



تحلیل تراکنش‌های امانت و گردش منابع کتابخانه‌های دانشگاه علوم پزشکی بیرجند با الگوریتم‌های داده‌کاوی

حسین بیدختی^۱ / نصرت ریاحی نیا^۲

چکیده

مقدمه: داده‌کاوی فرآیندی رایانه‌ای برای کشف روابط و الگوهای معنادار بین داده‌های بزرگ است. شناسایی الگوهای رفتاری کاربران به بهبود تصمیم‌گیری در کتابخانه‌ها کمک می‌کند. هدف این پژوهش تحلیل تراکنش‌های گردش منابع کتابخانه‌های دانشگاه علوم پزشکی بیرجند با الگوریتم‌های داده‌کاوی است.

روش کار: در این مطالعه توصیفی، برای تعیین الگوهای پنهان در داده‌های گردش منابع کتابخانه‌های دانشگاه علوم پزشکی بیرجند، از روش‌های داده‌کاوی استفاده شد. اطلاعات اعضا و تراکنش‌های گردش منابع از نرم‌افزار کتابخانه استخراج شد و پس از پالایش و افزودن فیلدهای محاسباتی، در قالب فایل‌های اکسل ذخیره شدند. برای انجام مراحل مختلف کشف دانش، از افزونه داده‌کاوی Microsoft SQL Server 2008 در نرم‌افزار آفیس ۲۰۰۷ استفاده شد. ماتریس طبقه‌بندی، برای محاسبه میزان صحت مدل‌های تولیدشده، به‌کاربرده شد. یافته‌های پژوهش در قالب شکل و جدول گزارش شدند.

یافته‌ها: ۳۹۴۰۱۱ تراکنش امانت و بازگشت کتاب توسط ۵۶۰۰ عضو تا پایان سال ۱۳۹۲ در نرم‌افزار کتابخانه ثبت شده بودند. گردش منابع در هر نیمسال تحصیلی از توالی منظمی برخوردار بود. ۸۷ درصد اعضا، کمتر از ۴۶ روز کتاب‌های امانتی را بازگشت می‌دهند و ۹۶ درصد دانشجویان کمتر از ۱۸ روز در عودت کتاب‌ها تأخیر دارند. اعضا به‌طور متوسط ۳۵ مدرک در مدت عضویتشان به امانت می‌گیرند. با استفاده از قوانین وابستگی می‌توان ارتباط موضوعی مدارک موردعلاقه اعضا را کشف کرد.

نتیجه‌گیری: داده‌کاوی منجر به شناسایی رفتار اطلاعاتی اعضا، طبقه‌بندی اعضای فعال و شناسایی عوامل مؤثر بر مدت امانت و دیرکرد مدارک گردید. الگوهای آشکارشده جهت بهبود تصمیم‌گیری در مدیریت نیروی انسانی و تدوین مقررات خدمات گردش منابع کتابخانه پیشنهاد می‌شود.

کلیدواژه‌ها: داده‌کاوی، اداره کتابخانه، بررسی‌های پیمایشی کتابخانه، خودکارسازی کتابخانه، مدیریت اطلاعات

• وصول مقاله: ۹۴/۰۱/۱۴ • اصلاح نهایی: ۹۴/۰۶/۲۳ • پذیرش نهایی: ۹۴/۰۸/۰۵

۱. مربی گروه اطلاع‌رسانی پزشکی، دانشکده پزشکی، دانشگاه علوم پزشکی بیرجند، بیرجند، ایران؛ دانشجوی دکتری علم اطلاعات و دانش‌شناسی، دانشگاه خوارزمی، تهران، ایران؛ نویسنده مسئول (beydokhti54@bums.ac.ir)
۲. استاد گروه علم اطلاعات و دانش‌شناسی، دانشکده روانشناسی و علوم تربیتی، دانشگاه خوارزمی، تهران، ایران

مقدمه

فناوری داده‌کاوی، یکی از روش‌های سودمند در یافتن اطلاعات مفید در بین حجم عظیمی از داده‌ها است. داده‌کاوی با کشف الگوها و روابط بین داده‌ها، ارزش‌های پنهانی آن‌ها را آشکار می‌نماید. با کشف این دانش می‌توان ارزش‌های دیگر متغیرها را پیش‌بینی کرد و در امر تصمیم‌گیری از آن بهره‌جست.

امروزه افزایش تعداد و حجم پایگاه‌های داده، توجه سازمان‌ها به دانش ارزشمند مستور در میان انبوه داده‌ها و پیشرفت ابزارهای قدرتمند داده‌کاوی، زمینه استفاده گسترده از داده‌کاوی در سازمان‌ها را فراهم کرده است [۱].

طبقه‌بندی، تخمین، پیش‌بینی، خوشه‌بندی، گروه‌بندی شباهت یا قوانین وابستگی و نمایه‌سازی از کارکردها و وظایف عمده داده‌کاوی هستند [۲، ۳]. سه مورد اول در گروه داده‌کاوی هدایت‌شده قرار می‌گیرند که هدف آن یافتن ارزش یک متغیر خاص است. خوشه‌بندی و گروه‌بندی شباهت جزو داده‌کاوی غیر هدایت‌شده هستند که برای یافتن ساختار پنهان درون داده‌ها بدون توجه به یک متغیر خاص مورد استفاده قرار می‌گیرند. نمایه‌سازی، عملی توصیفی است که می‌تواند هم هدایت‌شده و هم غیر هدایت‌شده باشد [۱، ۴].

کتابخانه‌ها در طی دهه‌های گذشته از سیستم‌های نرم‌افزاری برای مدیریت اعضا، فهرست‌نویسی، جستجو و گردش منابع استفاده کرده‌اند. هر کدام از این کارکردها با انواع مختلفی از داده‌ها سروکار دارند که به‌طور سنتی به دودسته رکوردهای اصلی و رکوردهای عملیاتی تقسیم و معمولاً به‌طور جداگانه پردازش می‌شوند [۵]. استفاده از آمار پیشرفته و تکنیک‌های داده‌کاوی برای تحلیل این گونه داده‌ها می‌تواند مفید باشد اما متأسفانه تعداد کمی از کتابخانه‌ها، از مزایای الگوهای متداول رفتار کاربران برای ارتقای کیفیت خدمات، مدیریت منابع مالی و تصمیم‌گیری‌های راهبردی استفاده می‌کنند [۶].

مطالعات گذشته در خصوص داده‌کاوی تراکنش‌های امانت و بازگشت مدارک حاکی از سودمندی این روش در مدیریت و ارتقای کیفیت خدمات کتابخانه‌ها است. کائو و همکارانش در

مطالعه‌ای، داده‌کاوی و کشف اطلاعات ناشناخته از داده‌های تاریخی امانت را در تصمیم‌گیری و تخصیص بودجه تأمین منابع کتابخانه ارزشمند یافتند [۷]. هانگ و همکارانش ارتباط بین الگوهای جستجو و منابع کتابخانه را بر اساس تحلیل قوانین وابستگی در یک ساختار سلسله‌مراتبی مصور، بررسی کردند [۸]. گوسومی و همکارانش در مقاله‌ای نشان دادند استخراج الگوها و روابط از داده‌های گردش منابع و مصور کردن آن‌ها می‌تواند در مدیریت کتابخانه بسیار مؤثر باشد [۹].

