نشریه علمی مطالعات مدیریت کسب و کار هوشمند- سال نهم ـ شماره ۳۳ـ پاییز ۱۳۹۹
مقاله پژوهشی/ صفحات ۳۸- ۵

Doi: 10.22054/IMS.2019.7824.1092

مدلی برای بخش‌بندی یادگیرندگان و بهبود عملکرد آموزشی با استفاده از الگوریتم‌های داده‌کاوی

سینا رئیسی وانانی *
ایمان رئیسی وانانی **
محمدتقی تقوی فرد ***

چکیده

ارزیابی عملکرد آموزشی از طریق شناسایی و تحلیل داده‌های حاصل از فعالیت‌های یادگیرندگان، می‌تواند به بهبود مؤثر عملکرد آموزشی منجر گردد. در پژوهش حاضر، داده‌های مربوط به دانش پذیران بین‌المللی، بر اساس روش تحقیق علم طراحی و با استفاده از روش‌های داده‌کاوی مورد بررسی قرار گرفته است. در این راستا تحقیقات انجام‌گرفته داخلی و بین‌المللی در دهه گذشته بررسی و مرور شده است و داده‌های تحصیلی و غیر تحصیلی یادگیرندگان در سه دسته خانوادگی، حمایتی و رفتار تحصیلی با استفاده از داده‌کاوی، خوشه‌بندی شده است. پس از اعتبارسنجی خروجی الگوریتم‌ها توسط شاخص‌های مرتبط و تعیین تعداد خوشه بهینه در هر بخش، خوشه‌ها نام‌گذاری و تحلیل شدند. تحلیل خوشه‌های شناسایی‌شده، نشان‌دهنده تجربه موفقیت یا شکست تحصیلی دانش پذیران و ریشه‌های عملکرد مؤثر در هر بخش است و روش نام‌گذاری ارائه‌شده، روشی نوین و قابل‌استفاده در اغلب مراکز آموزشی جهت تفکیک و تبیین عملکرد آموزشی است.

کلیدواژگان: آموزش، عملکرد تحصیلی، داده‌کاوی، خوشه‌بندی، طبقه‌بندی.

---

* کارشناسی ارشد، مدیریت، دانشکده مدیریت و حسابداری، دانشگاه علامه طباطبائی، تهران، ایران.

** استادیار، گروه مدیریت صنعتی، دانشکده مدیریت و حسابداری، دانشگاه علامه طباطبائی، تهران، ایران.(نویسنده مسئول)؛imanraeesi@atu.ac.ir

*** دانشیار، گروه مدیریت صنعتی، دانشکده مدیریت و حسابداری، دانشگاه علامه طباطبائی، تهران، ایران.

تاریخ دریافت:۱۳۹۸/۱۱/۱۲ **تاریخ پذیرش:۱۳۹۹/۰۳/۲۵**

مقدمه

در هر کشوری، آموزش از ارکان مهم رشد و توسعه آن کشور به‌حساب می‌آید که اگرچه آموزش در تمدن‌ها و کشورهای دوران گذشته، شکلی ساده داشته است اما در دوران کنونی، با ظهور فنّاوری‌ها و شکل‌گیری روابط درهم‌تنیده ملت‌ها و اقشار جوامع، ساختاری پیچیده پیدا کرده است. در چنین شرایطی یکی از ابزارهای مهمی که می‌تواند ما را در حل مسائل مربوط به آموزش و همچنین تبیین صحیح‌تر مسائل آن یاری نماید، ابزارهای مبتنی بر فناوری اطلاعات است. همچنین از میان فرآیندهای آموزشی، نظیر استعدادیابی، شیوه‌های تدریس، ارزیابی آموخته‌ها، مدیریت انضباط و ...، پیش‌بینی عملکرد تحصیلی دانش آموزان جایگاه ویژه‌ای در میان مقالات و تحقیقات بین‌المللی داشته است (کانسپسیون برگوس، ۲۰۱۸؛ آسیف و همکاران[1]، ۲۰۱۷؛ مربوطی و همکاران[2]، ۲۰۱۶؛ شهیری و همکاران[3]، ۲۰۱۵؛ کالاریز و همکاران، ۲۰۱۴)؛ چراکه این شاخص، شاخص تعیین‌کننده‌ای در ارزیابی تصمیم‌گیری‌های مدیریتی و اهداف آموزشی است و اگر چنانچه بتوان راهکارهای متناسبی را در جهت بهبود عملکرد تحصیلی بر اساس داده‌های مبتنی بر واقعیت ارائه نمود، تصمیم‌گیری مدیران را اثربخش‌تر خواهیم ساخت. اجمالاً در این قسمت سه موضوع عملکرد تحصیلی، داده‌کاوی آموزشی و راهکار پیش‌بینی آینده‌نگر را جهت تبیین دقیق‌تر موضوع مورد بررسی واقع می‌شود.

یکی از عناصر مهم در مسئله آموزش، عملکرد تحصیلی است. عملکرد تحصیلی فراگیران برای مدیریت آموزش عالی، دولتمردان، والدین و سرمایه‌گذاران اهمیت داشته و عامل مهمی در توسعه ملی شناخته می‌شود (نکولا و همکاران، ۲۰۱۵). به همین دلیل مراکز آموزشی به‌صورت فزاینده‌ای نیاز دارند عملکرد تحصیلی یادگیرندگان خود را مطالعه کرده و نسبت به بهبود آن راهکارهایی را ارائه دهند. این مطالعات در دو جهت می‌تواند مؤثر باشد:

1. Asif et al.
2. Marbouti et al.
3. Shahiri et al.

شناسایی یادگیرندگان مستعد، بی‌انگیزه و ناموفق و شناسایی راهکارهای قابل‌ارائه در جهت بهبود عملکرد آموزشی.

در دهه اخیر، تصمیم‌گیری مدیران به‌صورت چشمگیری از طریق روش‌های داده‌کاوی بهبود یافته است و داده‌کاوی آموزشی نیز یکی از زمینه‌هایی است که محققان بسیاری از طریق آن، عملکرد تحصیلی را پیش‌بینی کرده‌اند. روش‌هایی از قبیل ماشین بردار پشتیبان، شبکه‌های عصبی، نایوبیز، K میانگین و درخت تصمیم، یادگیری عمیق، شبکه‌های عصبی و دیگر روش‌ها. این روش‌ها در راستای محاسبه میزان تأثیرگذاری ویژگی‌های روان‌شناختی یادگیرندگان، ساختاری مراکز آموزشی و همچنین اجتماعی یادگیرندگان بر شاخص‌های داخلی و بین‌المللی تحصیلی نظیر معدل کل بکار گرفته شده‌اند (شهیری و همکاران، ۲۰۱۵). همچنین یکی از ضرورت‌های هر تصمیم‌گیری مدیریتی، در دست داشتن راهکارهای عملیاتی به همراه احتمال موفقیت و اثربخشی آن راهکارها است که استفاده از تکنیک‌های داده‌کاوی می‌تواند این مهم را محقق سازد.

یکی از کلیدواژه‌های مهم در این حیطه پژوهشی، عملکرد یادگیرنده به‌صورت خاص و همچنین انگیزش یادگیرنده و خروج از یادگیری از جمله مفاهیمی هستند. به‌عنوان نمونه می‌توان به افزایش عملکرد تحصیلی یادگیرندگان در دروس کمی (راحال و همکاران، ۲۰۱۵)، پیش‌بینی عملکرد تحصیلی با استفاده از درخت تصمیم و الگوریتم ژنتیک (همسا و همکاران[1]، ۲۰۱۶)، مدل پیش‌بینی شکست تحصیلی (گراد و همکاران ۲۰۱۸) اشاره کرد. این مدل‌ها سعی دارند الگوهای رفتار تحصیلی را به دست آورند تا تصمیم‌گیری مدیران نسبت به عملکرد آموزشی یادگیرندگان را بهبود دهند. این پژوهش از میان مباحث مطرح‌شده در تحقیقات میان‌رشته‌ای داده‌کاوی آموزشی، به ارائه راهکارهای پیش‌بینی آینده‌نگر پرداخته است. این راهکارها بر اساس داده‌های پیشین هستند و رفتار آینده یادگیرندگان را بر اساس رفتارهای گذشته پیش‌بینی کرده و نهایتاً جهت بهبود تصمیم‌گیری‌ها مدیریتی مورد استفاده قرار می‌گیرند.

1. Hamsa et al.

پیشینه پژوهش

توسعه و پیشرفت روزافزون فناوری اطلاعات و ارتباطات و ظرفیت بالقوه آن در بهبود آموزش، پژوهشگران حوزه آموزش را بر آن داشته تا در فرآیند آموزش نیز استفاده از فن‌آوری را به‌مثابه یک مزیت رقابتی در اولویت قرار دهند. از طرف دیگر نیاز جوامع درحال‌توسعه به استفاده بهینه از زمان و امکانات و توسعه انعطاف‌پذیر آموزش، منجر به گرایش بیشتر به استفاده از آموزش الکترونیکی شده است (ملک محمدی و شیروانی، ۱۳۹۳).

یکی از روش‌های آموزشی که در آن فراگیر بر اساس توانایی‌های خود به فعالیت و یادگیری می‌پردازد روش نوین آموزش الکترونیکی و نقش فناوری‌های نوین در این حوزه است (زاهد بابلان، معینی کیا و درخشان فرد، ۱۳۹۵) این نوع آموزش جستجوی اطلاعات را در دنیای فن‌آور امروز متحول کرده و باوجود چالش‌هایی از قبیل تقاضای روزافزون برای آموزش عالی و عدم‌کفایت بودجه، کمبود مدرس تمام‌وقت و نیاز به حذف محدودیت‌های جغرافیایی مورد توجه جدی قرار گرفته است (رضایی و زاهدی، ۱۳۹۷).

انعطاف‌پذیری آموزش الکترونیک، مدیریت محتوا، تعامل هم‌زمان استاد و دانشجویان و همچنین سازمان‌دهی ساختار دوره‌ها و ارزیابی، سبب شده است که آموزش از حالت آموزش محور به حالت یادگیری محور تغییر پیدا کند (زاهد بابلان، معینی کیا و درخشان فرد، ۱۳۹۵).

شکل ۱: سیستم مدیریت آموزشی (زاهد بابلان، معینی کیا و درخشان فرد، ۱۳۹۵)

همچنین در عصر امروز یادگیری از طریق گوشی‌های هوشمند مفهوم آموزش را به‌طورکلی تغییر داده است. امروزه ارائه‌دهندگان خدمات الکترونیکی آموزش آنلاین از ارائه راهکارهای یادگیری از طریق تلفن همراه سود بسیار زیاد و قابل‌توجهی را به دست می‌آورند. دلایل و مزایای استفاده از گوشی‌های همراه برای یادگیری از دیدگاه دانش آموزان به شرح زیر است:

- محتواهای آموزشی را می‌توان "در حال حرکت" بدون نیاز به برنامه‌ریزی و تنظیم زمان دقیقی برای یک درس مطالعه کرد.
- یک روش یادگیری به شکل گروهی است. جامعه دانش‌اندوزان به‌اندازه کافی برای کمک به حل مسائل چالشی و تقویت بحث‌های سازنده میان اعضا بزرگ است.

