

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 دانش‌پذیران 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 دانش‌پذیران و ریشه‌های عملکرد مؤثر در هر بخش است و روش نام‌گذاری ارائه‌شده، روشی نوین و قابل استفاده 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.

** دانشیار، 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 پذیرش: ۱۳۹۹/۰۳/۲۵

تاریخ دریافت: ۱۳۹۸/۱۱/۱۲

مقدمه

در هر کشوری، آموزش از ارکان مهم رشد و توسعه آن کشور به حساب می آید که اگرچه آموزش در تمدن‌ها و کشورهای دوران گذشته، شکلی ساده داشته است اما در دوران کنونی، با ظهور فناوری‌ها 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 دانش‌آموزان جایگاه ویژه‌ای در میان مقالات و تحقیقات بین‌المللی داشته است (کانسپسیون برگوس، ۲۰۱۸؛ آسیف و همکاران^۱، ۲۰۱۷؛ مربوطی و همکاران^۲، ۲۰۱۶؛ شهیری و همکاران^۳، 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 به صورت 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 قابل ارائه در جهت بهبود عملکرد آموزشی.

در دهه اخیر، تصمیم‌گیری مدیران به صورت 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، نایویز، 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 به صورت خاص و همچنین انگیزش یادگیرنده و خروج از یادگیری از جمله مفاهیمی هستند. به عنوان 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، 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.

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 فن آوری را به مثابه یک مزیت رقابتی در اولویت قرار دهند. از طرف دیگر نیاز جوامع در حال توسعه به استفاده بهینه از زمان و امکانات و توسعه انعطاف پذیر آموزش، منجر به گرایش بیشتر به استفاده از آموزش الکترونیکی شده است (ملک محمدی و شیروانی، ۱۳۹۳). یکی از روش های آموزشی که در آن فراگیر بر اساس توانایی های خود به فعالیت و یادگیری می پردازد روش نوین آموزش الکترونیکی و نقش فناوری های نوین در این حوزه است (زاهد بابلان، معینی کیا و درخشان فرد، ۱۳۹۵) این نوع آموزش جستجوی اطلاعات را در دنیای فن آور امروز متحول کرده و با وجود چالش هایی از قبیل تقاضای روزافزون برای آموزش عالی و عدم کفایت بودجه، کمبود مدرس تمام وقت و نیاز به حذف محدودیت های جغرافیایی مورد توجه جدی قرار گرفته است (رضایی و زاهدی، ۱۳۹۷).

انعطاف پذیری آموزش الکترونیک، مدیریت محتوا، تعامل هم زمان استاد و دانشجویان و همچنین سازمان دهی ساختار دوره ها 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 گوشی های هوشمند مفهوم آموزش را به طور کلی تغییر داده است. امروزه ارائه دهندگان خدمات الکترونیکی آموزش آنلاین از ارائه راهکارهای یادگیری از طریق تلفن همراه سود بسیار زیاد و قابل توجهی را به دست می آورند. دلایل و مزایای استفاده از گوشی های همراه برای یادگیری از دیدگاه دانش آموزان به شرح زیر است:

- محتواهای آموزشی را می توان "در حال حرکت" بدون نیاز به برنامه ریزی و تنظیم زمان دقیقی برای یک درس مطالعه کرد.
- یک روش یادگیری به شکل گروهی است. جامعه دانش اندوزان به اندازه کافی برای کمک به حل مسائل چالشی و تقویت بحث های سازنده میان اعضا بزرگ است.

- دانش آموزان همواره به اطلاعات موردنیاز خود دسترسی دارند و در مواقع لزوم به راحتی می‌توانند از آن‌ها استفاده کنند.
 - محتویات موبایل به صورت فشرده در اختیار دانش آموزان قرار می‌گیرد بنابراین آن‌ها را در مدت زمان 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 به کاررفته 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 مورد استفاده در آموزش، تکنیک‌های داده‌کاوی در دهه اخیر مورد استقبال قابل توجهی قرار گرفته است. مطمئناً تکنیک‌های داده‌کاوی می‌توانند در اختیار سیاست‌گذاران آموزش و پرورش 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، 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 به‌دست آمده از فرآیند آموزشی را جمع‌آوری کرده و پس از تجزیه و تحلیل، گزارشی از آن‌ها ارائه نمایند که در جهت بهبود عملکرد تحصیلی کارآمد باشد.
- داده‌کاوی آموزشی و تجزیه و تحلیل پیشرفته یادگیری این امکان را به ما می‌دهد که مدل‌های گذشته آموزش و یادگیری را دوباره‌سازی نماییم و راهکارهای جدیدی برای مسائل