پینگ یو در سال ۲۰۱۱ از نرم‌افزار SQL Server 2005 به‌عنوان ابزار داده‌کاوی برای شناخت خصوصیات امانت‌گیرندگان و کشف روابط بین داده‌های امانت‌مدارک، اطلاعات اعضا و منابع کتابخانه استفاده کرد. وی در این مطالعه، استفاده از تکنیک‌های داده‌کاوی را برای طبقه‌بندی امانت‌گیرندگان بر اساس خصوصیات هر گروه، تحلیل کتاب‌های مورد توجه خوانندگان و بررسی میزان استفاده اعضا از بخش امانت کتابخانه، مفید گزارش کرد [۱۰].

تسوجی و همکارانش برای ارزیابی منابع پیشنهادشده به کاربران از سه روش مشارکت اعضا، داده‌کاوی سوابق تراکنش‌ها و مدل به‌کار گرفته‌شده در سایت آمازون استفاده کردند. نتایج این مطالعه نشان داد چنانچه بر اساس سابقه گردش منابع کتاب به اعضای کتابخانه پیشنهاد شود، بهترین روش، استفاده از داده‌کاوی بر اساس قوانین وابستگی خواهد بود [۱۱].

مطالعات دیگری نیز از تکنیک‌های داده‌کاوی برای بررسی رفتار اعضا در استفاده از منابع به‌منظور کشف نیازهای اطلاعاتی آن‌ها استفاده کرده‌اند [۱۲-۱۵].

در مدیریت خدمات امانت و گردش منابع کتابخانه‌ها، تدوین آیین‌نامه و به‌کارگیری صحیح آن برای توزیع عادلانه منابع بین کاربران همواره مورد توجه مسئولین و کارشناسان کتابخانه‌ها بوده است. از سوی دیگر تسهیل فرایند گردش منابع و ارائه خدمات متناسب با نیاز کاربران و امکانات کتابخانه موجب افزایش رضایت‌مندی اعضا و ارتقای کیفیت خدمات خواهد شد. برای مدیریت خدمات امانت و گردش منابع، تصمیم‌گیری در مورد تعداد و مدت امانت منابع، تعداد دفعات تمدید مجاز،

ولیعصر (عج) و از ۱۳۸۶ در کتابخانه بیمارستان امام رضا (ع) راه‌اندازی شده است و تا پایان سال ۱۳۹۲ اطلاعات ۷۶۱۱ عضو و تعداد ۵۱۱۳۲۲ مورد تراکنش امانت و گردش منابع در این نرم‌افزار ثبت شده است. به منظور بررسی و تحلیل رفتار اعضا در امانت‌گیری منابع، داده‌های امانت و گردش کتاب‌ها از ابتدای سال ۱۳۸۳ لغایت پایان سال ۱۳۹۲ انتخاب شدند.

برای دسترسی به اطلاعات کتابشناختی، گردش کتاب‌ها و اعضا، دستورات پرس جوی مناسب، بر اساس جداول اطلاعاتی ONVAN, AMANAT, MAINREG, PERSON_PROFILE, PERSON استفاده از بخش گزارش‌های تعبیه‌شده در رابط کاربر نرم‌افزار جامع کتابخانه پارس آذرخش، داده‌های موردنیاز استخراج و به صورت فایل‌های اکسل جداگانه ذخیره شدند.

برای پالایش و تصفیه اطلاعات، فایل‌های اکسل بررسی شد و داده‌های مخدوش و ناکارآمد شامل کلیه تراکنش‌های قبل از سال ۱۳۸۳ و تراکنش‌های امانت مربوط به فعالیت‌های مدیریتی کتابخانه شامل منابع ارسال شده به صحافی، مطالعه در سالن و گردش منابع مربوط به گروه‌های آموزشی حذف شدند. یکسان‌سازی برخی داده‌ها از جمله عنوان رشته‌ها، نام دانشکده‌ها، مقاطع تحصیلی و تاریخ امانت انجام شد. با استفاده از نرم‌افزار Microsoft Excel 2007، نشانه رده و عدد صحیح رده اصلی از شماره راهنمای کتاب‌ها استخراج شدند. همچنین با استفاده از قابلیت‌های این نرم‌افزار مدت عضویت، تعداد روزهای مجاز امانت، تعداد روزهایی که یک مدرک در امانت عضو بوده، مدت تأخیر و تعجیل محاسبه و کلیه تاریخ‌های شمسی به میلادی تبدیل و در قالب فیلدهای اطلاعاتی به جداول افزوده شدند. از اول مهرماه تا پایان شهریورماه سال بعد به عنوان یک سال تحصیلی در نظر گرفته شد و این فیلد نیز به داده‌ها اضافه شد. به طور مثال، برجسب سال تحصیلی ۱۳۸۳ به کلیه تراکنش‌های اول مهرماه ۱۳۸۳ لغایت ۳۱ شهریورماه ۱۳۸۴ افزوده شد. برای اعضای که تاریخ اعتبارشان در زمان استخراج داده‌ها منقضی نشده بود و همچنین برای مدارکی که در امانت اعضا بودند، مدت عضویت، مدت امانت و تأخیر تا پایان سال ۱۳۹۲ محاسبه و درج شد.

زمان مناسب برای پیگیری دیرکرد منابع امانتی و شناسایی منابع پراستفاده و تهیه نسخه‌های تکراری آن‌ها ضرورت دارد.

تصمیم‌گیری برای همه موارد فوق کاملاً بستگی به شرایط و امکانات کتابخانه و شناخت صحیح از اعضا و نیازهای اطلاعاتی آن‌ها دارد. امروزه اکثر کتابخانه‌ها از نرم‌افزارهای کتابخانه‌ای برای مدیریت گردش منابع خود استفاده می‌کنند. دسترسی به گزارش تراکنش‌های گردش منابع و توسعه فناوری داده کاوی و نرم‌افزارهای قدرتمندی که بدین منظور تهیه شده‌اند، برای طبقه‌بندی و کشف روابط بین داده‌ها در کتابخانه‌ها، موقعیت مناسبی را فراهم کرده است. شناسایی روابط و الگوهای نهفته در داده‌های گردش منابع، پیش‌بینی تاریخ بازگشت منابع، شناسایی منابع مرتبط با رشته و مقطع تحصیلی اعضا، شناسایی منابع پراستفاده و اطلاع از مدت امانت، مسئولین و کارشناسان کتابخانه را قادر خواهد ساخت تا در تعامل با اعضا تصمیمات بهتری را اتخاذ نمایند. از این‌رو در این پژوهش که احتمالاً نخستین بررسی علمی و عملی کاربرد ابزارهای داده کاوی در مدیریت داده‌های کتابخانه‌ای ایران است، سوابق امانت و گردش منابع کتابخانه‌های دانشگاه علوم پزشکی بیرجند از سال ۱۳۸۳ لغایت ۱۳۹۲ با استفاده از روش‌های داده کاوی به منظور کشف دانش نهفته در داده‌ها، ارتقای کیفیت خدمات و بهبود تصمیم‌گیری در کتابخانه مورد تحلیل قرار گرفت.

