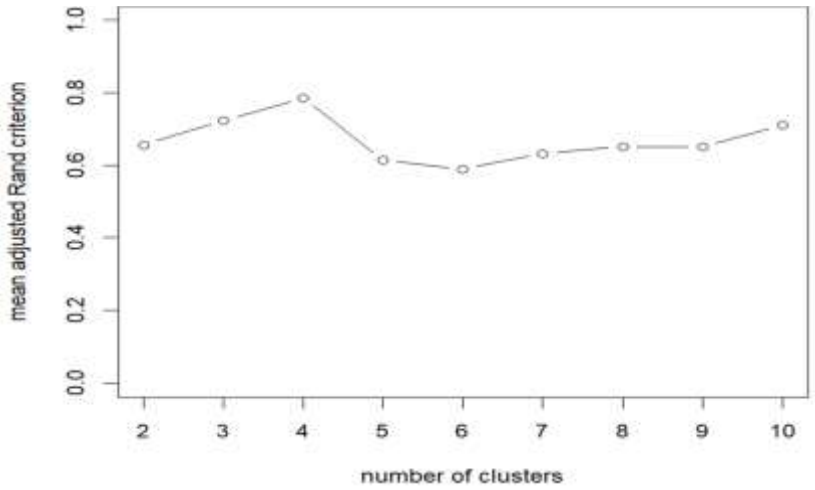
براساس ماتریس‌های اجماع ماتریس رتبه‌ای که آشفتگی کمتری دارد و تقریباً شمایل یک ماتریس قطری را نشان می‌دهد، آن رتبه با تعداد خوشه‌ها برابر است. در تصویر بالا مشاهده می‌شود که هرچه تعداد رتبه‌ها افزایش می‌یابد، آشفتگی ماتریس‌ها نیز کمتر شده است. اما اگر از رتبه دو شروع کنیم و به ترتیب پیش برویم، ماتریس برابر با رتبه ۶ به نسبت ماتریس‌های قبلی‌اش مرتب‌تر بوده و تا حدی نشان دهنده یک ماتریش قطبی است. ماتریس مربوط به رتبه ۱۰ بهترین حالت را در بین ماتریس‌ها دارد.

در ادامه تعیین تعداد خوشه‌ها از طریق روش تحلیل مؤلفه‌های اصلی بررسی شده است. برای این منظور از دو نمودار پایداری پارتیشن‌ها و نمودار پراکندگی شاخص راند تعدیل شده[2] استفاده شده است. برای رسم این نمودارها از دستورات زیر استفاده شد:

1. consensus matrix
2. Dispersion of the ajusted Rand index

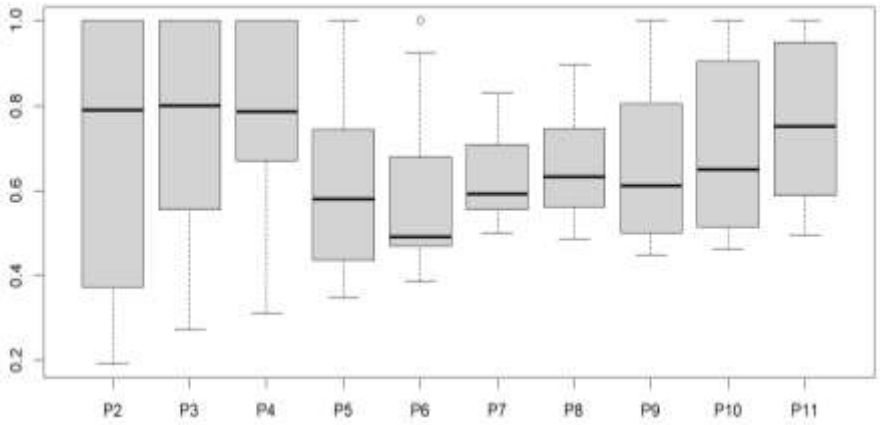
```
tree <- hclustvar(Hesaban)
plot(tree)
stab<-stability(tree,B=40)
plot(stab,nmax=10)
boxplot(stab$matCR[,1:10])
```

شکل ۳. نمودار برآورد پارتیشن‌ها برای سؤالات حسابان



براساس نمودار پایداری پارتیشن‌ها، نقطه‌ای که در آن نمودار بیشترین برافراشتگی را دارد برابر است با تعداد خوشه‌ها (Chavent, 2011) که در اینجا نمودار در نقطه ۴ دارای بیشترین برافراشتگی است. بنابراین تعداد خوشه‌های برابر با ۴ براساس این نمودار بهینه است.