- دانش آموزان همواره به اطلاعات موردنیاز خود دسترسی دارند و در مواقع لزوم به‌راحتی می‌توانند از آن‌ها استفاده کنند.
- محتویات موبایل به‌صورت فشرده در اختیار دانش آموزان قرار می‌گیرد بنابراین آن‌ها را در مدت‌زمان کوتاهی مطالعه می‌کنند.

علاوه بر این، یکی از روش‌های جذابیت آفرینی در حوزه آموزش، بازی وارسازی است. در واقع، در این روش، با استفاده از عناصر و عوامل به‌کاررفته در بازی‌ها، توسعه‌دهندگان محتواهای آموزشی توجه مخاطبان را به محتوای ارائه شده جلب می‌کنند و مشارکت آن‌ها را نیز افزایش می‌دهند. با استفاده از رویکرد بازی وارسازی، توسعه‌دهندگان نرم‌افزارها قادر به تبدیل محتوای پیچیده به یک محتوای ساده می‌باشند و سرگرمی را با فرآیند یادگیری ترکیب می‌کنند تا انگیزه دانش آموزان را تقویت کنند.

از میان روش‌ها و تکنیک‌های نوین مورداستفاده در آموزش، تکنیک‌های داده‌کاوی در دهه اخیر مورد استقبال قابل‌توجهی قرار گرفته است. مطمئناً تکنیک‌های داده‌کاوی می‌توانند در اختیار سیاست‌گذاران آموزش‌وپرورش قرار بگیرند و مدل‌های مبتنی داده و شواهد، پشتیبانی برای حمایت از افزایش کارایی و کیفیت آموزش و یادگیری قرار بگیرند. همچنین استفاده از تکنیک‌های داده‌کاوی مختلف می‌تواند تغییرات سیستمی را ایجاد کند و ابزاری مناسب برای حل کمک به حل اغلب مسائل و دغدغه‌های مسئولین سازمان‌های آموزش عالی باشد (وان بارنولد و همکاران[1]، ۲۰۱۲). داده‌کاوی آموزشی (EDM) و تحلیل‌های پیشرفته یادگیری(LA)، دو حوزه علمی هستند که در رابطه با تأثیر داده‌کاوی در آموزش رسمی و غیررسمی است. ابزارهای داده‌کاوی آموزشی، اکوسیستمی هستند می‌توانند داده‌های به‌دست‌آمده از فرآیند آموزشی را جمع‌آوری کرده و پس از تجزیه‌وتحلیل، گزارشی از آن‌ها ارائه نمایند که در جهت بهبود عملکرد تحصیلی کارآمد باشد.

داده کاوی آموزشی و تجزیه‌وتحلیل پیشرفته یادگیری این امکان را به ما می‌دهد که مدل‌های گذشته آموزش و یادگیری را دوباره‌سازی نماییم و راهکارهای جدیدی برای مسائل

1. Van Barneveld et al.

گذشته داشته باشیم. (براون و همکاران[1]، ۲۰۱۵) سیستم‌های مدیریت آموزش (LMS) اگرچه یک محیط یادگیری مجازی را برای دانشجویان فراهم می‌کند اما درک محدودی از موضوعات مرتبط با یادگیری ارائه می‌نماید و تحلیل‌های پیشرفته‌ای از فضای یادگیری و الگوهای آن به مدیران آموزشی و اساتید ارائه نمی‌کنند. این خلأ را می‌توان از طریق استفاده از داده کاوی آموزشی و تحلیل‌های پیشرفته آموزش جبران کرد. این دو حوزه می‌توانند یادگیری را به‌صورت شخصی‌سازی‌شده محقق سازند و تعامل بهتری را در فضای آموزش ایجاد نمایند. به همین دلیل اثربخشی آموزش افزایش پیدا کرده و نقشه عملکرد مدرسین و دانشجویان بهتر و دقیق ترسیم می‌شود(آلدووا و همکاران[2]، ۲۰۱۹). به همین جهت یادگیرندگان به‌خصوص دانشجویان، انتظار دارند تحلیل‌های پیشرفته یادگیری بتواند از برنامه‌ریزی درسی آن‌ها پشتیبانی کند، خودارزیابی را تقویت نماید، توصیه‌هایی که قابلیت شخصی‌سازی داشته باشند، ارائه شوند و تحلیل‌های یادگیری نیز به‌صورت شخصی‌سازی‌شده ارائه گردند(شوماخر و افنتالر[3]، ۲۰۱۸).

تلاش‌های فعلی برای تعمیم استفاده از EDM و LA در آموزش عالی منجر به مطالعات بسیاری در مورد اثربخشی چنین استفاده‌هایی در زمینه آنلاین شده است. این کاربرد گسترده باعث شد دانشگاه‌ها حجم زیادی از داده‌های مربوط به دانشجویان خود و فرایند یادگیری ذخیره‌شده در سیستم‌های مدیریت یادگیری و محتوا (LMS / CMS)را جمع‌آوری کنند (تیر و ال هالیس، ۲۰۱۲). همچنین ادبیات فعلی در مورد استفاده از داده کاوی در بخش آموزش عالی عمدتاً در استفاده از تکنیک‌هایی از قبیل طبقه‌بندی، خوشه‌بندی، قوانین انجمن، آمار و تجسم برای پیش‌بینی، گروه، مدل و نظارت متمرکز است.

محققان EDM و LA ابعاد دیگری را در رابطه با فعالیت‌های دانشگاهی مطرح می‌کنند که از جمله آن‌ها یادگیری مشارکتی با پشتیبانی رایانه باهدف استخراج الگوهای مشارکتی در مباحث آموزشی (دیملو و همکاران، ۲۰۱۰؛ پررا و همکاران، ۲۰۰۹)، حمایت مدرسان در شناخت مشارکتی مدل یادگیری دانشجویان (گادیوسو و همکاران، ۲۰۰۹)، ارزیابی مطالب

1. Brown
2. Aldowah et al
3. Schumacher & Ifenthaler

آموزشی دانشگاه و پیشرفت برنامه درسی (کامپاگنی و همکاران[1]، ۲۰۱۴؛ جیانگ و همکاران، ۲۰۱۶)، عوامل مرتبط با موفقیت، شکست و نیت ترک تحصیل دانشجویان (کمبروزی و همکاران[2]، ۲۰۱۵؛ لیکورنتزو و همکاران، ۲۰۰۹؛ مارکزز ورا و همکاران، ۲۰۱۶)، برنامه‌ریزی و استراتژی گزاری مؤسسات آموزشی (کاپوتی و گاریدو[3]، ۲۰۱۵؛ مانکاد، ۲۰۱۶) و همچنین شناخت مدرسان و نحوه تصمیم‌گیری مدیران آموزشی. همچنین نستور و همکاران[4] (۲۰۱۸) در تحقیق مروری خود، ۶ دسته داده‌های مورداستفاده در فرآیند داده کاوی آموزشی را این‌طور طبقه‌بندی کردند: ۱- تحلیل ردیابی‌های چشم یادگیرندگان، تحلیل تعاملات آنلاین، تحلیل نظرسنجی‌های مدرسه‌ای، تحلیل داده‌های لاگ در سطح فردی و مشارکتی، تحلیل یادگیری مصور در ارتباط با اینترنت اشیاء.

آلدووا و همکاران در سال ۲۰۱۹ پژوهش مروری درباره داده کاوی آموزشی انجام داده‌اند که بررسی ۴۰۲ مقاله منتخب از میان ۱۲۰۰ مقاله بوده و در پاسخ به دو سؤال اساسی تدوین شده است:

چطور می‌توان از تکنیک‌های داده کاوی آموزشی و تحلیل‌های پیشرفته یادگیری در حل مسائل کاربردی حوزه آموزش استفاده کرد؟

کدام‌یک از تکنیک‌های داده کاوی برای حل مسائل آموزشی مناسب هستند؟

این مقالات در سال‌های ۲۰۰۰ تا ۲۰۱۷ به چاپ رسیده‌اند که از منابعی نظیر Scopus - Web of Science- Google Scholar – ERIC - Science Direct - DBLP – ACM Digital Library – IEEE Xplore – Springer Linkگرداوری شده‌اند.

همچنین بر اساس تحقیق انجام‌شده در سال ۲۰۱۹ (آلدووا و همکاران، ۲۰۱۹) چهار حوزه اساسی در زمینه تحقیقات داده کاوی آموزشی وجود دارد که به شرح زیر می‌باشند:

1. Campagni
2. Cambruzzi et al.
3. Caputi & Garrido
4. Nistor et al.

تجزیه‌وتحلیل یادگیری مبتنی بر کامپیوتر[1](CSLA)

CSLA به معنای استفاده از تکنیک‌های داده کاوی جهت به دست آوردن اطلاعات مناسب برای اقدامات عملیاتی در زمینه آموزش است که این اطلاعات از تعامل یادگیرندگان با LMS ایجاد می‌گردد. یکی از نیازهای مهم مدرسانی که در فرآیند مدیریت آموزش هستند، ابزار سنجش تعامل یادگیرندگان با یکدیگر است که به‌وسیله آن بتوانند واکنش مناسبی به فرآیند یادگیری داشته باشند و اثربخشی آموزش را به‌درستی ارزیابی نمایند (ریتیلس و همکاران[2]، ۲۰۰۶). معمولاً EDM و LA برای ارزیابی مشکلات یادگیری از طریق ارزیابی تعامل دانش آموزان و نتایج یادگیری استفاده می‌شوند. داده‌های حاصل از این ارزیابی‌ها به‌طور بالقوه می‌توانند در تخمین یا تغییر سطح پشتیبانی موردنیاز برای افزایش خودآگاهی دانش آموزان در مورد فعالیت و محتوا کمک کنند. به‌عنوان‌مثال، داده‌های LMS با بررسی فعالیت‌های مرتبط با دوره، مانند انجمن‌های بحث، تحویل محتوا و ارزیابی، می‌توانند تنظیمات سیستم یادگیری را متناسب با هر یادگیرنده شخصی‌سازی نمایند.

تجزیه‌وتحلیل آینده‌نگر مبتنی بر کامپیوتر(CSPA[3])

ازآنجایی‌که می‌توان با استفاده از داده کاوی آموزشی، یادگیری را بهبود بخشید، می‌توان از EDM و LA برای پیش‌بینی عملکرد و حفظ یادگیرندگان در دوره‌های درسی بر اساس سنجش میزان پیشرفت، مشارکت، نمرات و دانش نهفته در فرآیند یادگیری، استفاده کرد. این تحلیل‌ها شامل ارزیابی مطالب یادگیری برای ارزیابی پیچیدگی تمرین‌ها و ارائه بازخورد برای پشتیبانی از یادگیری تصمیم‌گیری از طریق برنامه‌ریزی برای استراتژی‌های جدید است که نتایج کلی یادگیری را تقویت می‌کند. (سمبیرینگ و همکاران، ۲۰۱۱). لوان (۲۰۰۲) اظهار داشت که با استفاده از تکنیک‌های داده کاوی در زمینه یادگیری می‌تواند به کشف دانش و الگوهای پنهان در حجم زیادی از داده‌ها کمک کند و پیش‌بینی نتایج یا رفتارها را انجام دهد. بارادواج

1. Computer-Supported Learning Analytics
2. Retalis et al.
3. Computer Supported Predictive Analytics

و پال (۲۰۱۲) اظهار داشتند که از EDM و LA می‌توان برای کشف دانش استفاده کرد که به مدرسان کمک کند تا سریع‌تر یادگیرندگانی را که در معرض ترک تحصیل هستند شناسایی کنند و مشخص کنند چه کسانی نیازمند توجه بیشتر هستند.