روش کار

کتابخانه مرکزی دانشگاه علوم پزشکی بیرجند از نرم‌افزار جامع کتابخانه پارس آذرخش استفاده می‌نماید؛ که یکی از نرم‌افزارهای پراستفاده در کتابخانه‌های دانشگاه‌های علوم پزشکی کشور است. معماری این نرم‌افزار در دانشگاه علوم پزشکی بیرجند به صورت متمرکز بوده و خدمات فهرست‌نویسی، رده‌بندی، عضویت و گردش منابع کتابخانه مرکزی و کتابخانه‌های بیمارستان‌های ولیعصر (عج) و امام رضا (ع) به صورت متمرکز و یکپارچه توسط یک سیستم نرم‌افزاری تحت شبکه ارائه می‌شود. این سامانه نرم‌افزاری از سال ۱۳۸۱ در کتابخانه مرکزی، ۱۳۸۲ در کتابخانه بیمارستان

بیشتر امانت گرفته‌اند، از تحلیل سبد خرید (Shopping Basket Analysis) در افزونه داده‌کاوی اکسل استفاده شد. این ابزار روابط بین موضوع‌ها را بر اساس الگوریتم قوانین وابستگی ماکروسافت تحلیل می‌کند که در این مطالعه سطح اطمینان ۹۰ درصد برای پیش‌بینی روابط در نظر گرفته شد.

یافته‌ها

تا پایان سال ۱۳۹۲، تعداد ۷۶۱۱ نفر در کتابخانه‌های دانشگاه علوم پزشکی بیرجند عضو شده بودند که ۵۶۰۰ نفر به‌عنوان اعضای فعال، مدت عضویت بیش از ۱۱۹ روز و تعداد تراکنش امانت و تمدید کتاب بیشتر از چهار مورد بود (جدول یک). تعداد کل تراکنش‌های گردش کتاب اعضای فعال ۳۹۴۰۱۱ مورد و بیش از ۷۷ درصد آن‌ها مربوط به دانشجویان بود. سرانه تراکنش‌های امانت به تعداد اعضا، برای دانشجویان ۶۹، کارکنان ۷۳، اساتید و اعضای هیئت‌علمی ۸۹ و اعضای میهمان ۹۵ مورد برای هر عضو بود. میانگین تعداد عناوین امانتی در سال برای دانشجویان 13 ± 8 عنوان و برای کارکنان 8 ± 7 ، اساتید و اعضای هیئت‌علمی دانشگاه 7 ± 6 و سایر اعضا 11 ± 7 عنوان بوده است. ۷۷۷۹ عنوان کتاب فارسی و ۲۳۵۱ عنوان کتاب لاتین در طی سال‌های ۱۳۸۳ تا ۱۳۹۲ توسط اعضا امانت گرفته شده است. ۹۰ درصد کتاب‌های فارسی و ۷۲ درصد کتاب‌های لاتین توسط دانشجویان مطالعه شده و سهم اساتید و اعضای هیئت‌علمی دانشگاه از کل کتاب‌های امانتی، ۲۸۹۱ (۳۷ درصد) عنوان فارسی و ۱۰۶۰ (۴۵ درصد) عنوان لاتین بوده است.

فایل‌های اکسل پس از اصلاح و افزودن برخی فیلدهای محاسباتی و اطلاعاتی به نرم‌افزار Microsoft SQL Server 2008 R2 منتقل شدند. از قابلیت‌های نرم‌افزار SQL در تدوین پرس‌وجوهای مناسب برای استخراج داده‌های موردنیاز هر یک از مراحل داده‌کاوی استفاده شد. داده‌های استخراج شده از هریک از پرس‌وجوها، در قالب فایل‌های Microsoft Excel 2007 ذخیره شدند.

در مرحله دوم پالایش داده‌ها، با استفاده از ابزار Clean Data در افزونه داده‌کاوی اکسل ۲۰۰۷، اعضای که مدت عضویشان کمتر از ۱۲۰ روز و تعداد تراکنش امانت و تمدید منابعشان کمتر از پنج مورد بود به‌عنوان اعضای غیرفعال در نظر گرفته شدند و از فایل کارکرد اعضا حذف شدند.

برای مدل‌سازی و انجام مراحل مختلف داده‌کاوی نرم‌افزار Microsoft SQL Server 2008 Data Mining Add-Ins for Office 2007 به کار گرفته شد. الگوریتم درخت تصمیم ماکروسافت برای کشف بهترین نسبت بین فیلدها و عوامل مؤثر بر مدت امانت، دیرکرد مدارک و تعداد مدارک امانتی استفاده شد. مدل‌ها در بخش آموزش معادل ۳۰ درصد داده‌ها، ساخته و در بخش آزمایش معادل ۷۰ درصد داده‌ها، آزمون شدند و درخت تصمیم بر اساس هر یک از فیلدهای خروجی استخراج شدند. برای اعتباربخشی و محاسبه میزان صحت مدل‌های تولید شده از ماتریس طبقه‌بندی (Classification Matrix) استفاده شد. این ماتریس ابزاری مفید برای تحلیل نتایج مدل‌های دسته‌بندی و پیش‌بینی است [۱۶]. برای بررسی این که اعضا به‌طور معمول چه موضوع‌هایی را

جدول ۱: کارکرد امانت و گردش کتاب اعضای کتابخانه‌های دانشگاه علوم پزشکی بیرجند (۱۳۹۲-۱۳۸۳)

نوع عضویت	فراوانی اعضا		دفعات امانت و تمدید	
	فعال (درصد)	غیرفعال* (درصد)	کتاب فارسی (درصد)	کتاب لاتین (درصد)
دانشجویان	۴۴۰۴ (۷۸/۶)	۱۴۵۰ (۲۲/۱)	۲۸۱۵۴۱ (۷۷/۵)	۲۲۲۷۶ (۷۲/۹)
کارکنان	۹۹۹ (۱۷/۸)	۴۹۷ (۲۴/۷)	۶۸۲۷۱ (۱۸/۸)	۴۲۳۰ (۱۳/۸)
اساتید	۱۷۱ (۳/۱)	۲۹ (۱/۵)	۱۱۲۶۶ (۳/۱)	۳۹۶۴ (۱۳)
سایر	۲۶ (۰/۵)	۳۵ (۱/۷)	۲۳۶۸ (۰/۶)	۹۵ (۰/۳)
جمع	۵۶۰۰ (۱۰۰)	۲۰۱۱ (۱۰۰)	۳۶۳۴۴۶ (۱۰۰)	۳۰۵۶۵ (۱۰۰)

* اعضای که مدت عضویت کمتر از ۱۲۰ روز یا تعداد کارکرد کمتر از پنج مورد داشتند به‌عنوان اعضای غیرفعال در نظر گرفته شدند.