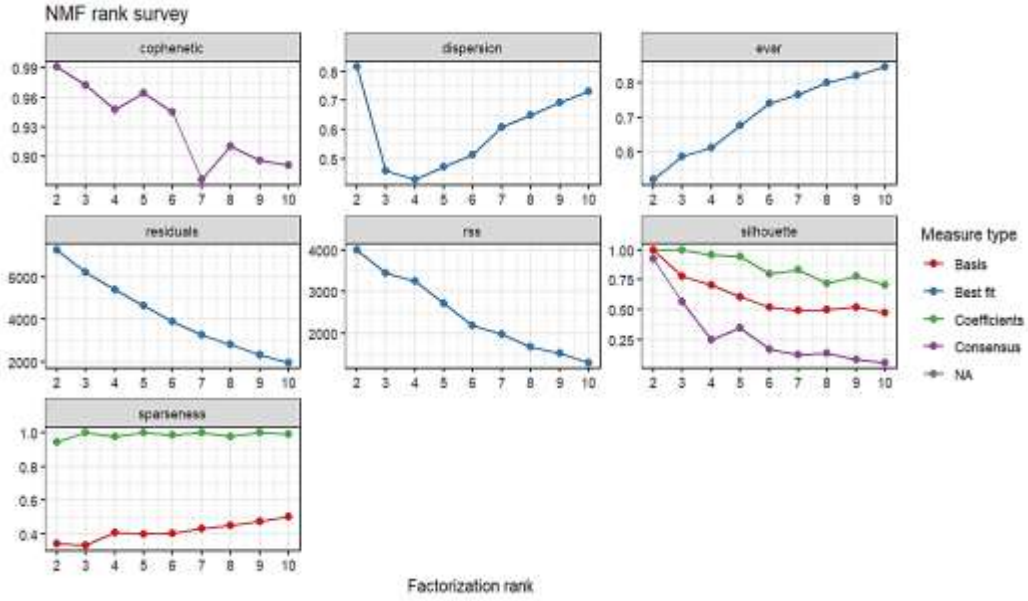
شکل ٤. نمودار جعبه‌ای پراکندگی شاخص راند تعدیل‌شده برای سؤالات حسابان



در این نمودار تعداد خوشه‌ها برابر با مقداری است که در آن شاهد کمترین پراکندگی باشیم. براساس این معیار مناسب‌ترین تعداد خوشه برابر با ۷ یا ۸ خوشه است.

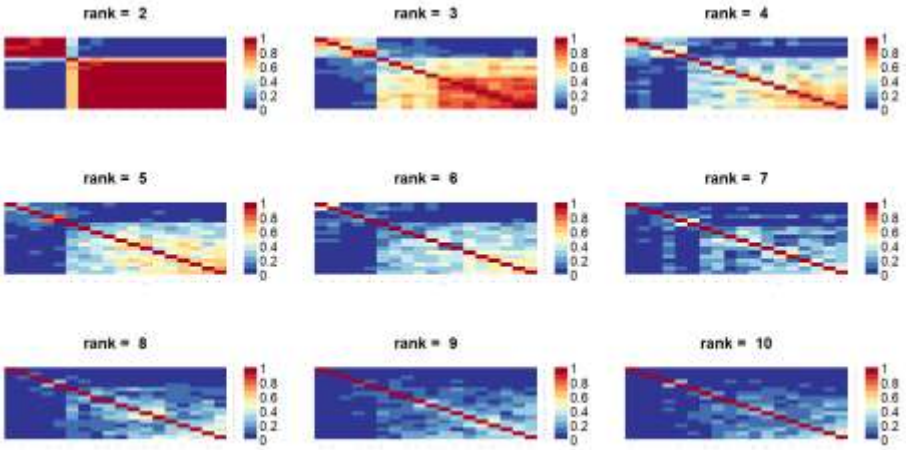
هندسه: در اینجا ۱۸ سؤال مربوط به حوزه محتوایی هندسه با همان دستورات قبلی اجرا و همان خروجی‌ها حاصل شد که در ادامه گزارش شده است.