تجزیه‌وتحلیل رفتار تحصیلی مبتنی بر کامپیوتر (CSBA[1])

استفاده از تکنیک‌های داده کاوی می‌تواند بینش قابل‌توجهی داشته و الگوهای ارزشمندی را در رفتارهای یادگیری دانش آموزان نشان دهد (هه[2]، ۲۰۱۳). هونگ و ژانگ (۲۰۰۸) از داده کاوی برای شناسایی الگوهای رفتاری و ترجیحی رفتاری یادگیرندگان هنگام شرکت در فعالیت‌های یادگیری آنلاین استفاده کردند. آن‌ها دریافتند که استفاده از EDM و LA منجر به بهبود تجربه یادگیری یادگیرندگان هنگام همکاری از راه دور می‌شود. در حال حاضر، بیشتر تمرکز بر روی EDM و LA بر استفاده از داده‌های زمان واقعی برای تنظیم یادگیری اطلاعات جدید است تا یادگیرندگان بتوانند مشکلات را با سطوح مختلف پیچیدگی حل کنند. به‌عنوان‌مثال، بیواز و جئونگ (۲۰۰۸) با درج اطلاعاتی در مورد دانش، انگیزه، فراشناخت و نگرش، الگوی دانشجویی را برای پیش‌بینی برخی فرایندهای یادگیری طراحی کردند. طبق رومرو و همکاران (۲۰۱۰)، EDM می‌تواند برای ارزیابی رفتار و فعالیت‌های نامنظم یادگیرندگان در یک محیط آنلاین مانند Moodle با ارزیابی ارتباط بین فعالیت‌های آنلاین یادگیرندگان و علائم نهایی آن‌ها استفاده شود. علاوه بر این، مک کویگ و بالدوین (۲۰۱۲) از EDM استفاده کردند تا یادگیرندگان موفق را بر اساس رفتار تعامل خود شناسایی کنند. گزارش شده است که داده‌های ورود به سیستم از LMS می‌توانند برای پیش‌بینی عملکرد یادگیرندگان در دوره آموزش الکترونیکی (موفقیت یا شکست) بدون نیاز به نتایج ارزیابی رسمی استخراج شوند. همچنین هرناندز و همکاران[3] (۲۰۱۹) در تحقیقات خود به این نتیجه

1. Computer-Supported Behavioral Analytics
2. He
3. Hernández et al.

رسیدند که استفاده از تحلیل‌های پیشرفته تحصیلی و داده کاوی در بررسی تعاملات آنلاین در رابطه با یادگیری، تأثیر بالایی در کشف قواعد یادگیری و خروجی‌های کلاس‌های درسی دارد.

تجزیه‌وتحلیل اطلاعات بصری مبتنی بر کامپیوتر (CSVA[1])

CSVA نوعی بررسی وضعیت است که تکنیک‌های مصورسازی اطلاعات را با روش‌های پیشرفته داده کاوی و نمایش دانش ترکیب می‌کند و اغلب برای ارائه یک تحلیل بصری از رفتارهای تحصیلی یادگیرندگان بکار می‌رود. در محیط‌های آموزشی، CSVA به معنای استفاده از ابزارهای مصورسازی برای یافتن بینشی از فرآیند آموزش و تجربه یادگیری دانشجویان است (پنا آیالا، ۲۰۱۴). به‌عنوان‌مثال، نگاشت مباحثات آنلاین و ارزیابی کیفیت تک‌تک پست‌ها (تعامل) بر اساس ویژگی‌های ساختاری موضوع می‌تواند به دانشجویان در شناسایی نکات و مباحث مربوطه کمک کند. طبق گفته جین و همکاران. (۲۰۰۹)، داده کاوی بصری مورد استفاده در یک سیستم ارزیابی آموزش عالی می‌تواند روش ارزیابی را منعطف‌تر، متنوع‌تر و بصری‌تر کند که در آن می‌توان بهره‌وری از فرایندهای یادگیری را بهبود بخشید. از طرف دیگر کومار و کادهار (۲۰۱۱) به پتانسیل استفاده از EDM برای استخراج دانش و اطلاعات معنی‌دار از مجموعه داده‌های بزرگ پرداختند و از این اطلاعات برای کشف الگوهای پنهان و روابطی استفاده کردند که می‌تواند برای تصمیم‌گیری بالاتر باشد. از نمودارها می‌توان برای نشان دادن تعامل دانشجویان با تمرین‌ها استفاده کرد که می‌تواند به مدرسان کمک کند که درک بهتری از رفتار آنلاین دانشجویان خود داشته و نسبت به آنچه در محیط آنلاین اتفاق می‌افتد، آگاه باشند (رومرو و همکاران، ۲۰۰۸). علاوه بر این، از ابزارهای تجسم داده می‌توان در آموزش عالی برای ساده‌سازی داده‌های پیچیده و ردیابی داده‌های چندبعدی دانش آموزان که از تعامل آن‌ها با سیستم‌های آموزشی آنلاین به دست آمده است، استفاده کرد. (رومرو و ونتورا، ۲۰۰۷).

1. Computer-Supported Visualization Analytics

ویرا و همکاران[1] (۲۰۱۸) پس از مطالعه مروری خود بر روی مقالات و تحقیقات انجام‌شده در این حوزه، چهار پیشنهاد ارائه کرده‌اند:

تحقیقات اندکی در زمینه کاربرد تحلیل‌های بصری برای کلاس درس انجام شده است و می‌توان آن‌ها را گسترش داد.

تحقیقات اندکی، داده‌ها را در ارتباط با عملکرد گذشته یادگیرندگان و اطلاعات جمعیت شناختی آن‌ها نشان می‌دهد.

همچنان از تکنیک‌های گذشته نظیر نمودارهای نقطه‌ای برای تحلیل وضعیت استفاده می‌شود و روش‌های نوین مورد استفاده قرار نمی‌گیرد.

خلأ قابل‌توجهی میان ابزارهای مصورسازی داده‌ها و تئوری‌های نظری تبیین شده وجود دارد که باید این خلأ رفع شود.

برخی از مقالات بین‌المللی و داخلی منتشرشده در این زمینه عبارت هستند از:

جدول ۱: مقالات بین‌المللی و داخلی در زمینه داده کاوی آموزشی

ردیف	عنوان	سال انتشار	نویسنده
۱	چالش‌های پیشروی داده کاوی آموزشی در آینده	۲۰۱۹	رایان و بیکر
۲	تحلیل آینده‌نگر عملکرد تحصیلی دانش آموزان در مدارس پایتخت برزیل	۲۰۱۹	اردواردو فرناندس و همکاران
۳	داده کاوی آموزشی: مروری بر ارزیابی فرآیند یادگیری الکترونیکی	۲۰۱۸	مرکوس و همکاران
۴	استفاده از داده کاوی جهت مدل‌سازی عملکرد تحصیلی: برنامه درسی برای جلوگیری از ترک تحصیل	۲۰۱۸	برگس و همکاران
۵	پیش‌بینی ترک تحصیل با استفاده از تکنیک‌های یادگیری ماشینی	۲۰۱۸	لورنز کمپر
۶	تحقیقی پیرامون ارزیابی عملکرد تحصیلی، تحلیل روند تحصیلی و پیش‌بینی استعداد یادگیرندگان بر اساس تکنیک‌های داده کاوی	2018	فن یانگ
۷	مدلی برای پیش‌بینی شکست تحصیلی دانشجویان رشته پزشکی در اولین سال تحصیلی	۲۰۱۷	گراد و همکاران
۸	تحلیل عملکرد تحصیلی دانشجویان کارشناسی با استفاده از روش‌های داده کاوی	۲۰۱۷	آسیف و همکاران
۹	ارزیابی کارایی تکنیک‌های داده کاوی جهت پیش‌بینی شکست تحصیلی یادگیرندگان	۲۰۱۷	اواندرو کاستا و همکاران

1. Vieira et al.

۱۰	روشی نوین برای داده کاوی آموزشی تحصیلات عالیه	۲۰۱۵	تویجری و همکاران
۱۱	پیش‌بینی یادگیرندگانی که سرعت کمی در یادگیری دارند با استفاده از روش‌های داده کاوی مبتنی بر طبقه‌بندی و پیش‌بینی	۲۰۱۵	پارنیت کور و همکاران
۱۲	تحلیل طول زمان تحصیل دانشجویان بر اساس روش‌های خوشه‌بندی و کشف قواعد انجمنی	۲۰۱۴	پتروس بلسیس و همکاران
۱۳	استفاده از داده کاوی آموزشی جهت گروه‌بندی یادگیرندگان در محیط یادگیری الکترونیکی به‌منظور شخصی‌سازی برنامه آموزش	۱۳۹۷	دی پیر و رابو
۱۴	کاربرد داده کاوی در سیستم آموزشی	۱۳۹۴	خیرخواه و جوانمرد
۱۵	ارائه و ارزیابی روشی کارآمد جهت بهبود تعیین رشته دانش آموزان مقطع متوسطه با استفاده از روش های رده بندی داده کاوی و یادگیری ماشین	۱۳۹۴	عطارزاده و همکاران
۱۶	ارزیابی بازخورد دانشجویان از دوره آموزشی و عملکرد مدرسین از طریق ترکیب روش های خوشه بندی و الگوریتم درخت تصمیم	۱۳۹۴	بنی رستم و طاهری پور
۱۷	توسعه مدلی برای پیش بینی آسیب پذیری تحصیلی دانشجویان مقطع کارشناسی رشته های مهندسی، مبتنی بر الگوریتم شبکه عصبی	۱۳۹۴	امین بیدختی و همکاران
۱۸	روش‌های داده کاوی برای سیستم مدیریت سطح علمی دانشجویان و مؤسسات آموزش عالی	۱۳۹۳	افرایش و همکاران
۱۹	پیش بینی فارغ التحصیلی دانشجویان در آموزش عالی با استفاده از مدل های داده کاوی	۱۳۹۳	بالافر و همکاران
۲۰	تحلیل و ارزیابی عملکرد فراگیران در سیستم‌های آموزش الکترونیکی با استفاده از تکنیک‌های داده کاوی	۱۳۹۲	نحوی و همکاران
۲۱	ارائه چارچوبی برای پیش‌بینی عدم موفقیت دانش آموزان با استفاده از تکنیک‌های مختلف داده کاوی	۱۳۹۲	فراهی و مرادی
۲۲	انتخاب الگوریتم داده کاوی مناسب برای تشخیص دلایل ترک تحصیل دانش آموزان(مورد کاوی: مدارس استان اردبیل)	۱۳۹۲	فراهی ومختاری
۲۳	چارچوبی برای الگوکاوی زمانی در سیستم‌های آموزش الکترونیکی: موردکاوی دانشگاه مجازی امام خمینی	۱۳۹۱	مینایی و همکاران
۲۴	ارتقای کیفیت آموزش در سامانه‌های آموزش الکترونیکی با استفاده از داده کاوی آموزشی	۱۳۹۱	مقصودی و همکاران
۲۵	کاربرد تکنیک‌های داده‌کاوی در محیط‌های آموزش الکترونیکی	۱۳۹۰	نریمی ساری و همکاران
۲۶	ارائه چارچوب جدیدی برای تعیین عوامل تأثیرگذار بر نمره ارزشیابی اساتید با ترکیب روش‌های داده کاوی	۱۳۹۰	مینایی بیگدلی و همکاران
۲۷	کاربردهای داده کاوی آموزشی در جهت بهبود عملکرد سیستم‌های آموزشی و مدیریت آموزش در سازمان‌ها	۱۳۸۹	فیروزه و همکاران
۲۸	مقایسه تکنیک‌های داده کاوی جهت شخصی‌سازی مطلوب‌تر در آموزش الکترونیک	۱۳۸۷	پاینده فر و همکاران