گذشته داشته باشیم. (پراون و همکاران^۱، 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 به صورت شخصی سازی شده 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^۲، ۲۰۱۹). به همین جهت یادگیرندگان به خصوص دانشجویان، انتظار دارند تحلیل‌های پیشرفته یادگیری بتواند از برنامه‌ریزی درسی آن‌ها پشتیبانی کند، خودارزیابی را تقویت نماید، توصیه‌هایی که قابلیت شخصی سازی داشته باشند، ارائه شوند و تحلیل‌های یادگیری نیز به صورت شخصی سازی شده ارائه گردند (شوماخر و افتتالر^۳، 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 ذخیره شده 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؛ پرا 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^۱، ۲۰۱۴؛ جیانگ و همکاران، ۲۰۱۶)، عوامل مرتبط با موفقیت، شکست و نیت ترک تحصیل دانشجویان (کمبروزی و همکاران^۲، ۲۰۱۵؛ لیکورنتزو و همکاران، ۲۰۰۹؛ مارکزور و همکاران، ۲۰۱۶)، برنامه‌ریزی و استراتژی گزارشی مؤسسات آموزشی (کاپوتی و گاریدو^۳، ۲۰۱۵؛ مانکاد، ۲۰۱۶) و همچنین شناخت مدرسان و نحوه تصمیم‌گیری مدیران آموزشی. همچنین نستور و همکاران^۴ (۲۰۱۸) در تحقیق مروری خود، ۶ دسته داده‌های مورد استفاده 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 اشیا.

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؟

کدام یک 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 Link گردآوری 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. Cambrozzi et al.
3. Caputi & Garrido
4. Nistor et al.

تجزیه و تحلیل یادگیری مبتنی بر کامپیوتر^۱ (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 به وسیله آن بتوانند واکنش مناسبی به فرآیند یادگیری داشته باشند و اثربخشی آموزش را به درستی ارزیابی نمایند (ریتلس و همکاران^۲، 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 مورد نیاز برای افزایش خودآگاهی دانش آموزان در مورد فعالیت و محتوا کمک کنند. به‌عنوان مثال، داده‌های LMS با بررسی فعالیت‌های مرتبط با دوره، مانند انجمن‌های بحث، تحویل محتوا و ارزیابی، می‌توانند تنظیمات سیستم یادگیری را متناسب با هر یادگیرنده شخصی سازی نمایند.

تجزیه و تحلیل آینده‌نگر مبتنی بر کامپیوتر^۳ (CSPA)

از آنجایی که 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. (سمیرینگ 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 یادگیرندگان را که در معرض ترک تحصیل هستند شناسایی کنند و مشخص کنند چه کسانی نیازمند توجه بیشتر هستند.

تجزیه و تحلیل رفتار تحصیلی مبتنی بر کامپیوتر (CSBA)^۱

استفاده از تکنیک‌های داده کاوی می‌تواند بیش از قبیل توجهی داشته و الگوهای ارزشمندی را در رفتارهای یادگیری دانش آموزان نشان دهد (هه^۲، 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. به‌عنوان مثال، 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 هرناوندز و همکاران^۳ (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.

تجزیه و تحلیل اطلاعات بصری مبتنی بر کامپیوتر (CSVA)^۱

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). به‌عنوان مثال، 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).

ویرا و همکاران^۱ (۲۰۱۸) پس از مطالعه مروری خود بر روی مقالات و تحقیقات انجام شده 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.