اعضای را بر اساس ویژگی‌ها و خواص مشابه به سطح یک با کمتر از ۳۵ روز فاصله بین امانت هر عنوان کتاب، سطح دو بین ۳۶ تا ۸۰ روز، سطح سه بین ۸۱ تا ۱۵۵ روز، سطح چهار بین ۱۵۶ تا ۲۳۷ روز و سطح پنج با بیشتر از ۲۳۷ روز فاصله بین امانت هر عنوان کتاب طبقه‌بندی کرد. جدول دو فراوانی اعضای فعال را در سطح‌های پنج‌گانه به تفکیک نوع عضویت نشان می‌دهد.

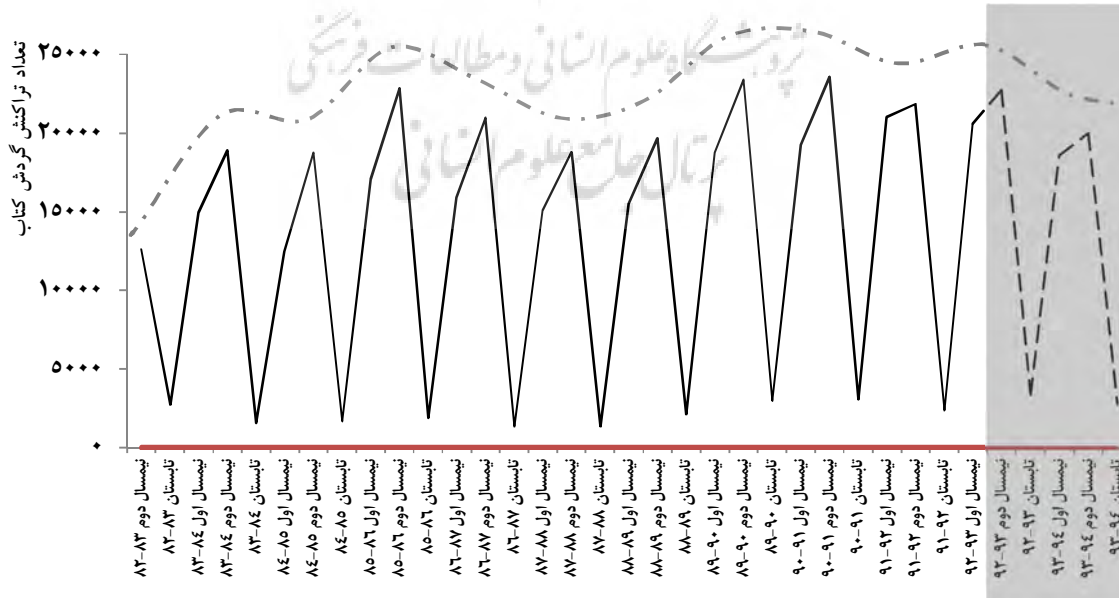
برای شناسایی و توصیف اعضای فعال از بخش طبقه‌بندی (Detect Categories) در ابزار تحلیل جدول (Table Analysis Tools) افزونه داده‌کاوی نرم‌افزار اکسل، استفاده شد. اعضای فعال توسط الگوریتم تحلیل خوشه‌بندی، بر اساس میانگین فاصله زمانی بین امانت هر عنوان کتاب در پنج سطح طبقه‌بندی شدند. این الگوریتم به صورت غیر هدایت‌شده [۱]،

جدول ۲: طبقه‌بندی اعضای فعال بر اساس میانگین تعداد روزهای بین امانت هر عنوان کتاب در کتابخانه‌های دانشگاه علوم پزشکی بیرجند (۱۳۹۲-۱۳۸۳)

نوع عضویت	سطح				
	سطح ۱ < ۳۵ روز (درصد)	سطح ۲ ۳۵-۸۰ روز (درصد)	سطح ۳ ۸۱-۱۵۵ روز (درصد)	سطح ۴ ۱۵۶-۲۳۷ روز (درصد)	سطح ۵ > ۲۳۷ روز (درصد)
دانشجویان	۳۲۶۰ (۷۴/۱)	۹۹۱ (۲۲/۵)	۱۳۲ (۳)	۱۵ (۰/۳)	۶ (۰/۱)
کارکنان	۳۸۴ (۳۸/۵)	۳۶۳ (۳۶/۳)	۱۸۶ (۱۸/۶)	۴۵ (۴/۵)	۲۱ (۲/۱)
اساتید	۴۶ (۲۶/۹)	۵۶ (۳۲/۷)	۴۲ (۲۴/۶)	۲۳ (۱۳/۵)	۴ (۲/۳)
سایر	۱۳ (۵/۰)	۱۰ (۳۸/۵)	۲ (۷/۷)	۱ (۳/۸)	۰
جمع	۳۷۰۳ (۶۶/۱)	۱۴۲۰ (۲۵/۳)	۳۶۲ (۶/۵)	۸۴ (۱/۵)	۳۱ (۰/۶)

کتابخانه‌های دانشگاه علوم پزشکی بیرجند از سال ۱۳۸۲ لغایت ۱۳۹۴ نشان داده شده است. روند پیش‌بینی شده برای نیمسال دوم سال تحصیلی ۹۳-۹۲ لغایت تابستان ۱۳۹۴ به صورت خط چین نشان داده شده است.

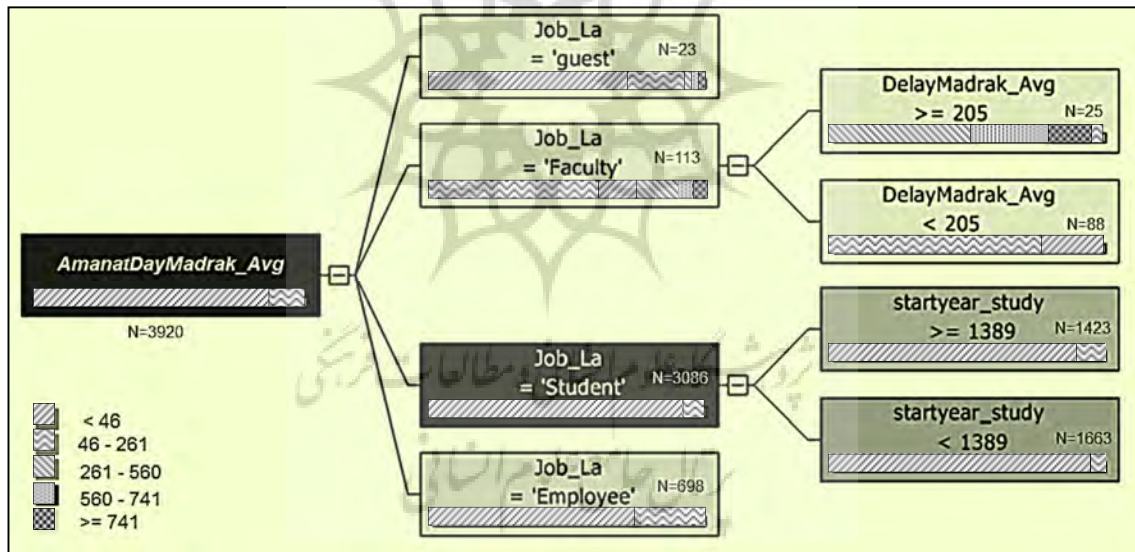
برنامه‌ریزی تقاضا، مشتمل بر سری‌های زمانی است و مدیران کتابخانه‌ها را قادر می‌سازد تا از نوسانات و تغییرات رخ داده در کارکرد امانت و گردش منابع در طی زمان، درک بهتری داشته باشند. در شکل یک با به کارگیری امکانات پیش‌بینی تعبیه شده در افزونه داده‌کاوی اکسل، روند امانت کتاب‌ها در