روش‌شناسی تحقیق

روش تحقیق پژوهش حاضر بر اساس علم طراحی است. روش تحقیق علم طراحی مبتنی بر ارائه یک فرآورده از جنس روش، فرآیند و محصولی است که بتواند یک مشکل را نسبت به تحقق یک هدف، حل نماید. همچنین مشکلات در ماتریسی نسبت به نوآوری در راهکار، چهار حالت پیدا می‌کنند که یکی از آن حالات، حل مشکلات گذشته از طریق راهکارهای جدید است. تحقیق حاضر، مبتنی بر روش تحقیق "علم طراحی[1]" بنا شده است که در آن مشکلی از گذشته، یعنی پیش‌بینی عملکرد تحصیلی را به‌عنوان هدف قرار داده است و در راستای حل این مسئله از تکنیک‌های داده کاوی بهره برده و کارآمدی برخی از الگوریتم‌ها را با معیارهای کمی نشان داده است.

پیش از بررسی گام‌های تحقیق، سؤالاتی که از نظر هیونر[2] (۲۰۱۰) لازم است به آن‌ها در فرآیند پژوهش بر اساس علم طراحی در حوزه فناوری اطلاعات پاسخ داده شود، بررسی می‌نماییم. در روش تحقیق علم طراحی، برای شکل گرفتن مناسب یک فرآورده باید پاسخ به هشت پرسش داده شود که بر اساس تحقیق حاضر، پاسخ پرسش‌ها به شرح جدول ۲ خواهد بود:

1. Design Sience
2. Hevner

جدول ۲: سؤالات سنجش فرآورده تولیدشده در علم طراحی

شماره	سؤال	پاسخ بر اساس پژوهش حاضر
۱	سؤال تحقیق چیست؟	چطور می‌توان با استفاده از داده‌های گذشته دانش آموزان، عملکرد تحصیلی آینده آن‌ها را پیش‌بینی کرد و راهکارهایی برای بهبود عملکرد آنان ارائه نمود؟
۲	فرآورده مورد نظر چیست؟ در واقع فرآورده قرار است به چه چیزی در محیط پاسخگو باشد؟	سیستم تحلیل آینده‌نگری که راهکارهایی جهت بهبود عملکرد تحصیلی هرکدام از دانش آموزان ارائه نماید.
۳	فرآیند شکل‌گیری فرآورده چیست؟	تحلیل عملکرد گذشته دانش آموزان و خوشه‌بندی آن‌ها در دسته‌های مختلف تحصیلی و یافتن شباهت میان عملکرد دانش آموزان فعلی و دانش آموزان گذشته و در نهایت تشخیص زمینه‌های ممکن ارتقاء تحصیلی دانش آموزان فعلی
٤	چه دانش زمینه‌ای، فرآیند ساخت فرآورده را در بر گرفته است؟ چه تئوری، پشتوانه ساخت فرآورده است؟	۱- تکنیک‌ها و روش‌های خوشه‌بندی ۲- تکنیک‌ها و روش‌های آماری ۳- دانش تعلیم و تربیت
۵	چه معیارهای ارزیابی در طول شکل‌گیری فرآورده تعیین شده است؟ چه بهبودی در چرخه تکامل فرآورده شناسایی شده است؟	معیارهای ارزیابی فرآورده این تحقیق، شاخص اعتبارسنجی دان بوده است که فرآورده، امتیاز بالایی را در آن به دست آورده است.
٦	چطور فرآورده موردنظر به در محیط آزمایش می‌شود؟ چه شاخص‌های برای برآورد کارایی فرآورده در محیط تعبیه‌شده است؟	این فرآورده از طریق تحلیل داده‌های دانش آموزان سایر مراکز تحصیلی مورد آزمایش قرار می‌گیرد. شاخص کارایی این فرآورده، افزایش عملکرد تحصیلی از طریق عمل به توصیه‌های پایانی است.
۷	چه دانشی به دانش گذشته اضافه شده است؟	بهبود توصیه‌های تحصیلی با استفاده از داده کاوی مبتنی بر داده‌های گذشته
۸	آیا به‌صورت رضایت بخشی به سؤال تحقیق پرداخته‌شده است؟	سیستم طراحی‌شده امکان بیان راهکارهایی را دارد که پیش از طراحی سیستم وجود نداشته است و نمی‌توان با تکیه‌بر نظرات تجربی آن‌ها را احصاء نمود.

توصیف فرآورده

فرآورده موردنظر در این تحقیق، یک فرآیند داده کاوی است. این فرآیند، توصیه‌های تحصیلی به یادگیرندگان و مدرسان خواهد داشت که منجر به بهبود عملکرد تحصیلی و کاهش شکست و یا ترک تحصیل می‌گردد.

فرآورده موردنظر در این تحقیق، از دو بخش اصلی تشکیل شده است:

خوشه‌بندی: در بخش اول داده‌ها بر اساس سه دسته از متغیرها در خوشه‌های مختلف، خوشه‌بندی خواهند شد. این سه دسته عبارت‌اند از: "متغیرهای شخصیتی و خانوادگی "، "متغیرهای حمایتی و غیرمستقیم "، "متغیرهای رفتار تحصیلی ". در خوشه‌بندی داده از سه الگوریتم "K-means"، "K-medoids" و "K-fast" استفاده شده است.

ارائه توصیه‌ها: بر اساس خوشه‌بندی‌های انجام شده بر روی داده‌ها، ترکیب خوشه‌های هر سه دسته، ۱۶ حالت را ایجاد خواهد کرد که با در نظر گرفتن احتمال نتایج پایانی هرکدام از ترکیب خوشه‌ها و همچنین بررسی امکان انتقال از یک ترکیب خوشه به ترکیب خوشه دیگر، می‌توان توصیه‌هایی را برای هر یک از داده‌هایی که در هرکدام از ترکیب خوشه‌ها قرار می‌گیرند، ارائه داد.

فرآورده به دست آمده امکان تحلیل یادگیرندگان را بر اساس سه دسته متغیرها فراهم کرده است. دسته اول متغیرها، غیرقابل تغییر هستند اما دسته‌های دوم و سوم، امکان تغییر دارند. همچنین در میان دسته‌های دوم و سوم، دسته سوم که مربوط به رفتارهای تحصیلی می‌شود، از قابلیت تغییر بیشتری برخوردار است. بااین‌حال، این تحقیق نشان داده است که الزاماً تمرکز بر درس و تحصیل نمی‌تواند سبب بالاتر رفتن نتایج تحصیلی شود. بلکه هر ترکیب خوشه از یادگیرندگان نیازمند میزان مشخصی از تمرکز بر درس و تحصیل است. ضمن اینکه میزان تحصیلات خانوادگی تأثیر بسیار بالایی بر ضریب اثرگذاری تمرکز بر درس و تحصیل داشته است.

همچنین نتایج به دست آمده در این تحقیق، مرتبط با پژوهش‌های پیش‌بینی عملکرد یادگیرندگان و ارائه توصیه‌های کارآمد است و دانش این حوزه را در زمینه استفاده از

الگوریتم‌ها و تحلیل داده‌ها گسترش داده است.

تحلیل داده‌های تحصیلی

بررسی پایگاه داده

پایگاه داده مورد بررسی در این پژوهش از مجموعه پایگاه‌های داده معتبر دانشگاه کالیفرنیا استخراج شده است[1]. مشخصات و جزئیات پایگاه داده مورد بررسی در این پژوهش به شرح جدول ۳ است:

جدول ۳: توصیف پایگاه داده دانش آموزان

گروه	شماره ستون	نوع داده	انواع داده‌های موجود در هر ستون				
شخصیتی و خانوادگی	۱	سن	از ۱۵ تا ۲۲ سال				
	۲	جنسیت	زن	مرد	–	–	–
	۳	جمعیت خانواده	بیشتر از ۳ نفر	کمتر از ۳ نفر	–	–	–
	۴	تحصیلات مادر	بی‌سواد	ابتدایی	سیکل	دیپلم	دانشگاهی
	۵	تحصیلات پدر	بی‌سواد	ابتدایی	سیکل	دیپلم	دانشگاهی
	۶	وضعیت سرپرستی	غیر از پدر و مادر	مادر	پدر	–	–
	۷	گذراندن مهدکودک	خیر	بله	–	–	–
	۸	وضعیت طلاق	طلاق گرفته	زندگی مشترک	–	–	–
	۹	دلیل انتخاب مدرسه	سایر دلایل	تجدیدی سال گذشته	آدرس منزل	سطح علمی	–
	۱۰	محل زندگی	خارج از شهر	داخل شهر	–	–	–
	۱۱	شغل پدر	سایر مشاغل	بیکار	خدمات	پزشکی و سلامت	معلم
	۱۲	شغل مادر	سایر مشاغل	بیکار	خدمات	پزشکی	معلم

1. https://archive.ics.uci.edu/ml/datasets/Student+Performance

						و سلامت	
حمایتی و غیرمستقیم	۱۳	وضعیت سلامتی	بسیار نامطلوب	نامطلوب	متوسط	مطلوب	بسیار مطلوب
	۱۴	دسترسی به اینترنت	دسترسی	عدم دسترسی	–	–	–
	۱۵	کیفیت روابط خانوادگی	بسیار نامطلوب	نامطلوب	متوسط	مطلوب	بسیار مطلوب
	۱۶	فاصله زمانی منزل تا مدرسه	کمتر از ۱۵ دقیقه	بین ۱۵ تا ۳۰ دقیقه	بین ۳۰ تا ۶۰ دقیقه	بیشتر از ۶۰ دقیقه	–
	۱۷	حمایت مالی خانواده	عدم‌حمایت	حمایت	–	–	–
	۱۸	پرداخت هزینه‌های مدرسه	عدم پرداخت	پرداخت	–	–	–
	۱۹	تجدیدی در سال گذشته	هیچ درسی	۱ درس	۲ درس	۳ درس	–
	۲۰	حمایت مالی مدرسه	عدم‌حمایت	حمایت	–	–	–
رفتار تحصیلی	۲۱	زمان آزاد	بسیار کم	کم	متوسط	زیاد	بسیار زیاد
	۲۲	ساعات تفریح با دوستان	بسیار کم	کم	متوسط	زیاد	بسیار زیاد
	۲۳	ساعات فعالیت‌های فوق‌برنامه	بله	خیر	–	–	–
	۲۴	زمان مطالعه	کمتر از ۲ ساعت	بین ۲ تا ۵ ساعت	بین ۵ تا ۱۰ ساعت	بیشتر از ۱۰ ساعت	–
	۲۵	تعداد غیبت کلاسی	از ۱ تا ۷۵ جلسه				
نتایج تحصیلی	۲۶	نمره اولین آزمون	نمره‌ای از ۰ تا ۲۰				
	۲۷	نمره دومین آزمون	نمره‌ای از ۰ تا ۲۰				
	۲۸	نمره نهایی	نمره‌ای از ۰ تا ۲۰				

در این پایگاه داده، ۱۰۶۷ ردیف و ۲۸ ستون یا ویژگی موجود است. پایگاه داده فوق

پس از تبدیل ساختار فایل، وارد نرم‌افزار Rapid miner شده و آمار فراوانی و نمودارهای مرتبط با آن مورد بررسی قرار گرفته است.