شکل ۵. نمودار برآورد تعداد خوشه‌های سؤالات هندسه



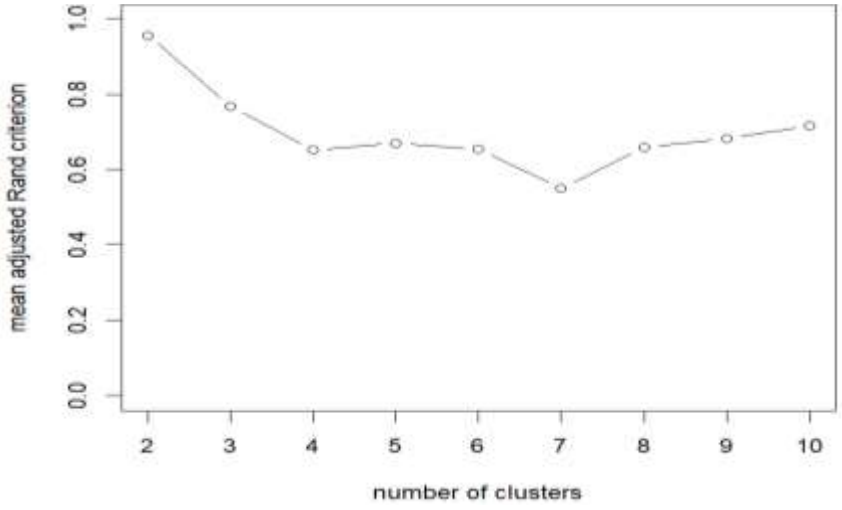
براساس نمودار مربوط به ضریب کوهنتیک تعداد خوشه برابر با ۲ مناسب‌ترین تعداد است و بعد از آن تعداد ۵ خوشه توصیه می‌شود.

شکل ٦. ماتریس‌های اجماع برآورد تعداد خوشه‌های سؤالات هندسه



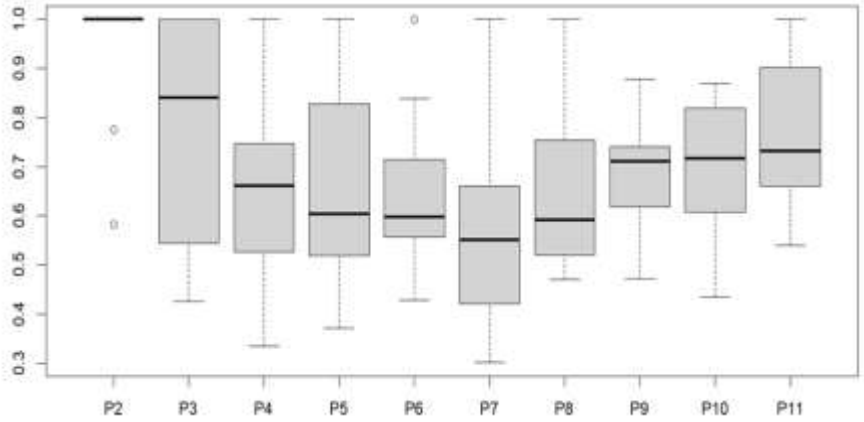
براساس ماتریس‌های اجماع، ماتریس مربوط به رتبه ۲ مرتب بوده و تقریباً شکل یک ماتریس قطری را نشان می‌دهد. همچنین از ماتریس رتبه ۷ به بعد نیز آشفتگی کم شده و تقریباً ماتریس‌ها به سمت قطری شدن پیش رفته‌اند. دراینجا نتیجه ماتریس اجماع با نتیجه ضریب کوهنتیک همخوانی دارد. در ادامه نتایج تحلیل مؤلفه‌های اصلی گزارش شده است.