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

برخی از مقالات بین‌المللی و داخلی منتشر شده 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 تحصيلات عاليه	۲۰۱۵	تویجری و همکاران
۱۱	پیش‌بینی یادگیرندگان که سرعت کمی در یادگیری دارند با استفاده از روش‌های داده کاوی مبتنی بر طبقه‌بندی و پیش‌بینی	۲۰۱۵	پارنیت کور و همکاران
۱۲	تحليل طول زمان تحصيل دانشجویان بر اساس روش‌های خوشه‌بندی و کشف قواعد انجمنی	۲۰۱۴	پتروس پلسیس و همکاران
۱۳	استفاده از داده کاوی آموزشی جهت گروه‌بندی یادگیرندگان در محیط یادگیری الکترونیکی به منظور شخصی سازی 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	روش های 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	تحليل 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 و مختاری
۲۳	چارچوبی برای الگوکاوی زمانی در سیستم‌های آموزش الکترونیکی: مورد کاوی 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 داده کاوی 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 شخصی سازی مطلوب تر در آموزش الکترونیکی	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" 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^۲ (۲۰۱۰) لازم است به آن‌ها در فرآیند پژوهش بر اساس علم طراحی در حوزه فناوری اطلاعات پاسخ داده شود، بررسی می‌نمایم. 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:

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

-
1. Design Science
 2. Hevner

جدول ۲: سؤالات سنجش فرآورده تولیدشده در علم طراحی

شماره	سؤال	پاسخ بر اساس پژوهش حاضر
۱	سؤال تحقیق چیست؟	چطور می توان با استفاده از داده های گذشته دانش آموزان، عملکرد تحصیلی آینده آن ها را پیش بینی 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 آینده نگری که راهکارهایی جهت بهبود عملکرد تحصیلی هر کدام از دانش آموزان ارائه نماید.
۳	فرآیند شکل گیری فرآورده چیست؟	تحلیل عملکرد گذشته دانش آموزان و خوشه بندی آن ها در دسته های مختلف تحصیلی و یافتن شباهت میان عملکرد دانش آموزان فعلی و دانش آموزان گذشته و در نهایت تشخیص زمینه های ممکن ارتقاء تحصیلی دانش آموزان فعلی
۴	چه دانش زمینه ای، فرآیند ساخت فرآورده را در بر گرفته است؟ چه تئوری، پشتوانه ساخت فرآورده است؟	۱- تکنیک ها و روش های خوشه بندی ۲- تکنیک ها و روش های آماری ۳- دانش تعلیم و تربیت
۵	چه معیارهای ارزیابی در طول شکل گیری 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.
۶	چطور فرآورده مورد نظر به در محیط آزمایش می شود؟ چه شاخص های برای برآورد کارایی فرآورده در محیط تعبیه شده است؟	این فرآورده از طریق تحلیل داده های دانش آموزان سایر مراکز تحصیلی مورد آزمایش قرار می گیرد. شاخص کارایی این فرآورده، افزایش عملکرد تحصیلی از طریق عمل به توصیه های پایانی است.
۷	چه دانشی به دانش گذشته اضافه شده است؟	بهبود توصیه های تحصیلی با استفاده از داده کاوی مبتنی بر داده های گذشته
۸	آیا به صورت رضایت بخشی به سؤال تحقیق پرداخته شده است؟	سیستم طراحی شده امکان بیان راهکارهایی را دارد که پیش از طراحی سیستم وجود نداشته است و نمی توان با تکیه بر نظرات تجربی آن ها 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 هر کدام 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 هر کدام 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. با این حال، 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^۱. 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	بی سواد	ابتدایی	سیکل	دیپلم دانشگاهی
	۵	تحصیلات پدر	بی سواد	ابتدایی	سیکل	دیپلم دانشگاهی
	۶	وضعیت سرپرستی	غیر از پدر و مادر	مادر	پدر	-
	۷	گذراندن مهد کودک	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						
بسیار مطلوب	مطلوب	متوسط	نامطلوب	بسیار نامطلوب	وضعیت سلامتی	۱۳	حمایتی و غیرمستقیم
-	-	-	عدم دسترسی	دسترسی	دسترسی به اینترنت	۱۴	
بسیار مطلوب	مطلوب	متوسط	نامطلوب	بسیار نامطلوب	کیفیت روابط خانوادگی	۱۵	
-	بیشتر از ۶۰ دقیقه	بین ۳۰ تا ۶۰ دقیقه	بین ۱۵ تا ۳۰ دقیقه	کمتر از ۱۵ دقیقه	فاصله زمانی منزل تا مدرسه	۱۶	
-	-	-	حمایت	عدم حمایت	حمایت مالی خانواده	۱۷	
-	-	-	پرداخت	عدم پرداخت	پرداخت هزینه‌های مدرسه	۱۸	
-	۳ درس	۲ درس	۱ درس	هیچ درسی	تجدیدی در سال گذشته	۱۹	
-	-	-	حمایت	عدم حمایت	حمایت مالی مدرسه	۲۰	
بسیار زیاد	زیاد	متوسط	کم	بسیار کم	زمان آزاد	۲۱	
بسیار زیاد	زیاد	متوسط	کم	بسیار کم	ساعات تفریح با دوستان	۲۲	
-	-	-	خیر	بله	ساعات فعالیت‌های فوق برنامه	۲۳	رفتار تحصیلی
-	بیشتر از ۱۰ ساعت	بین ۵ تا ۱۰ ساعت	بین ۲ تا ۵ ساعت	کمتر از ۲ ساعت	زمان مطالعه	۲۴	
از ۱ تا ۷۵ جلسه					تعداد غیبت کلاسی	۲۵	
نمره‌ای از ۰ تا ۲۰					نمره اولین آزمون	۲۶	نتایج تحصیلی
نمره‌ای از ۰ تا ۲۰					نمره دومین آزمون	۲۷	
نمره‌ای از ۰ تا ۲۰					نمره نهایی	۲۸	