پیش‌پردازش داده‌ها

در این قسمت داده‌ها را از حالت خام خارج کرده و طی فرآیندی تبدیل به داده‌هایی می‌نماییم که امکان انجام عملیات داده کاوی و مدل‌سازی بر روی آن وجود داشته باشد. از میان فرآیندهای مربوط به پیش‌پردازش داده‌ها، داده‌های نیازمند فرآیندهای «شناسایی داده‌های پرت»، «تبدیل»، «کاهش ابعاد» و «نرمال‌سازی» بوده‌اند. فرآیند حذف و تصمیم‌گیری درباره داده‌های گمشده به دلیل عدم وجود داده گمشده نیاز نبوده است.

داده‌های خارج از محدوده شناسایی شده‌اند. برای شناسایی داده‌های پرت، از الگوریتم مبتنی بر نزدیک‌ترین همسایگی استفاده شده است. برای این کار از الگوریتم Kmeans جهت شناسایی خوشه مربوط به داده‌های پرت استفاده شده است و تعداد داده‌های پرت مناسب در این روش بر اساس الگوریتم درخت گرادیانت[1]، ۷ داده است. RMSE در الگوریتم درخت گرادیانت برابر ۳/۴۱۸ و دقت پیش‌بینی الگوریتم Kmeans در شناسایی داده‌های پرت با تعداد ۷ نیز، ۰/۷۸ یا ۷۸٪ است. همچنین داده‌های کیفی از طریق توابع شرطی در نرم‌افزار اکسل به داده‌های عددی تبدیل شده است.

خوشه‌بندی

ایده اصلی مربوط به خوشه‌بندی، تقسیم کلی ویژگی‌های موجود در پایگاه داده به سه دسته "خانوادگی و شخصیتی "، "حمایتی و غیرمستقیم " و "رفتار تحصیلی " بوده است و نتایج مربوط به داده، در امتداد این سه ویژگی کلی ارزیابی می‌شود. هرکدام از سه دسته ویژگی فوق، می‌توانند در چند حالت خوشه‌بندی شوند و هر خوشه بیان‌گر شدت یا نوع خاصی از ویژگی‌های آن دسته باشد که در ادامه توضیح داده خواهد شد.

1. Gradient Booted Tree

نمایی از تعداد خوشه‌های مناسب برای هر کدام از دسته‌های سه‌گانه به همراه نرمال‌سازی و کاهش ابعاد را در جدول ۴ مشاهده می‌کنید:

جدول ٤: تعداد خوشه‌های مناسب برای هرکدام از ویژگی‌های تحصیلی به همراه نرمال‌سازی و کاهش ابعاد

	تعداد خوشه ها	بدون نرمال سازی			نرمال شده			نرمالسازی همراه با کاهش ابعاد		
		Kmeans	Kmedoids	K-fast	Kmeans	Kmedoids	K-fast	Kmeans	Kmedoids	K-fast
داده های خانوادگی و شخصیتی	۱	- ∞	- ∞	- ∞	- ∞	- ∞	- ∞	- ∞	- ∞	- ∞
	۲	۰.۱۷۴	۰.۲۲۱	۰.۱۷۴	۰.۲۵۳	۰.۱۹۹	۰.۲۵۲	۰.۲۷۷	۰.۲۲۳	۰.۲۷۶
	۳	۰.۱۷۶	۰.۲۰۴	۰.۱۷۶	۰.۲۱۳	۰.۲۰۰	۰.۲۱۳	۰.۲۳۲	۰.۲۱۴	۰.۲۳۳
	۴	۰.۱۷۲	۰.۲۴۸	۰.۱۷۴	۰.۲۰۰	۰.۲۳۲	۰.۲۰۰	۰.۲۱۸	۰.۲۰۷	۰.۲۱۸
	۵	۰.۱۹۱	۰.۲۴۴	۰.۱۷۸	۰.۲۰۳	۰.۲۱۸	۰.۱۹۴	۰.۲۱۹	۰.۲۲۷	۰.۲۰۹
	۶	۰.۱۸۶	۰.۲۲۸	۰.۱۸۷	۰.۱۹۸	۰.۲۱۷	۰.۱۹۵	۰.۲۱۱	۰.۴۳۲	۰.۲۱۰
	۷	۰.۱۹۸	۰.۲۶۲	۰.۱۸۷	۰.۲۰۸	۰.۳۳۴	۰.۱۹۷	۰.۲۰۹	۰.۳۷۴	۰.۲۰۵
	۸	۰.۱۹۴	۰.۲۳۰	۰.۱۸۶	۰.۱۸۴	۰.۳۱۹	۰.۱۹۵	۰.۲۲۷	۰.۴۲۰	۰.۱۹۷
	۹	۰.۲۰۰	۰.۲۳۲	۰.۱۹۷	۰.۲۰۷	۰.۳۱۸	۰.۱۹۵	۰.۲۱۴	۰.۳۳۷	۰.۲۰۷
	۱۰	۰.۲۰۱	۰.۲۱۶	۰.۲۰۶	۰.۲۰۷	۰.۳۰۴	۰.۲۰۳	۰.۲۱۸	۰.۳۵۱	۰.۲۰۰
داده های حمایتی و غیر مستقیم	۱	- ∞	- ∞	- ∞	- ∞	- ∞	- ∞	- ∞	- ∞	- ∞
	۲	۰.۱۵۹	۰.۱۶۲	۰.۱۵۹	۰.۲۵۹	۰.۱۸۳	۰.۲۰۸	۰.۲۳۴	۰.۱۹۸	۰.۲۶۲
	۳	۰.۱۸۳	۰.۱۷۷	۰.۲۲۵	۰.۲۱۶	۰.۲۲۳	۰.۲۱۸	۰.۱۹۲	۰.۲۳۷	۰.۲۳۸
	۴	۰.۱۹۶	۰.۱۸۶	۰.۲۰۸	۰.۱۸۸	۰.۲۱۰	۰.۱۹۳	۰.۲۳۳	۰.۲۲۱	۰.۲۰۵
	۵	۰.۱۹۵	۰.۲۰۸	۰.۲۰۶	۰.۱۷۶	۰.۲۱۳	۰.۱۷۶	۰.۱۹۲	۰.۲۲۳	۰.۱۹۲
	۶	۰.۱۸۵	۰.۲۰۶	۰.۱۸۹	۰.۱۷۷	۰.۱۹۹	۰.۱۸۴	۰.۱۹۱	۰.۲۲۰	۰.۱۸۷
	۷	۰.۱۷۳	۰.۱۹۱	۰.۱۸۴	۰.۱۸۰	۰.۲۳۵	۰.۱۷۹	۰.۱۷۷	۰.۲۴۴	۰.۱۸۰
	۸	۰.۱۷۰	۰.۲۱۶	۰.۱۷۳	۰.۱۵۸	۰.۲۴۹	۰.۱۶۹	۰.۱۷۲	۰.۲۲۹	۰.۱۷۳
	۹	۰.۱۸۱	۰.۲۲۹	۰.۱۸۴	۰.۱۷۳	۰.۲۷۱	۰.۱۸۰	۰.۱۶۸	۰.۲۱۲	۰.۱۹۸
	۱۰	۰.۱۸۰	۰.۲۰۰	۰.۱۹۵	۰.۱۸۷	۰.۲۲۷	۰.۱۸۱	۰.۲۱۵	۰.۲۰۸	۰.۱۹۲
داده های رفتار تحصیلی	۱	- ∞	- ∞	- ∞	- ∞	- ∞	- ∞	- ∞	- ∞	- ∞
	۲	۰.۱۲۳	۰.۴۱۹	۰.۱۲۳	۰.۳۷۳	۰.۳۲۰	۰.۳۷۳	۰.۴۲۰	۰.۴۱۳	۰.۴۲۰
	۳	۰.۱۴۶	۰.۲۲۶	۰.۱۴۶	۰.۳۲۶	۰.۲۹۵	۰.۳۰۳	۰.۳۶۶	۰.۳۳۳	۰.۳۳۹
	۴	۰.۱۶۲	۰.۲۲۷	۰.۱۶۲	۰.۲۷۹	۰.۴۰۶	۰.۲۷۸	۰.۳۱۳	۰.۳۳۹	۰.۳۰۵
	۵	۰.۱۷۴	۰.۲۲۵	۰.۱۷۴	۰.۲۷۹	۰.۳۷۵	۰.۲۷۱	۰.۳۰۵	۰.۳۶۶	۰.۲۹۶
	۶	۰.۲۱۶	۰.۳۲۲	۰.۱۹۳	۰.۲۴۹	۰.۳۷۷	۰.۲۵۷	۰.۲۸۳	۰.۴۱۰	۰.۲۸۲
	۷	۰.۲۴۶	۰.۲۹۶	۰.۲۲۳	۰.۲۵۳	۰.۳۴۰	۰.۲۵۴	۰.۲۷۲	۰.۴۶۳	۰.۲۸۰
	۸	۰.۲۳۷	۰.۲۹۰	۰.۲۳۶	۰.۲۵۰	۰.۳۹۳	۰.۲۴۹	۰.۲۷۵	۰.۴۲۰	۰.۲۷۳
	۹	۰.۲۶۱	۰.۲۸۴	۰.۲۵۵	۰.۲۶۱	۰.۳۸۶	۰.۲۴۹	۰.۲۶۶	۰.۴۴۳	۰.۲۶۹
	۱۰	۰.۲۴۳	۰.۳۱۸	۰.۲۴۲	۰.۲۴۴	۰.۴۷۹	۰.۲۴۸	۰.۲۶۲	۰.۴۱۱	۰.۲۷۰

نتایج به دست آمده از جدول ۴ را می‌توان در سه بند خلاصه کرد:

تأثیر نرمال‌سازی و کاهش ابعاد: نرمال‌سازی و کاهش ابعاد هر دو سبب افزایش پایه‌ای عدد دیویس شده‌اند و میزان خطا را به‌طورکلی افزایش می‌دهند. به همین دلیل به نظر می‌رسد خوشه‌بندی داده‌ها بدون نرمال‌سازی و کاهش ابعاد نتیجه بهتری داشته باشد.