شکل ۷. نمودار برآورد پارتیشن‌ها برای سؤالات هندسه



براساس نمودار پایداری پارتیشن‌ها در نقطه ۲ نمودار دارای بیشترین برافراشتگی است. بنابراین تعداد خوشه‌های برابر با ۲ پیشنهاد می‌شود.

شکل ٤. نمودار جعبه‌ای پراکندگی شاخص راند تعدیل‌شده برای سؤالات هندسه



براساس این نمودار نیز کمترین پراکندگی در تعدا خوشه‌های برابر با ۲ مشاهده می‌شود و بعد از آن تعداد خوشه برابر با ۶ و ۹ پراکندگی کمتری را نشان داده‌اند.

بحث و نتیجه‌گیری

یافتن صفت و ویژگی‌های پنهان با استفاده از داده‌های آشکار و مشاهده‌شده همواره یکی از دغدغه‌های حوزه روانسنجی و سنجش آموزشی بوده است. به همین دلیل نظریه‌ها و روش‌های مختلفی در طول زمان توسط افراد مختلفی ارائه شده است و همه این‌ها در تلاش برای رسیدن به نتیجه مطلوب و معتبر و تکمیل روش‌های پیشین هستند. اما تنوع این روش‌ها باوجود فوایدی که دارد گاهی باعث سردرگمی پژوهشگران در انتخاب روش مناسب و بهینه می‌شود. به همین دلیل در سال‌های اخیر برخی از پژوهش‌ها به این سمت (مقایسه روش‌ها) حرکت کرده‌اند. مثلاً Barnes (2003) طی پژوهشی روش ماتریس Q را با تحلیل عاملی و تحلیل خوشه مقایسه کرده است یا در پژوهشی دیگر مدل‌های کلی تشخیصی[1] با شبکه‌های بیزی[2] مورد مقایسه قرارگرفته است (Wu, 2013) که همه این موارد در جهت رواسازی نتایج صورت می‌گیرد. در واقع رواسازی فرآیندی است در جهت ارزیابی

1. General Diagnostic Models (GDM)
2. bayesian networks

استدلال‌های منطقی و شواهد علمی که ادعا را تأیید می‌کنند (تیلور؛ ترجمه یونسی، ۱۳۹۸). بنابراین یکی از وظائف پژوهشگران حوزه سنجش، رواسازی، بررسی و ارزیابی این روش‌ها و در صورت امکان معرفی روش بهینه برای پژوهش‌های آتی است.

PeñaAyala (2013) در کتاب خود با عنوان "داده‌کاوی آموزشی: کاربردها و روندها[1] ضمن معرفی رویکرد داده‌کاوی آموزشی یادآور می‌شود که به دلیل شرایط خاص سیستم‌های آموزشی لازم است در کنار تکنیک‌های داده‌کاوی آموزشی از رشته‌های نزدیک به آن شامل آمار، یادگیری ماشین، متن‌کاوی[2]، روانسنجی و غیره کمک گرفته شود. برای مثال در پژوهشی ابتدا با روش خوشه‌بندی K-MEANS دانش‌آموزان را خوشه‌بندی کرده و سپس از طریق نظریه سؤال پاسخ به برآورد توانایی آن‌ها پرداخته شده است (Chang & Yang, 2009).

داده‌کاوی تحلیل مجموعه‌ای از داده‌های مشاهده شده (اغلب بزرگ) برای یافتن روابط غیرمنتظره و خلاصه سازی داده‌ها به روش‌های جدیدی است که برای صاحب داده هم قابل درک و هم مفید باشد. (Hand et al., 2001). فرآیندی که تحت عنوان داده‌کاوی شناخته می‌شود، تفاوت قابل‌ملاحظه‌ای با تحلیل‌های آماری سنتی دارد، زیرا در این نوع از تحلیل، تحلیل جداده‌ها اکتشافی بوده و به‌وسیله فرضیات قبلی هدایت نمی‌شود (Schumann, 2005). مثلاً برای استخراج ماتریس Q از طریق تکنیک‌های داده‌کاوی و بدون دخالت متخصصان نیز پیشنهادهایی شده و پژوهش‌هایی (Casalino et al., 2017؛ Pardos & Dadu, 2018) در این حوزه شکل گرفته است. تکنیک مورد استفاده این پژوهشگران طبقه‌بندی ماتریس‌های نامنفی بود. این تکنیک با وجود آنکه در زمره روش‌های بدون نظارت قرار می‌گیرد نیاز به تعیین تعداد خوشه‌ها دارد و اگر در این زمینه پژوهشگر صاحب نظر نباشد با مشکل مواجه خواهد شد.