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

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

پیش پردازش 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 گرادینت^۱، ۷ داده است. RMSE در الگوریتم درخت گرادینت 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. هر کدام 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.

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

جدول ۴: تعداد خوشه‌های مناسب برای هر کدام از ویژگی‌های تحصیلی به همراه نرمال‌سازی و کاهش ابعاد

شخصیتی و خنراکی های داده	ها خوشه تعداد	سازي نرمال بدون			شده نرمال			ابعاد کاهش یا هموار نرمال‌سازی		
		Kmeans	Kmedoids	K-fast	Kmeans	Kmedoids	K-fast	Kmeans	Kmedoids	K-fast
شخصیتی و خنراکی های داده	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	۰.۲۵۱	۰.۲۰۰
مستقیم غیر حمایتی های داده	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	۰.۲۳۷	۰.۱۸۱	۰.۲۱۵	۰.۲۰۸	۰.۱۹۲
تحصیلی رفتار های داده	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	۰.۲۲۶	۰.۲۹۵	۰.۲۰۳	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 می توان در سه بند خلاصه کرد:

تأثیر نرمال سازی و کاهش ابعاد: نرمال سازی و کاهش ابعاد هر دو سبب افزایش پایه ای عدد دیویس شده اند و میزان خطا را به طور کلی افزایش می دهند. به همین دلیل به نظر می رسد خوشه بندی داده ها بدون نرمال سازی و کاهش ابعاد نتیجه بهتری داشته باشد.

حداکثر تعداد خوشه ها: از آنجایی که تعدادهای خوشه های هر کدام از سه دسته فوق در یکدیگر ضرب خواهند شد و عدد به دست آمده در توصیه های مدیریتی استفاده می گردد، تعداد خوشه ها را نهایتاً تا ۵ عدد در نظر گرفته ایم تا در بیشترین حالت ۱۲۵ خوشه نهایی به دست آید و بتوان توصیه هایی را نسبت به آن ها در نظر گرفت. اگر چنانچه کمترین اعداد شاخص دیویس را به عنوان تعداد خوشه مطلوب در نظر بگیریم، برای داده های شخصیتی و خانوادگی، تعداد ۸ خوشه، برای داده های حمایتی و غیرمستقیم، تعداد ۹ خوشه و در داده های رفتار تحصیلی تعداد ۱۰ خوشه انتخاب خواهد شد که در نتیجه، ۷۲۰ حالت را ایجاد می کند. در حالی که تمام داده، ۱۰۳۸ داده است؛ بنابراین محدودیت حداکثر ۵ خوشه اعمال شده است. هر چند تعداد بیشتر خوشه ها نیز تفاوت چندان زیادی در شاخص دیویس را نشان نمی دهند.

الگوریتم های انتخاب شده: بنا بر جدول ۴ برای داده های شخصیتی و خانوادگی تعداد ۴ خوشه با الگوریتم Kmeans و برای داده های 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 داده های رفتار تحصیلی با الگوریتم Kmeans (دلیل مانند مورد گذشته است، تعیین شده است).