حداکثر تعداد خوشه‌ها: ازآنجایی‌که تعدادهای خوشه‌های هرکدام از سه دسته فوق در یکدیگر ضرب خواهند شد و عدد به دست آمده در توصیه‌های مدیریتی استفاده می‌گردد، تعداد خوشه‌ها را نهایتاً تا ۵ عدد در نظر گرفته‌ایم تا در بیشترین حالت ۱۲۵ خوشه نهایی به دست آید و بتوان توصیه‌هایی را نسبت به آن‌ها در نظر گرفت. اگر چنانچه کمترین اعداد شاخص دیویس را به‌عنوان تعداد خوشه مطلوب در نظر بگیریم، برای داده‌های شخصیتی و خانوادگی، تعداد ۸ خوشه، برای داده‌های حمایتی و غیرمستقیم، تعداد ۹ خوشه و در داده‌های رفتار تحصیلی تعداد ۱۰ خوشه انتخاب خواهد شد که در نتیجه، ۷۲۰ حالت را ایجاد می‌کند. درحالی‌که تمام داده، ۱۰۳۸ داده است؛ بنابراین محدودیت حداکثر ۵ خوشه اعمال شده است. هرچند تعداد بیشتر خوشه‌ها نیز تفاوت چندان زیادی در شاخص دیویس را نشان نمی‌دهند.

الگوریتم‌های انتخاب‌شده: بنا بر جدول ۴ برای داده‌های شخصیتی و خانوادگی تعداد ۴ خوشه با الگوریتم Kmeans و برای داده‌های حمایتی و غیرمستقیم تعداد ۲ خوشه با الگوریتم Kmeans (اگرچه دو الگوریتم Kmeans و Kmeans fast هر دو یک عدد را در شاخص دیویس دارند اما به دلیل اینکه سرعت Kmeans حدوداً ۸ برابر بیشتر بوده است، این الگوریتم انتخاب شده است) و تعداد ۲ خوشه برای داده‌های رفتار تحصیلی با الگوریتم Kmeans (دلیل مانند مورد گذشته است، تعیین شده است.

جدول نهایی توصیف هرکدام از دسته‌های خوشه‌ها را می‌توان در جدول ۵ مشاهده کرد:

جدول ۵: نام‌گذاری خوشه‌های مربوط به هرکدام از دسته‌ها

خوشه‌ها	متغیرهای خانوادگی و شخصیتی	متغیرهای حمایتی و غیرمستقیم	متغیرهای رفتار تحصیلی
۰	یادگیرندگان دارای تحصیلات بالای خانوادگی	روابط خانوادگی مناسب و سلامت بالا	تمرکز نسبی بر درس و تحصیل
۱	یادگیرندگان دچار مشکلات خانوادگی و کم‌جمعیت	روابط خانوادگی نامناسب و سلامت پایین	تمرکز کامل بر درس و تحصیل
۲	یادگیرندگان دارای تحصیلات پائین خانوادگی		
۳	یادگیرندگان دارای پایه‌های ضعیف تحصیلی		

تحلیل یادگیرندگان و نتایج تحصیلی

ازآنجایی‌که قصد داریم تأثیر هرکدام از خوشه‌ها را در نتایج تحصیلی بررسی کنیم لازم است به نمرات پایانی نیز معنا و مفهومی بدهیم تا بتوان در تحلیل‌ها، توصیف دقیق‌تری داشت. بر همین اساس نمرات پایانی یادگیرندگان به چهار دسته تقسیم خواهد شد که به شرح زیر است:

یادگیرندگان نخبه: نمرات بالای ۱۳,۳۳

یادگیرندگان متوسط و خوب: نمرات بین ۱۱,۶۶ تا ۱۳,۳۳

یادگیرندگان نزدیک به تجدید: نمرات بین ۹,۶ تا ۱۱,۳۳

یادگیرندگان نزدیک به اخراج یا ترک تحصیل: نمرات کمتر از ۹,۳۳

این تقسیم‌بندی بر اساس توزیع داده‌های گذشته است. به‌نحوی‌که در هرکدام از این ۴ خوشه، تعداد داده‌ها تقریباً یکسان است. تعداد داده‌ها در هرکدام از محدوده‌ها فوق عبارت است از:

جدول ٦: نام‌گذاری بازه‌های نتایج پایانی به همراه تعداد داده‌ها

شماره محدوده	بازه نمرات محدوده	نام محدوده	تعداد داده‌ها
۱	کمتر از ۹,۳۳	نخبه و استعداد	۲۸۱
۲	۱۱,۶۶ تا ۱۳,۳۳	متوسط و خوب	۲۶۵
۳	۹,۶۶ تا ۱۱,۳۳	نزدیک به تجدید	۲۴۱
٤	بیشتر ۱۳,۳۳	نزدیک به اخراج یا ترک تحصیل	۲۵۱

اگر هر کدام از یادگیرندگان را ازلحاظ خوشه‌های خانوادگی و حمایتی و رفتار تحصیلی در برابر نتایج تحصیلی قرار دهیم، ۱۶ حالت ایجاد می‌شود (جایگشت ۴ خوشه برای دسته اول، ۲ خوشه برای دسته دوم و ۲ خوشه نیز برای دسته سوم.) که به‌صورت جدول زیر می‌باشند:

جدول ۷: احتمال موفقیت و شکست تحصیلی هر کدام از ترکیب خوشه‌های تعیین شده

توصیف نتیجه تحصیلی	موفقیت تحصیلی	شکست تحصیلی	رفتار تحصیلی	حمایتی و غیرمستقیم	خانوادگی و شخصیتی	حالت‌های ترکیب خوشه‌ها
نزدیک به تجدید	47%	53%	0	0	0	1
نخبه و استعداد	68%	32%	1	0	0	2
یادگیرنده نمونه	54%	46%	0	1	0	3
نخبه و استعداد	60%	40%	1	1	0	4
نخبه و استعداد	54%	46%	0	0	1	5
نزدیک به تجدید	45%	55%	1	0	1	6
یادگیرنده نمونه	50%	50%	0	1	1	7
نزدیک به تجدید	45%	55%	1	1	1	8
نزدیک به تجدید	24%	76%	0	0	2	9
نزدیک به ترک تحصیل یا اخراج	43%	57%	1	0	2	10
نزدیک به ترک تحصیل یا اخراج	34%	66%	0	1	2	11
یادگیرنده نمونه	51%	49%	1	1	2	12
نزدیک به ترک تحصیل یا اخراج	33%	67%	0	0	3	13
نزدیک به تجدید	47%	53%	1	0	3	14
نزدیک به ترک تحصیل یا اخراج	40%	60%	0	1	3	15
نزدیک به تجدید	47%	53%	1	1	3	16

اگر به‌جای شماره خوشه، نام خوشه‌ها گذاشته شود و وضعیت نهایی در مقابل آن‌ها قرار گیرد جدولی به شکل زیر خواهیم داشت:

جدول ۸: توصیف وضعیت موفقیت یا شکست تحصیلی هرکدام از ترکیب خوشه‌های تعیین شده

نتیجه تحصیلی احتمالی	رفتار تحصیلی	حمایتی و غیرمستقیم	خانوادگی و شخصیتی	حالت‌های ترکیب خوشه‌ها
شکست: تجدید	تمرکز نسبی بر درس و تحصیل	روابط خانوادگی مناسب و سلامت بالا	تحصیلات بالای خانوادگی	**1**
موفقیت: نخبه	تمرکز کامل بر درس و تحصیل	روابط خانوادگی مناسب و سلامت بالا	تحصیلات بالای خانوادگی	**2**
موفقیت: نمونه	تمرکز نسبی بر درس و تحصیل	روابط خانوادگی نامناسب و سلامت پایین	تحصیلات بالای خانوادگی	**3**
موفقیت: نخبه	تمرکز کامل بر درس و تحصیل	روابط خانوادگی نامناسب و سلامت پایین	تحصیلات بالای خانوادگی	**4**
موفقیت: نخبه	تمرکز نسبی بر درس و تحصیل	روابط خانوادگی مناسب و سلامت بالا	مشکلات خانوادگی و جمعیت کم خانواده	**5**
شکست: تجدید	تمرکز کامل بر درس و تحصیل	روابط خانوادگی مناسب و سلامت بالا	مشکلات خانوادگی و جمعیت کم خانواده	**6**
موفقیت: نمونه	تمرکز نسبی بر درس و تحصیل	روابط خانوادگی نامناسب و سلامت پایین	مشکلات خانوادگی و جمعیت کم خانواده	**7**
شکست: تجدید	تمرکز کامل بر درس و تحصیل	روابط خانوادگی نامناسب و سلامت پایین	مشکلات خانوادگی و جمعیت کم خانواده	**8**
شکست: تجدید	تمرکز نسبی بر درس و تحصیل	روابط خانوادگی مناسب و سلامت بالا	تحصیلات پایین خانوادگی	**9**
شکست: اخراج یا ترک تحصیل	تمرکز کامل بر درس و تحصیل	روابط خانوادگی مناسب و سلامت بالا	تحصیلات پایین خانوادگی	**10**
شکست: اخراج یا ترک تحصیل	تمرکز نسبی بر درس و تحصیل	روابط خانوادگی نامناسب و سلامت پایین	تحصیلات پایین خانوادگی	**11**
موفقیت: نمونه	تمرکز کامل بر درس و تحصیل	روابط خانوادگی نامناسب و سلامت پایین	تحصیلات پایین خانوادگی	**12**
شکست: اخراج یا ترک تحصیل	تمرکز نسبی بر درس و تحصیل	روابط خانوادگی مناسب و سلامت بالا	پایه تحصیلی ضعیف یادگیرنده	**13**
شکست: تجدید	تمرکز کامل بر درس و تحصیل	روابط خانوادگی مناسب و سلامت بالا	پایه تحصیلی ضعیف یادگیرنده	**14**
شکست: اخراج یا ترک تحصیل	تمرکز نسبی بر درس و تحصیل	روابط خانوادگی نامناسب و سلامت پایین	پایه تحصیلی ضعیف یادگیرنده	**15**
شکست: تجدید	تمرکز کامل بر درس و تحصیل	روابط خانوادگی نامناسب و سلامت پایین	پایه تحصیلی ضعیف یادگیرنده	**16**

تحلیل یافته‌ها

در جدول زیر خلاصه‌ای از شرح هرکدام از حالت‌های ترکیب خوشه‌ها آورده شده است:

جدول ۹: تحلیل یافته‌های مربوط به توصیف وضعیت موفقیت یا شکست هرکدام از ترکیب خوشه‌های تعیین شده