شکل ۱: روند امانت کتاب‌ها در کتابخانه‌های دانشگاه علوم پزشکی بیرجند بر حسب هر نیمسال تحصیلی

کرده‌اند. ۵۱ درصد از اعضای این گروه که تأخیر بیش از ۲۰۵ روز داشته‌اند، مدارک امانتی را بین ۲۶۱ تا ۵۶۰ روز برگشت داده‌اند. ۹۳ درصد دانشجویان، مدارک امانتی را کمتر از ۴۶ روز عودت داده‌اند. قبل از سال تحصیلی ۸۹-۹۰ تنها سه درصد از دانشجویان کتاب‌ها را بین ۴۶ تا ۲۶۱ روز به امانت گرفته‌اند که بعد از آن سال، به ۱۱ درصد افزایش یافته است. نمودار شبکه وابستگی (Dependent Network) نشان داد که نوع عضویت، میزان تأخیر در عودت مدارک و تاریخ عضویت نسبت به سایر عوامل تأثیر بیشتری بر مدت زمان امانت کتاب‌ها دارند. بررسی شبکه وابستگی نشان داد که ارتباط عوامل تأثیرگذار بر مدت امانت یک‌طرفه بوده و وابستگی آن به نوع عضویت بیشتر از سایر عوامل بوده است.

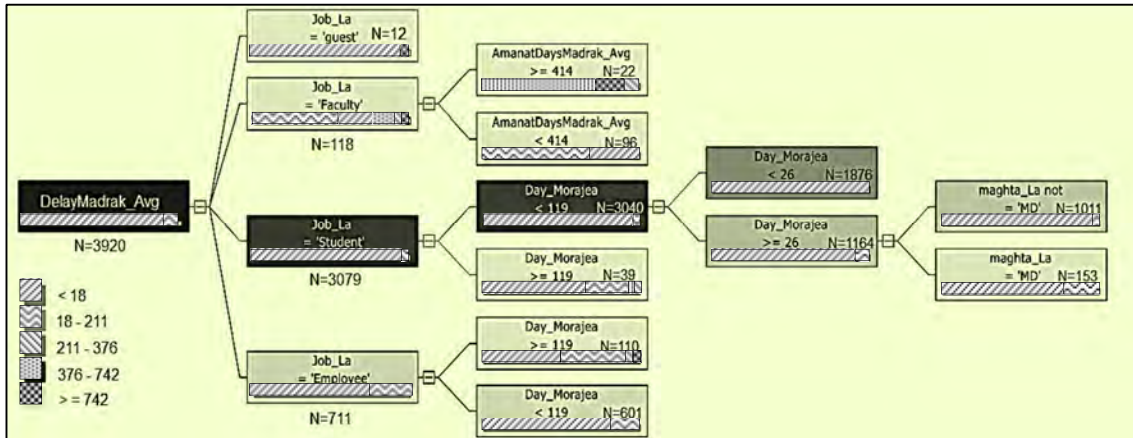
برای بررسی و تصمیم‌گیری در خصوص مدت مناسب برای امانت مدارک از مدل دسته‌بندی بر اساس الگوریتم درخت تصمیم ماکروسافت استفاده شد. فیلدهای اطلاعاتی نوع عضویت، جنس، سن، رشته و مقطع تحصیلی، دفعات مراجعه، امانت و تمدید، تعداد عناوین و کتابخانه‌های مورد استفاده، مدت عضویت، میانگین دیرکرد هر عنوان و فاصله هر دفعه مراجعه به کتابخانه‌ها برای گروه‌بندی میانگین مدت امانت، مورد استفاده قرار گرفت. شکل دو درخت تصمیم را برای مدت امانت نشان می‌دهد که صحت مدل آن بر اساس ماتریس طبقه‌بندی ۸۹/۲۷ درصد محاسبه شد. ۸۷ درصد اعضا، کتاب‌های امانتی را کمتر از ۴۶ روز و ۱۱ درصد بین ۴۶ تا ۲۶۱ روز عودت داده‌اند. ۵۹ درصد اعضای هیئت‌علمی، امانت بین ۴۶ تا ۲۶۱ روز را تجربه



شکل ۲: درخت تصمیم مدت امانت کتاب‌های فارسی و لاتین در کتابخانه‌های دانشگاه علوم پزشکی بیرجند

نسبت به سایر اعضا در عودت مدارک امانتی دیرکرد بیشتری داشته‌اند. اعتبار مدل بر اساس ماتریس طبقه‌بندی ۹۱/۱۶ درصد محاسبه شد. شبکه وابستگی، به ترتیب رابطه مؤثر گروه عضویت، مدت امانت مدارک، فاصله مراجعه و مقطع تحصیلی را بر میزان تأخیر در عودت مدارک امانتی نشان داد.

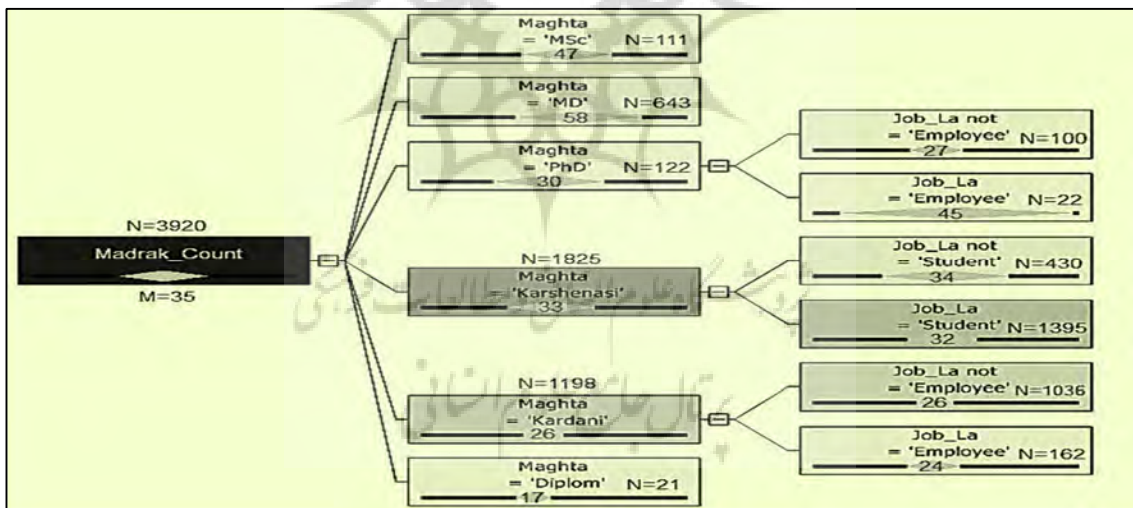
شکل سه عوامل مرتبط با میزان دیرکرد مدارک را نشان می‌دهد. ۹۶ درصد دانشجویان، مدارک امانتی را با کمتر از ۱۸ روز دیرکرد به کتابخانه عودت داده‌اند. دانشجویان دکترای حرفه‌ای که فاصله مراجعه آن‌ها به کتابخانه بین ۲۶ تا ۱۱۹ روز بود، نسبت به سایر مقاطع تحصیلی دیرکرد بیشتری داشته‌اند. اعضای هیئت‌علمی