دراین پژوهش روش‌های برآورد تعداد خوشه‌ها در تکنیک طبقه‌بندی ماتریس‌ها نامنفی مورد مطالعه قرار گرفت و با روش‌های تعیین تعداد تکنیک مؤلفه‌های اصلی نیز مقایسه گردید. در بررسی سؤالات حسابان مشاهده شد که نتیجه برآورد تعداد خوشه از طریق روش ضریب کوهنتیک و ماتریس اجماع با یکدیگر متفاوت هستند؛ اما نکته جالب این بود که نتیجه روش ضریب کوهنتیک با روش پایداری پارتیشن‌ها از تحلیل مؤلفه‌های اصلی یکسان

1. Educational data mining: applications and trends
2. text mining

بود و نتیجه ماتریس اجماع با پراکندگی نمودار جعبه‌ای مربوط به تحلیل مؤلفه‌های اصلی همخوانی داشت. اما در بررسی سؤالات هندسه مشاهده شد که نتایج تمامی روش‌ها وجود تعداد ۲ خوشه را تأیید می‌کنند؛ در حقیقت در مورد این سؤالات بین روش‌ها اختلاف نظری وجود ندارد.

همان‌طور که در نتایج مشاهده شد بین نتایج به دست آمده از دو مجموعه سؤال تفاوت وجود دارد. یکی از دلایل این تفاوت را می‌توان در نوع سؤالات این مجموعه‌ها جستجو کرد. در واقع سؤالات هندسه فقط از دو کتاب هندسه ۱ و هندسه ۲ طرح شده و این کتاب‌ها به لحاظ محتوایی از سؤالات حسابان که از کتاب‌های ریاضی و حسابان هر سه سال مقطع متوسطه دوم طرح می‌شوند همگن‌تر هستند و سؤالات حسابان طیف وسیع‌تری از سؤالات را در بر دارند. همچنین با بررسی درصد پاسخگویی به سؤالات این دو مجموعه واضح است که تقریباً بیشتر سؤالات هندسه برای آزمودنی‌ها دشوار است در صورتی که سؤالات حسابان به لحاظ دشواری یکسان نبوده و برخی از سؤالات این حوزه نسبت به سؤالات دیگر آن برای پاسخ‌گویی آسان‌تر هستند که دلیل آن در دشورای متفاوت مباحث ارائه شده در کتاب‌های سال اول نسبت به کتاب‌های سال آخر است. در واقع دروس سال آخر کتاب حسابان به لحاظ دشواری نسبت به کتاب ریاضی ۱ و ۲ سنگین‌تر بوده و فهم ریاضی بالاتری را می‌طلبد؛ بنابراین سؤالات طراحی شده از کتاب سال دهم به مراتب راحت‌تر از سؤالات طراحی شده از کتاب حسابان سال آخر است. در نهایت یکی از محدودیت‌های این پژوهش وابسته بودن آن به سؤالات مورد بررسی است. به همین خاطر پیشنهاد می‌شود که برای پژوهش‌های آینده به کمک شبیه‌سازی داده‌ها مقایسه این روش‌ها انجام شده و مورد بررسی قرار گیرد.

منابع

تیلور، کاترین اس. (۱۳۹۸). *روایی و رواسازی*. ترجمه جلیل یونسی. تهران: انتشارات دانشگاه علامه طباطبائی.

سرمد، زهره، بازرگان، عباس و حجازی، الهه. (۱۳۹۱). *روش‌های تحقیق در علوم رفتاری*. تهران: نشر آگاه.

References

Baepler, P., & Murdoch, C. J. (2010). Academic analytics and data mining in higher education. *International Journal for the Scholarship of Teaching and Learning*, *4*(2), 1-9.