شماره حالت ترکیب خوشه‌ها	تحلیل نتیجه
۱	تمرکز نسبی بر تحصیل سبب بی‌اثر شدن تحصیلات خانوادگی و شرایط مناسب حمایت تحصیلی شده است و احتمال شکست تحصیلی در این حالت از ترکیب خوشه‌ها وجود دارد.
۲	تمرکز کامل بر درس و تحصیل با شرایط مناسب خانوادگی و سلامتی همراه شده است و احتمال نخبه شدن و شکوفایی تحصیلی در این حالت از ترکیب خوشه‌ها بسیار بالاست.
۳	علی‌رغم شرایط نامناسب روابط خانوادگی و سلامتی و همچنین تمرکز نسبی بر تحصیل، در اغلب موارد موفقیت تحصیلی ثبت شده است. احتمال دارد تأثیر بالای تحصیلات خانوادگی و استعداد یادگیرنده سبب بازده بالای تحصیلی شده باشد و یادگیرنده با صرف زمان کمتر، نتیجه بهتری کسب شده باشد.
٤	اگرچه وضعیت روابط خانوادگی و سلامتی در این داده‌ها مناسب نبوده است، اما تمرکز کامل یادگیرنده بر تحصیل و سطح تحصیلی بالای خانوادگی سبب موفقیت چشمگیر تحصیلی شده است.
٥	یادگیرندگانی که مشکلات خانوادگی دارند، در صورت داشتن روابط خانوادگی مناسب با والدین و برخورداری از سلامتی، می‌توانند نتایج بسیار بالای تحصیلی را کسب نمایند. ضمن اینکه تمرکز نسبی بر درس و تحصیل الزاماً به معنای رها کردن درس و تحصیل نیست. چه‌بسا پرداختن متعادل به درس سبب جبران آسیب‌های ناشی از مشکلات خانوادگی شده باشد و بازده تحصیلی را افزایش داده باشد. نمودار ساعات تفریح این دسته از یادگیرندگان بر اساس نتایج تحصیلی نیز نشان می‌دهد که یادگیرندگانی با ساعات تفریح متعادل، احتمال بیشتری برای کسب نتایج بالای تحصیلی داشته‌اند.

٦	همان‌طور که در حالت گذشته نیز اشاره شد، تمرکز کامل بر تحصیل در یادگیرندگانی که مشکلات خانوادگی دارند، تأثیر معکوس دارد و تمرکز نسبی بر تحصیل می‌تواند نتایج بهتری داشته باشد.
۷	این دو حالت نیز مانند دو حالت گذشته، نشان‌دهنده اثرگذاری تمرکز نسبی بر درس و تحصیل در بهبود نتایج تحصیلی برای یادگیرندگانی که دارای مشکلات خانوادگی هستند، است.
۸	
۹	تمرکز نسبی بر درس و تحصیل به همراه سطح پایین تحصیلات خانوادگی، تأثیر بسیار بالایی بر کاهش نتایج تحصیلی داشته و یادگیرندگانی که در این حالت ترکیب خوشه هستند، احتمال زیادی دارد که دچار شکست تحصیلی شوند.
۱۰	به نظر می‌رسد تأثیر سطح پایین تحصیلات خانوادگی بر نتایج تحصیلی بیشتر از تمرکز کامل یادگیرنده بر تحصیل و کیفیت بالای روابط خانوادگی است. تا اندازه‌ای که با احتمال نسبتاً بالایی، یادگیرندگان این ترکیب خوشه، دچار ترک تحصیل یا اخراج می‌شوند.
۱۱	جمع شدن سه عامل سطح تحصیلات پایین خانواده، وضعیت ارتباطی نامناسب و عدم تمرکز بر درس و تحصیل، سبب شده است بدترین حالت ممکن یعنی اخراج یا ترک تحصیل بیشترین احتمال را برای نتیجه تحصیلی داشته باشد.
۱۲	این حالت از ترکیب خوشه‌ها، تنها حالتی است که نشان‌دهنده موفقیت یادگیرندگانی با تحصیلات پایین خانوادگی را نشان می‌دهد. در واقع این حالت نشان می‌دهد امکان موفقیت تحصیلی در میان یادگیرندگانی که خانواده‌هایشان از سطح تحصیلات پایینی برخوردار هستند و روابط خانوادگی نیز مطلوب نیست، وجود دارد. هرچند با توجه به درصد موفقیت این ترکیب خوشه، احتمال موفقیت یک درصد بیشتر از شکست است.
۱۳ تا ۱٦	ترکیب‌های خوشه‌ای که متغیرهای شخصیتی و خانوادگی که پایه‌های ضعیف تحصیلی دارند و از جمعیت خانوادگی اندکی برخوردار هستند، به‌طورکلی نتایج نامطلوبی در تحصیل داشته‌اند. به نظر می‌رسد در این خوشه، روابط خانوادگی و سلامتی چندان بر موفقیت تحصیلی اثرگذار نباشد و عامل مهم، تمرکز کامل یادگیرندگان بر درس و تحصیل باشد. هرچند این تمرکز نیز الزاماً نمی‌تواند سبب موفقیت تحصیلی شود اما تا ۱۴ درصد می‌تواند احتمال شکست را کاهش دهد. در واقع داده‌های این ترکیب خوشه‌ها، نه متغیرهایی مؤثر بر افزایش عملکرد تحصیلی را دارند و نه متغیرهای مؤثر بر کاهش عملکرد تحصیلی را و احتمال دارد عدم وجود عامل اثرگذار بر تحصیل سبب از بین رفتن انگیزه‌های تحصیلی شده و نتیجه را همواره نزدیک به شکست رسانده باشد.

بنا بر جدول فوق می‌توان عامل موفقیت تحصـیلی در هر کدام از ترکیب‌های خوشـه‌ها را به‌صورت طبقه‌بندی کرد:

جدول ۱۰: بررسی عامل موفقیت هرکدام از ترکیب خوشه‌های تعیین شده

حالت‌های ترکیب خوشه‌ها	شرح عامل موفقیت
۱ تا ٤	تمرکز کامل بر درس و تحصیل
٥ تا ۸	تمرکز نسبی بر درس و تحصیل
۹ تا ۱۲	برای حالت‌های ۹ و ۱۰: تمرکز نسبی بر درس و تحصیل برای حالت‌های ۱۱ و ۱۲: تمرکز کامل بر درس و تحصیل
۱۳ تا ١٦	تمرکز کامل بر درس و تحصیل

همچنین می‌توان مقایسه‌ای میان متغیرهای حمایتی و غیرمستقیم به‌صورت جدول ۱۱ انجام داد:

جدول ۱۱: مقایسه تأثیر متغیرهای حمایتی و غیرمستقیم در حالت‌های مختلف ترکیب خوشه‌های تعیین شده

شماره حالت ترکیب خوشه‌ها	تحلیل نتیجه
۱ و ۳	فرزندان خانواده‌هایی که روابط مناسبی دارند و از سلامتی بالاتری نیز برخوردار هستند، موفقیت تحصیلی کمتری به دست آورده‌اند.
۲ و ٤	عدم ارتباط معنادار.
٥ و ۷	فرزندان خانواده‌هایی که وضعیت ارتباطی و سلامتی مناسب‌تری دارند، نتایج تحصیلی بهتری را نیز کسب کرده‌اند.
٦ و ۸	عدم ارتباط معنادار.
۹ و ۱۱	فرزندان خانواده‌هایی که وضعیت ارتباطی و سلامتی مناسب‌تری دارند، نتایج تحصیلی بهتری را نیز کسب کرده‌اند.
۱۰ و ۱۲	فرزندان خانواده‌هایی که روابط مناسبی دارند و از سلامتی بالاتری نیز برخوردار هستند، موفقیت تحصیلی کمتری به دست آورده‌اند.
۱۳ و ۱٥	عدم ارتباط معنادار.
١٤ و ١٦	عدم ارتباط معنادار.

همان‌طور که ملاحظه می‌شود، در دو حالت از حالت‌های مقایسه‌ای جدول ۴-۴۵، فرزندان خانواده‌هایی که از روابط خانوادگی باکیفیت‌تر و سلامتی بیشتر برخوردار هستند موفقیت تحصیلی بیشتر داشته‌اند و در دو حالت موفقیت تحصیلی کمتری را به خود اختصاص داده‌اند. در سایر حالات نیز ارتباط معناداری دیده نشده است. این نشان می‌دهد عضویت یادگیرندگان در خوشه‌هایی باکیفیت روابط خانوادگی و سلامتی بالا به‌طورکلی تأثیر محسوسی در کاهش یا افزایش موفقیت تحصیلی ندارد.

نتیجه‌گیری و ارائه پیشنهادها

نتایج حاصل از پژوهش بر اساس یادگیرندگان در چند بخش است که شرح آن عبارت است از:

شرایط خانوادگی یادگیرندگان به لحاظ تحصیلات، کیفیت روابط و جمعیت، عواملی هستند که به‌صورت زمینه‌ای بر روی بازدهی تحصیل یادگیرندگان تأثیر می‌گذارند و نمی‌توان آن‌ها را تغییر داد. بلکه لازم است مدیران و مدرسان، این عوامل را در نظر بگیرند و برای موفقیت هرکدام از یادگیرندگان، راهبردهایی را در نظر داشته باشند. به طور کلی یادگیرندگان در چهار دسته تقسیم شده و توصیه‌های کلی در رابطه با هرکدام از آن‌ها عبارت است از:

یادگیرندگانی که تحصیلات خانوادگی بالایی دارند.

به نظر می‌رسد در این ترکیب خوشه از یادگیرندگان، تمرکز نسبی بر درس و تحصیل، می‌تواند نتایج تحصیلی را کاهش دهد؛ اما به‌طورکلی این دسته از یادگیرندگان، در سطح بالای تحصیلی قرار دارند.

یادگیرندگانی که تحصیلات خانوادگی پایینی دارند.

به‌طورکلی احتمال بالایی وجود دارد که این دسته از یادگیرندگان دچار شکست تحصیلی یا ترک تحصیل شوند. به همین دلیل لازم است مدرسان از یادگیرندگانی که در این دسته قرار می‌گیرند، مراقبت بیشتری به عمل آورند.

یادگیرندگانی که شرایط خانوادگی نامناسبی دارند.

اگرچه فرزندان طلاق به لحاظ خانوادگی دارای خلأ جدی هستند، اما بر اساس تجربه گذشته، درصورتی‌که این دسته بر روی تحصیل تمرکز داشته باشند، می‌توانند نتایج تحصیلی بالایی کسب نمایند.

یاد گیرندگانی پایه تحصیلی ضعیفی دارند و از جمعیت اندک خانوادگی برخوردار هستند.

در این گروه، یادگیرندگان به‌نوعی تحت تأثیر عامل مثبت یا منفی به خصوصی قرار نمی‌گیرند و به همین جهت احتمال شکست و ترک تحصیل به‌صورت بالقوه بالا است. این گروه نیز درصورتی‌که موردتوجه واقع نشوند، نتایج مناسبی کسب نخواهند کرد.

همچنین پیشنهادهای پژوهشی عبارت‌اند از:

مقایسه الگوریتم‌های بکار گرفته‌شده در طبقه‌بندی

الگوریتم‌های بکار گرفته‌شده در این پژوهش را می‌توان به لحاظ برنامه‌نویسی بررسی کرد و آن‌ها را به‌دقت بالاتری رساند.