شکل ۳: درخت تصمیم مدت دیرکرد مدارک امانتی در کتابخانه‌های دانشگاه علوم پزشکی بیرجند

اعضا به‌طور متوسط ۳۵ مدرک در مدت عضویشان به امانت گرفته‌اند. همان‌طور که شکل چهار نشان می‌دهد تعداد مدارک امانت بر اساس مقطع تحصیلی متفاوت بوده و تعداد آن‌ها با افزایش مقطع تحصیلی از دیپلم تا دکتری حرفه‌ای

افزایش یافته است. اعضای کارمند دانشگاه با مقطع دکتری تخصصی نسبت به اعضای هیئت‌علمی و دانشجویان این مقطع تحصیلی تعداد مدارک بیشتری به امانت گرفته‌اند.



شکل ۴: درخت تصمیم تعداد مدارک امانت گرفته‌شده در کتابخانه‌های دانشگاه علوم پزشکی بیرجند

پیش‌بینی ارتباط بین امانت یک مدارک با امانت گرفتن سایر مدارک، یکی دیگر از کاربردهای داده‌کاوی در کتابخانه‌ها است. با استفاده از ابزار تحلیل سبد خرید در افزونه داده‌کاوی اکسل بر اساس الگوریتم قوانین وابستگی ماکروسافت، موضوع‌های کتاب‌های امانت گرفته‌شده توسط اعضا در سطح اطمینان ۹۰ درصد بررسی شد. در جدول دو تعدادی از

ترکیب‌های موضوعی دوگانه و سه‌گانه کتاب‌ها گزارش شده است که بیشتر مورد استفاده اعضا قرار گرفته است. همان‌طور که در جدول سه مشخص است کتاب‌های فارسی با موضوع کلیات فیزیولوژی انسانی به همراه کتاب‌های فارسی کلیات کالبدشناسی توسط ۱۱۳۸ عضو به امانت گرفته‌شده‌اند و هر عضو به‌طور متوسط ۱۱ مدرک از این موضوعات را

امانت گرفته یا اینکه ۱۱ دفعه آن‌ها را تمدید کرده است. استفاده توأم کتاب‌های این دو موضوع بیشترین میزان تراکنش امانت و تمدید (۱۲۴۱۰ دفعه) را به خود اختصاص داده‌اند. همین‌طور استفاده از ترکیب موضوعی سه‌گانه اطلس‌های پزشکی، مجموعه سؤالات جراحی و سؤالات دستیاری و پره‌انترنی توسط ۵۳ نفر از اعضا ۱۶۰۲ دفعه مورد امانت و تمدید قرار گرفته است.

جدول ۳: تحلیل موضوع کتاب‌های امانت گرفته‌شده با استفاده از الگوریتم قوانین وابستگی

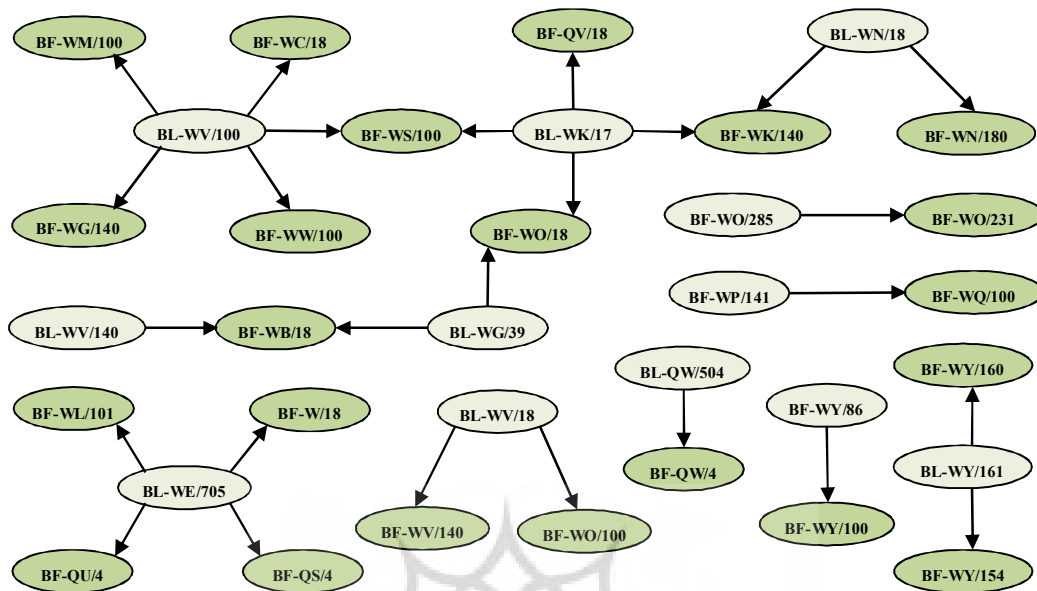
کل دفعات امانت و تمدید (درصد)	میانگین دفعات امانت	اعضا تعداد (درصد)	موضوع کتاب‌ها
۱۲۴۱۰ (۳/۱)	۱۰/۹	۱۱۳۸ (۲۰/۳)	کلیات فیزیولوژی انسانی (فارسی) * کلیات کالبدشناسی (فارسی)
۱۱۰۲۰ (۲/۸)	۱۷/۲	۶۴۱ (۱۱/۴)	چکیده‌ها و مجموعه سؤالات پرستاری (فارسی) * پرستاری مراقبت‌های ویژه (فارسی)
۱۰۹۴۸ (۲/۷)	۱۱/۸	۹۲۶ (۱۶/۵)	اطلس کالبدشناسی (لاتین) * کلیات کالبدشناسی (فارسی)
۱۰۷۸۵ (۲/۶)	۱۰/۸	۱۰۰۱ (۱۷/۸)	کلیات زیست‌شیمی (فارسی) * کلیات کالبدشناسی (فارسی)
۱۰۴۵۳ (۲/۶)	۱۳/۴	۷۸۱ (۱۳/۹)	پرستاری مراقبت‌های ویژه (فارسی) * مهارت‌های پرستاری و مراقبت (فارسی)
۱۰۳۰۳ (۲/۵)	۱۱	۹۳۵ (۱۶/۷)	کلیات زیست‌شیمی (فارسی) و کلیات فیزیولوژی انسانی (فارسی)
۹۸۶۰ (۲/۴)	۱۶/۸	۵۸۷ (۱۰/۴)	چکیده‌ها و مجموعه سؤالات پرستاری (فارسی) * مهارت‌های پرستاری و مراقبت (فارسی)
۹۵۵۵ (۲/۳)	۱۷/۱	۵۵۹ (۹/۹)	پرستاری کودکان (فارسی) * چکیده‌ها و مجموعه سؤالات پرستاری (فارسی)
۹۲۹۲ (۲/۲)	۱۲	۷۷۲ (۱۳/۷)	اطلس کالبدشناسی (لاتین) * کلیات فیزیولوژی انسانی (فارسی)
۹۷۸۹ (۲/۱)	۱۶/۱	۵۴۵ (۹/۷)	پرستاری اعصاب و روان (فارسی) * چکیده‌ها و مجموعه سؤالات پرستاری (فارسی)
⋮	⋮	⋮	⋮
۱۶۰۲ (۰/۴۱)	۳۰/۲	۵۳ (۰/۹۵)	اطلس کلیات پزشکی (لاتین) * چکیده‌ها و مجموعه سؤالات جراحی و بیهوشی (فارسی) * چکیده‌ها و سؤالات دستیاری و پره‌انترنی (فارسی)
۱۵۴۷ (۰/۳۹)	۳۶/۸	۴۲ (۰/۷۵)	اطلس کلیات پزشکی (فارسی) * چکیده‌ها و سؤالات دستیاری و پره‌انترنی (فارسی) * چکیده‌ها و سؤالات علوم پایه (فارسی)
۱۵۲۲ (۰/۳۹)	۴۳/۴	۳۵ (۰/۶۳)	فلسفه و نظریه‌های پرستاری (فارسی) * چکیده‌ها و مجموعه سؤالات پرستاری (فارسی) * مهارت‌های پرستاری و مراقبت (فارسی)
۱۵۱۹ (۰/۳۹)	۳۱	۴۹ (۰/۸۸)	اطلس کلیات پزشکی (فارسی) * کلیات آسیب‌شناسی (فارسی) * چکیده‌ها و سؤالات دستیاری و پره‌انترنی (فارسی)
۱۴۳۶ (۰/۳۶)	۳۵	۴۱ (۰/۷۳)	فلسفه و نظریه‌های پرستاری (فارسی) * پرستاری مراقبت‌های ویژه (فارسی) * مهارت‌های پرستاری و مراقبت (فارسی)
⋮	⋮	⋮	⋮