Barnes, T. M. (2003). The q-matrix method of fault-tolerant teaching in knowledge assessment and data mining.

Berkhin, P. (2006). A survey of clustering data mining techniques. In *Grouping multidimensional data* (pp. 25-71). Springer, Berlin, Heidelberg.

Brunet, J. P., Tamayo, P., Golub, T. R., & Mesirov, J. P. (2004). Metagenes and molecular pattern discovery using matrix factorization. *Proceedings of the national academy of sciences*, *101*(12), 4164-4169.

Carthy, A., Gray, G., McGuinness, C., & Owende, P. (2014). A review of psychometric data analysis and applications in modelling of academic achievement in tertiary education.

Casalino, G., Castiello, C., Del Buono, N., Esposito, F., & Mencar, C. (2017, July). Q-matrix extraction from real response data using nonnegative matrix factorizations. In *International Conference on Computational Science and Its Applications* (pp. 203-216). Springer, Cham.

Chang, W. C., & Yang, H. C. (2009). Applying IRT to Estimate Learning Ability and K-means Clustering in Web based Learning. *JSW*, *4*(2), 167-174.

Chavent, M., Kuentz, V., Liquet, B., & Saracco, L. (2011). ClustOfVar: An R package for the clustering of variables. *arXiv preprint arXiv:1112.0295*.

Freitas, A. A. (2002). *Data mining and knowledge discovery with evolutionary algorithms*. Springer Science & Business Media.

Hand, D. J., Mannila, H., & Smyth, P. (2001). Retrieval by Content.

Hutchins, M. J., & Sutherland, J. W. (2008). An exploration of measures of social sustainability and their application to supply chain decisions. *Journal of cleaner production*, *16*(15), 1688-1698.

Jose, P. E. (2013). *Doing statistical mediation and moderation*. Guilford Press.

Kerr, D., & Chung, G. K. (2012). Identifying key features of student performance in educational video games and simulations through cluster analysis. *JEDM| Journal of Educational Data Mining*, *4*(1), 144-182.

Larose, D. T., & Larose, C. D. (2014). *Discovering knowledge in data: an introduction to data mining* (Vol. 4). John Wiley & Sons.

Man, K., Harring, J. R., & Sinharay, S. (2019). Use of data mining methods to detect test fraud. *Journal of Educational Measurement*, *56*(2), 251-279.

Mislevy, R. J., Behrens, J. T., Dicerbo, K. E., & Levy, R. (2012). Design and discovery in educational assessment: Evidence-centered design, psychometrics, and educational data mining. *Journal of educational data mining*, *4*(1), 11-48.

Nguyen, D. Q., Nguyen, T. D., Nguyen, D. Q., & Phung, D. (2017). A novel embedding model for knowledge base completion based on convolutional neural network. *arXiv preprint arXiv:1712.02121*.

Peña-Ayala, A. (Ed.). (2013). *Educational data mining: applications and trends* (Vol. 524). Springer.

Pardos, Z. A., & Dadu, A. (2018). dAFM: Fusing Psychometric and Connectionist Modeling for Q-matrix Refinement. *JEDM Journal of Educational Data Mining*, *10*(2), 1-27.

Reimann, C., Filzmoser, P., Garrett, R., & Dutter, R. (2011). *Statistical data analysis explained: applied environmental statistics with R*. John Wiley & Sons.

Schumann, J. A. (2005). *Data mining methodologies in educational organizations*. University of Connecticut.

Wu, H. (2013). Comparison of General Diagnostic Models (GDM) and Bayesian Networks Using a Middle School Mathematics Test.
Yu, Z. (2009). Optimization techniques in data mining with applications to biomedical and psychophysiological data sets. *Theses and Dissertations*, 274.

استناد به این مقاله: رحیمی، ریحانه، یونسی، جلیل، مقدم‌زاده، علی و عسگری، محمد. (۱۴۰۱). برآورد تعداد خوشه‌ها در تکنیک طبقه‌بندی ماتریس‌های نامنفی .*فصلنامه اندازه‌گیری تربیتی*، ۱۳(۵۰)، ۱۳۵-۱۵۲.
doi: 10.22054/jem.2023.68624.3384

Educational Measurement is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.