توسعه یک سیستم خبره جهت خوشه‌بندی و پیش‌بینی نتایج تحصیلی و ارائه توصیه‌های آموزشی

تمام مسیر طی شده در این پژوهش را می‌توان به‌صورت یک سیستم خبره طراحی کرد. به این صورت که اطلاعات یادگیرندگان گرفته شود و سپس سیستم، با استفاده از خوشه‌بندی و طبقه‌بندی، نتایج پیش‌بینی کرده و یادگیرنده را در گروه خود قرار دهد و نهایتاً بسته توصیه‌های مربوطه را ارائه نماید.

پیش‌بینی نتایج تحصیلی بر اساس توازن و عدم توازن نتایج در محدوده‌های مختلف

نتایج به دست آمده در قسمت ترکیب خوشه‌ها می‌تواند از طریق به‌صورت احتمال فاصله گرفتن نتایج از میانگین مورد بررسی قرار گیرد. با این تفسیر که هرکدام از ترکیب خوشه‌ها اگر احتمال برابری در کسب نتایج پایین یا بالای تحصیلی داشته باشند، توازن کامل برقرار خواهد بود. هرچقدر نتایج تحصیلی با سمت بالا یا پایین باشد، عدم توازن ایجاد شده است و لازم است نسبت به آن حالت از ترکیب خوشه‌ها، تصمیم‌گیری آموزشی لازم به عمل آید.

بررسی اثرگذاری متغیرهای حذف‌شده در ترکیب خوشه‌ها

برخی از متغیرهای نظیر شغل والدین یا محل زندگی از متغیرهای مدل‌سازی حذف شده‌اند. می‌توان تأثیر آن‌ها را بر ترکیب خوشه‌ها مورد بررسی قرار داد.

به‌کارگیری دیگر الگوریتم‌های خوشه‌بندی و اعتبارسنجی آن‌ها به‌وسیله سایر روش‌ها

هرچند احتمال اندکی وجود دارد که نتایج متفاوتی با اعمال روش‌های دیگر برای محقق حاصل شود اما می‌توان به‌عنوان کاری تمرینی در داده کاوی این پیشنهاد را مورد بررسی قرار داد.

پیشنهادهای کاربردی عبارت‌اند از:

اجرای ایده پژوهش حاضر در بستر مدارس ایران

این پژوهش بر اساس یک پایگاه داده بین‌المللی تنظیم شده است. می‌توان روش پیشنهادشده در این تحقیق را در یکی از مدارس داخل کشور پیاده‌سازی کرده و نتایج آن را در شرایط بومی ایران بررسی کرد.

شخصی‌سازی آموزش با استفاده از ایده پیش‌بینی تحصیلی مبتنی بر خوشه‌بندی

روش پیشنهادشده در این پژوهش می‌تواند فراتر از صرفاً پیش‌بینی نتایج تحصیلی، روش تحصیل یا مسیر آموزش مناسب برای هرکدام از دانش آموزان را از طریق خوشه‌بندی روش‌های تحصیل (نظیر آنلاین، غیرحضوری، تقویتی، آموزش دانش‌آموزی، رقابتی و سایر روش‌ها) پیش‌بینی نماید. به این صورت که پس از ارزیابی میزان موفقیت تحصیلی هرکدام از دانش آموزان در هر روش تحصیل، خوشه‌هایی از دانش آموزان موفق و نیمه موفق و همچنین خوشه‌هایی از دانش آموزان ناموفق یا تقریباً ناموفق قرار گیرند و پس از اتمام فرآیند یادگیری ماشینی، برآورد شود که یک دانش‌آموز جدید در کدام‌یک از خوشه‌های تحصیلی قرار می‌گیرد تا بر اساس شرایط او، روش تحصیل متناسب پیشنهاد شود.

تحلیل رفتار تحصیلی دانشجویان در دانشگاه‌های ایران بر اساس ایده محوری پژوهش

داده‌های سه‌گانه دانشجویان از قبیل داده‌های شخصیتی، داده‌های رفتار تحصیلی (انتخاب واحد و سایر داده‌ها) و داده‌های نتایج پایانی ترم‌های تحصیلی، در کنار یکدیگر می‌توانند نشان‌دهنده خوشه‌های رفتار موفق، رفتار نیمه موفق، رفتار تقریباً ناموفق و رفتار ناموفق باشند. می‌توان بر اساس داده‌های دانشجویان در ابتدای ورود به دانشگاه، مسیر احتمالی موفقیت را به آن‌ها پیشنهاد کرد. البته رسیدن به الگوی موفقیت برای دانشجویان یک دانشگاه به عوامل محیطی بسیاری بستگی دارد و بدون دسترسی به تعداد قابل‌توجهی از متغیرها و داده‌های تحصیلی، آموزش مستمر الگوریتم‌های یادگیری ماشینی و شناسایی الگوهای در حال تغییر، نسبتاً دشوار خواهد بود.

منابع

دی پیر، محمود و رابو، احمد. (۱۳۹۷). استفاده از داده کاوی آموزشی جهت گروه‌بندی یادگیرندگان در محیط یادگیری الکترونیکی به‌منظور شخصی‌سازی برنامه آموزش. *مدیریت و برنامه‌ریزی در نظام‌های آموزشی*، ۱۱(۱).۱۰۸-۸۳.

رضائی، عباسعلی و زاهدی، محمدهادی. (۱۳۹۷). نقش فن‌آوری‌های نوین در پیشرفت آموزش‌های الکترونیکی (با نگاهی به فرصت‌ها و چالش‌های پیش رو در دانشگاه‌ها و مراکز آموزش عالی داخل). *نشریه پژوهش در نظام‌های آموزشی*،۱۲(۴۰).۲۲۴-۲۰۷.

امین بیدختی علی‌اکبر، فتحیان بروجنی، محمد و نامنی، احمد. (۱۳۹۶). مدلی برای پیش‌بینی آسیب‌پذیری تحصیلی مقطع کارشناسی مبتنی بر شبکه عصبی. *مدیریت و برنامه ریزی در نظام های آموزشی*، ۱۰(۱)، ۱۰۲-۸۱.

زاهدبابلان، ع.؛ معینی کیا، م؛ و درخشانفرد، س. (۱۳۹۵). نقش آموزش الکترونیکی در نظام آموزش عالی و چالش‌های پیش روی آن، *اولین کنفرانس بین‌المللی پژوهش‌های نوین در حوزه علوم تربیتی و روانشناسی و مطالعات اجتماعی ایران*، قم، دبیرخانه دائمی کنفرانس، موسسه بین‌المللی مطالعات و توسعه علم خاورمیانه.

امیر تیموری، محمدحسن و زارع، محمد. (۱۳۹۴). *بارشناختی و چندرسانه‌ای آموزشی*، دانشگاه علامه طباطبایی.

مقصودی بهروز؛ سلیمانی، صادق؛ امیری، علی و افشارچی، محسن. (۱۳۹۱). ارتقای کیفیت آموزش در سامانه‌های آموزش الکترونیکی با استفاده از داده کاوی آموزشی. *فناوری آموزش*، ۶(۴)، ۲۸۶-۲۲۷.

نریمی سایی، ژاله و شادگار، بیتا. (۱۳۹۰). کاربرد تکنیک‌های داده‌کاوی در محیط‌های آموزش الکترونیکی .*مطالعات کتابداری و علم اطلاعات.۱۸(۷)، ۳-۱.*

Aldowah, H., Al-Samarraie, H., & Fauzy, W. M. (2019). Educational Data Mining and Learning Analytics for 21stcentury higher education: A Review and Synthesis. *Telematics and Informatics*. 37, 13-49.

Hernández-Lara, A. B., Perera-Lluna, A., & Serradell-López, E. (2019). Applying learning analytics to students' interaction in business

simulation games. The usefulness of learning analytics to know what students really learn. *Computers in Human Behavior*, 92, 600-612.

Nistor, N., & Hernández-Garcíac, Á. (2018). What types of data are used in learning analytics? An overview of six cases. *Computers in Human Behavior*, 89, 335-338.

Schumacher, C., & Ifenthaler, D. (2018). Features students really expect from learning analytics. *Computers in Human Behavior*, 78, 397-407.

Vieira, C., Parsons, P., & Byrd, V. (2018). Visual learning analytics of educational data: A systematic literature review and research agenda. *Computers & Education*, 122, 119-135.

Asif, R., Merceron, A., Ali, S. A., & Haider, N. G. (2017). Analyzing undergraduate students' performance using educational data mining. *Computers & Education*, 113, 177-194.

Hamsa, H., Indiradevi, S., & Kizhakkethottam, J. J. (2016). Student academic performance prediction model using decision tree and fuzzy genetic algorithm. *Procedia Technology*, 25, 326-332.

Marbouti, F., Diefes-Dux, H. A., & Madhavan, K. (2016). Models for early prediction of at-risk students in a course using standards-based grading. *Computers & Education*, 103, 1-15.

Shahiri, A.M., & Husain, W. (2015). A review on predicting student's performance using data mining techniques. *Procedia Computer Science*, 72, 414-422.

Brown, S. J., White, S., & Power, N. (2015). Tracking undergraduate student achievement in a first-year physiology course using a cluster analysis approach. *Advances in physiology education*, 39(4), 278-282.

Cambruzzi, W. L., Rigo, S. J., & Barbosa, J. L. (2015). Dropout Prediction and Reduction in Distance Education Courses with the Learning Analytics Multitrail Approach. *J. UCS*, 21(1), 23-47.

Caputi, V., & Garrido, A. (2015). Student-oriented planning of e-learning contents for Moodle. *Journal of Network and Computer Applications*, 53, 115-127.

Campagni, R., Merlini, D., & Verri, M. C. (2014, April). An Analysis of Courses Evaluation Through Clustering. *In International Conference on Computer Supported Education* (pp. 211-224).

He, W. (2013). Examining students' online interaction in a live video streaming environment using data mining and text mining. *Computers in Human Behavior*, 29(1), 90-102.

Van Barneveld, A., Arnold, K. E., & Campbell, J. P. (2012). Analytics in higher education: Establishing a common language. *EDUCAUSE learning initiative*, 1(1), l-ll.

Ferguson, R., & Shum, S. B. (2012, April). Social learning analytics: five approaches. *In Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 23-33). ACM.

Hevner, A., & Chatterjee, S. (2010). Design science research in information systems. *In Design research in information systems* (pp. 9-22). Springer, Boston, MA.

Romero, C., & Ventura, S. (2010). Educational data mining: a review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 40(6), 601-618.

Chen, C. M., & Chen, M. C. (2009). Mobile formative assessment tool based on data mining techniques for supporting web-based learning. *Computers & Education*, 52(1), 256-273.

Romero, C., Ventura, S., & García, E. (2008). Data mining in course management systems: Moodle case study and tutorial. *Computers & Education*, 51(1), 368-384.

Romero, C., & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert systems with applications*, 33(1), 135-146.

Retalis, S., Papasalouros, A., Psaromiligkos, Y., Siscos, S., & Kargidis, T. (2006). Towards networked learning analytics–A concept and a tool. *In Proceedings of the fifth international conference on networked learning*, (pp. 1-8).