Min. Rule Probability = 90%

کلیات بیماری‌های گوش و حلق و بینی را به امانت گرفته‌اند، کتاب‌هایی با موضوعات چشم‌پزشکی، روان‌پزشکی، بیماری‌های کودکان، نوار قلب و چکیده‌ها و مجموعه سؤالات بیماری‌های واگیردار را نیز امانت خواهد گرفت. همچنین اعضایی که کتاب‌هایی با موضوع کلیات بیماری‌های کودکان

در شکل پنج وابستگی موضوع‌های کتاب‌های امانت گرفته شده با احتمال ۹۰ درصد نشان داده شده است. برخی از گره‌ها در این شکل پیش‌بینی کننده و برخی پیش‌بینی شونده هستند. الگوی کشف شده در داده‌های امانت و گردش کتاب‌ها بر اساس الگوریتم قوانین وابستگی نشان داد اعضایی که کتاب‌های

را امانت گرفته‌اند، قبلاً کتاب‌های اطلس دستگاه غدد درون‌ریز و کلیات بیماری‌های گوش و حلق و بینی را امانت گرفته‌اند.



شکل ۵: وابستگی موضوعی کتاب‌های امانت گرفته‌شده با استفاده از الگوریتم قوانین وابستگی (احتمال $\leq 90\%$)

تحصیلی به‌طور متوسط چهار تا پنج عنوان کتاب مطالعه می‌کنند. بینگ یو نیز در مطالعه‌اش اذعان کرد طبقه‌بندی امانت‌گیرندگان و بررسی خصوصیات هر گروه به‌منظور ارائه خدمات مناسب به آن‌ها مفید است و منجر به ارتقای کیفیت خدمات در کتابخانه خواهد شد [۱۰].

فعالیت‌های امانت و بازگشت کتاب، سفارشات، خدمات تأمین مدرک، استفاده کاربران از اینترنت از جمله خدماتی هستند که قابلیت بررسی و تحلیل از طریق سری زمانی را دارا می‌باشند. بررسی روند گردش منابع در یک دوره زمانی ده‌ساله نشان داد تقاضا برای امانت و بازگشت منابع در کتابخانه‌های دانشگاه علوم پزشکی بیرجند تقریباً از نظم یکسانی برخوردار بوده است (شکل یک). مطالعات گذشته نشان می‌دهد تحلیل گزارش تراکنش‌های امانت و گردش منابع در طی زمان و بررسی میزان استفاده اعضا از بخش امانت در مدیریت نیروی انسانی کتابخانه مؤثر است [۱۰]. در زمان اوج تقاضا می‌توان با تمرکز بر خدمات گردش منابع و در زمان کاهش تقاضا، می‌توان کارکنان را به مرتب کردن

بحث و نتیجه‌گیری

در این مطالعه تلاش شد با استفاده از افزونه داده کاوی در نرم‌افزار اکسل، داده‌های امانت و گردش منابع کتابخانه به‌صورت کاربردی با روش‌های ساده داده کاوی مورد تحلیل قرار گیرد تا الگوها و روابط پنهان داده‌ها کشف و مورد بهره‌برداری قرار گیرد. در ابتدا بررسی کلی کارکرد اعضا و مدارک استفاده‌شده نشان داد که بیش از ۷۷ درصد کل تراکنش‌های گردش منابع و ۹۰ درصد کل کتاب‌های مطالعه شده توسط دانشجویان انجام شده است که نشان‌دهنده اهمیت دانشجویان در کتابخانه‌های دانشگاهی است.

طبقه‌بندی اعضا بر اساس میانگین فاصله زمانی بین امانت هر عنوان کتاب نشان داد که ۶۶ درصد اعضا به‌طور متوسط کمتر از هر ۳۵ روز، یک عنوان کتاب به امانت گرفته‌اند (جدول دو). با شناخت رفتار اعضا در کتابخانه، می‌توان به نیاز آن‌ها پی برد، لذا از گروه‌بندی اعضا می‌توان نتیجه گرفت که ۶۶ درصد اعضای فعال که عمدتاً دانشجو هستند در هر نیمسال

کتاب‌ها، فعالیت‌های پژوهشی و استفاده بهینه از اوقات فراغت ترغیب کرد. با استفاده از الگوهای دوره‌ای در امانت و گردش منابع می‌توان خدمات گردش منابع را برنامه‌ریزی کرد. همچنین، می‌توان برای ارتقای کیفیت خدمات واحدهای مرتبط، یک برنامه علمی سالیانه تدوین کرد. در مطالعات سری زمانی به دلیل روش مشاهده، خطا انکارناپذیر است، زیرا اختلال و بی‌نظمی در کنار هم هستند. قسمت‌های تأییدشده و بخش‌های تصادفی در سری‌های زمانی وجود دارند، اما می‌توان بر اساس قسمتی که قابل پذیرش است یا قسمتی که به واقعیت نزدیک است، قضاوت کرد.