جدول نهایی توصیف هر کدام از دسته های خوشه ها را می توان در جدول ۵ مشاهده کرد:

جدول ۵: نام گذاری خوشه‌های مربوط به هر کدام 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 کم جمعیت	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

از آنجایی که قصد داریم تأثیر هر کدام 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. به نحوی که در هر کدام از این ۴ خوشه، تعداد داده‌ها تقریباً یکسان است. تعداد داده‌ها در هر کدام از محدوده‌ها فوق عبارت است از:

جدول ۶: 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
۴	بیشتر ۱۳,۳۳	نزدیک به اخراج یا ترک تحصیل	۲۵۱

اگر هر کدام از یادگیرندگان را از لحاظ 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 به صورت 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 غیر مستقیم	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 به جای 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 هر کدام 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 هر کدام از حالت‌های ترکیب خوشه‌ها آورده شده است:

جدول ۹: تحلیل یافته‌های مربوط به توصیف وضعیت موفقیت یا شکست هر کدام 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.
۴	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.
۵	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.

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.	۶
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.	۷
	۸
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.	۹
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.	۱۰
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.	۱۱
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.	۱۲
ترکیب‌های خوشه‌ای که متغیرهای شخصیتی و خانوادگی که پایه‌های ضعیف تحصیلی دارند و از جمعیت خانوادگی اندکی برخوردار هستند، به‌طور کلی 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.	۱۶ تا ۱۳

بنا بر جدول فوق می توان عامل موفقیت تحصیلی در هر کدام از ترکیب های خوشه ها را به صورت طبقه بندی کرد:

جدول ۱۰: بررسی عامل موفقیت هر کدام از ترکیب خوشه های تعیین شده

شرح عامل موفقیت	حالت های ترکیب خوشه ها
تمرکز کامل بر درس و تحصیل	۴ تا ۱
تمرکز نسبی بر درس و تحصیل	۸ تا ۵
برای حالت های ۹ و ۱۰: تمرکز نسبی بر درس و تحصیل برای حالت های ۱۱ و ۱۲: تمرکز کامل بر درس و تحصیل	۱۲ تا ۹
تمرکز کامل بر درس و تحصیل	۱۶ تا ۱۳

همچنین می توان مقایسه ای میان متغیرهای حمایتی و غیرمستقیم به صورت جدول ۱۱ انجام داد:

جدول ۱۱: مقایسه تأثیر متغیرهای حمایتی و غیرمستقیم در حالت های مختلف ترکیب خوشه های تعیین شده

تحلیل نتیجه	شماره حالت ترکیب خوشه ها
فرزندان خانواده هایی که روابط مناسبی دارند و از سلامتی بالاتری نیز برخوردار هستند، موفقیت تحصیلی کمتری به دست آورده اند.	۳ و ۱
عدم ارتباط معنادار.	۴ و ۲
فرزندان خانواده هایی که وضعیت ارتباطی و سلامتی مناسب تری دارند، نتایج تحصیلی بهتری را نیز کسب کرده اند.	۷ و ۵
عدم ارتباط معنادار.	۸ و ۶
فرزندان خانواده هایی که وضعیت ارتباطی و سلامتی مناسب تری دارند، نتایج تحصیلی بهتری را نیز کسب کرده اند.	۱۱ و ۹
فرزندان خانواده هایی که روابط مناسبی دارند و از سلامتی بالاتری نیز برخوردار هستند، موفقیت تحصیلی کمتری به دست آورده اند.	۱۲ و ۱۰
عدم ارتباط معنادار.	۱۵ و ۱۳
عدم ارتباط معنادار.	۱۶ و ۱۴

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 به‌طور کلی 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 هر کدام از یادگیرندگان، راهبردهایی را در نظر داشته باشند. به‌طور کلی یادگیرندگان در چهار دسته تقسیم شده و توصیه‌های کلی در رابطه با هر کدام 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 به‌طور کلی این دسته از یادگیرندگان، در سطح بالای تحصیلی قرار دارند.

یادگیرندگانی که تحصیلات خانوادگی پایینی دارند.

به‌طور کلی 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، در صورتی که این دسته بر روی تحصیل تمرکز داشته باشند، می‌توانند نتایج تحصیلی بالایی کسب نمایند.

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

در این گروه، یادگیرندگان به نوعی تحت تأثیر عامل مثبت یا منفی به خصوصی قرار نمی‌گیرند و به همین جهت احتمال شکست و ترک تحصیل به صورت بالقوه بالا است. این گروه نیز در صورتی که مورد توجه واقع نشوند، نتایج مناسبی کسب نخواهند کرد. همچنین پیشنهادهای پژوهشی عبارت‌اند از:

مقایسه الگوریتم‌های بکار گرفته شده در طبقه بندی

الگوریتم‌های بکار گرفته شده در این پژوهش را می‌توان به لحاظ برنامه‌نویسی بررسی کرد و آن‌ها را به دقت بالاتری رساند.

توسعه یک سیستم خبره جهت خوشه بندی و پیش بینی نتایج تحصیلی و ارائه توصیه‌های آموزشی

تمام مسیر طی شده در این پژوهش را می‌توان به صورت یک سیستم خبره طراحی کرد. به این صورت که اطلاعات یادگیرندگان گرفته شود و سپس سیستم، با استفاده از خوشه بندی و طبقه بندی، نتایج پیش بینی 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 به صورت احتمال فاصله گرفتن نتایج از میانگین مورد بررسی قرار گیرد. با این تفسیر که هر کدام 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.

شخصی سازی آموزش با استفاده از ایده پیش بینی تحصیلی مبتنی بر خوشه بندی

روش پیشنهادشده در این پژوهش می تواند فراتر از صرفاً پیش بینی نتایج تحصیلی، روش تحصیل یا مسیر آموزش مناسب برای هر کدام از دانش آموزان را از طریق خوشه بندی روش های تحصیل (نظیر آنلاین، غیر حضوری، تقویتی، آموزش دانش آموزی، رقابتی و سایر روش ها) پیش بینی نماید. به این صورت که پس از ارزیابی میزان موفقیت تحصیلی هر کدام از دانش آموزان در هر روش تحصیل، خوشه هایی از دانش آموزان موفق و نیمه موفق و همچنین خوشه هایی از دانش آموزان ناموفق یا تقریباً ناموفق قرار گیرند و پس از اتمام فرآیند یادگیری ماشینی، برآورد شود که یک دانش آموز جدید در کدام یک از خوشه های تحصیلی قرار می گیرد تا بر اساس شرایط او، روش تحصیل متناسب پیشنهاد شود.

تحلیل رفتار تحصیلی دانشجویان در دانشگاه های ایران بر اساس ایده محوری پژوهش

داده های سه گانه دانشجویان از قبیل داده های شخصیتی، داده های رفتار تحصیلی (انتخاب واحد و سایر داده ها) و داده های نتایج پایانی ترم های تحصیلی، در کنار یکدیگر می توانند نشان دهنده خوشه های رفتار موفق، رفتار نیمه موفق، رفتار تقریباً ناموفق و رفتار ناموفق باشند. می توان بر اساس داده های دانشجویان در ابتدای ورود به دانشگاه، مسیر احتمالی موفقیت را به آن ها پیشنهاد کرد. البته رسیدن به الگوی موفقیت برای دانشجویان یک دانشگاه به عوامل محیطی بسیاری بستگی دارد و بدون دسترسی به تعداد قابل توجهی از متغیرها و داده های تحصیلی، آموزش مستمر الگوریتم های 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 به منظور شخصی سازی برنامه آموزش. مدیریت و برنامه ریزی در نظام‌های آموزشی، ۱۱(۱)، ۱۰۸-۸۳.
- رضائی، عباسعلی و زاهدی، محمدهادی. (۱۳۹۷). نقش فن آوری‌های 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 علی اکبر، فتحیان بروجنی، محمد و نامنی، احمد. (۱۳۹۶). مدلی برای پیش‌بینی آسیب‌پذیری تحصیلی مقطع کارشناسی مبتنی بر شبکه عصبی. مدیریت و برنامه ریزی در نظام‌های آموزشی، ۱۰(۱)، ۱۰۲-۸۱.
- زاهدبابلان، ع؛ معینی کیا، م؛ و درخشانیفرد، 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 داده کاوی 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 21st century 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-García, 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), 1-11.
- 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).