اگرچه می‌توان نوسانات امانت و بازگشت منابع را پیش‌بینی کرد و با استفاده از تحلیل توالی‌های زمانی، نظم منطقی در کار ایجاد کرد و حداکثر بهره‌برداری از نیروی انسانی را داشت، اما پیش‌شرط استفاده از این روش این است که تعداد اعضا و کتاب‌ها نسبتاً پایدار و ثابت بماند. اگر اعضا تغییرات زیادی داشته باشند (از جمله تغییرات حرفه‌ای، تعداد، شیوه آموزش و آموختن، ساعات برگزاری کلاس‌ها) موجب تغییراتی در تقاضاهای امانت و بازگشت مدارک خواهد شد که در این صورت می‌باید تحلیل و پیش‌بینی دوباره انجام شود [۱۳].

درخت تصمیم‌گیری تقریباً یکی از رایج‌ترین تکنیک‌های داده‌کاوی است. معمول‌ترین زمینه کاری درخت تصمیم، دسته‌بندی است. در ساخت یک مدل، الگوریتم ارزیابی می‌کند که هر مشخصه ورودی در مجموعه داده چگونه در نتیجه مشخصه پیش‌بینی شده، تأثیر می‌گذارد و سپس از این مشخصه‌های ورودی، قوی‌ترین رابطه را به‌منظور ساخت مجموعه‌ای از تقسیم‌ها استفاده می‌کند، که گره نامیده می‌شوند. گره‌های جدید به مدل اضافه می‌شوند و این ترتیب یک ساختار درختواره شکل می‌گیرد [۱۶]. در این مطالعه تلاش شد با استفاده از الگوریتم درخت تصمیم ماکروسافت، عوامل مؤثر بر مدت امانت، دیرکرد و تعداد مدارک امانتی بررسی شود. همان‌طور که شکل‌های دو و سه نشان می‌دهند ۸۷ درصد اعضا، کتاب‌هایی که امانت گرفته بودند کمتر از ۴۶ روز به کتابخانه عودت داده‌اند و ۹۶ درصد دانشجویان

مدارک امانتی را با کمتر از ۱۸ روز تأخیر عودت داده‌اند. از آنجایی که مدت امانت کتاب‌ها در کتابخانه‌های دانشگاه علوم پزشکی بیرجند برای دانشجویان، کارکنان و اعضای میهمان ۱۴ روز تعیین شده است، به نظر می‌رسد افزایش مدت امانت کتاب‌هایی که دارای نسخه‌های تکراری می‌باشند از ۱۴ روز به ۳۰ روز، می‌تواند تراکنش‌های تمدید و درنهایت دیرکرد مدارک امانتی را به میزان قابل توجهی کاهش دهد. به عبارت دیگر افزایش مدت امانت کتاب‌ها، در صورتی که منجر به کاهش دسترسی به منابع نشود، می‌تواند موجب افزایش رضایت اکثریت اعضا گردد. آگاهی کتابداران از رفتار اعضا در امانت و بازگشت مدارک، می‌تواند مبنای زمان پیگیری دیرکرد منابع باشد، لذا از آنجایی که ۹۳ درصد دانشجویان کتاب‌های امانتی را کمتر از ۴۶ روز به کتابخانه عودت داده‌اند، می‌توان برای دانشجویانی که بیش از ۴۶ روز دیرکرد دارند، پیگیری دیرکرد را انجام داد. این تصمیم، صرفه جویی در نیروی انسانی و هزینه‌های جانبی را به همراه خواهد داشت.

شبکه وابستگی معرف قدرت مشخصه‌های ورودی در پیش‌بینی متغیر هدف می‌باشند. هر گره بیانگر یک مشخصه و هر بردار نشانگر وجود یک رابطه بین دو گره است [۱۶]. بررسی متغیرهای مؤثر بر مدت امانت و دیرکرد کتاب‌ها در پژوهش حاضر نشان داد که متغیر نوع عضویت، مهم‌ترین عامل در مدت امانت و دیرکرد مدارک است. در کتابخانه‌های دانشگاه علوم پزشکی بیرجند، مدت مجاز امانت کتاب‌ها برای اساتید و اعضای هیئت علمی دانشگاه، ۳۰ روز و برای سایر اعضا ۱۴ روز تعیین شده و قوانین جریمه دیرکرد مدارک برای دانشجویان سخت‌گیرانه‌تر اعمال می‌شود. از این رو بر اساس اطلاعات مندرج در شکل‌های دو و سه می‌توان نتیجه گرفت که عموماً اساتید و اعضای هیئت علمی دانشگاه نسبت به سایر اعضا مدت‌زمان بیشتری کتاب‌ها را در اختیار دارند و احتمال اینکه مدارک امانتی‌شان را در موعد تعیین شده عودت ندهند، زیاد است.

استفاده نشان داده شده است. پینگ یو به منظور کشف روابط بین داده‌های امانت مدارک کتابخانه از الگوریتم قوانین وابستگی استفاده کرد. وی دریافت که تحلیل کتاب‌های مورد توجه خوانندگان، می‌تواند به تأمین کتاب‌های مورد نیاز آن‌ها و مجموعه‌سازی کمک کند. وی در مطالعه خود نشان داد که میزان استفاده از کتاب‌های موضوع ادبیات همراه با کتاب‌های موضوع‌های زبان، تاریخ و جغرافیا، اقتصاد، فلسفه، فناوری و علوم اجتماعی بیشتر بوده است و بیش از ۷۵ درصد اعضایی که کتاب‌های ادبیات را امانت گرفته‌اند از کتاب‌های موضوع تاریخ و جغرافیا نیز استفاده کرده‌اند [۱۰]. زانگ و همکارانش نیز در تحلیل منابع امانت گرفته شده، گزارش کردند که دانشجویان مهندسی مکانیک هرگز کتاب‌های موضوع علوم سیاسی و حقوق را به امانت نگرفته‌اند و بیشترین استفاده آن‌ها از کتاب‌های ریاضیات و علوم وابسته بوده است. همچنین با افزایش سابقه تحصیلی دانشجویان دانشکده علوم انسانی، میزان استفاده آن‌ها از کتاب‌های ادبیات کاهش و امانت گرفتن کتاب‌های موضوع حقوق افزایش یافته است [۱۵]. بررسی مطالعات قبلی و نتایج مطالعه حاضر نشان می‌دهد تحلیل منابع امانت گرفته شده و شناخت علایق و رفتار امانت و گردش منابع که منشأ گرفته از نیاز اطلاعاتی اعضا است، می‌تواند در مدیریت کتابخانه‌ها و ارتقای کیفیت خدمات مناسب باشد.

استفاده از روش داده کاوی برای کشف الگوها و روابط بین داده‌های تراکشی امانت و گردش منابع در کتابخانه می‌تواند به استفاده بهینه از نیروی انسانی، شناخت نیازهای واقعی کاربران، تدوین قوانین و میزان خدمات قابل ارائه به اعضا، تخصیص مناسب بودجه برای تأمین منابع مورد نیاز گروه‌های مختلف، تحول در خدمات مرجع و ظهور خدمات جدیدی مانند پیشنهاد منابع مناسب به اعضا، منجر شود.

کتابخانه‌ها با استفاده از داده کاوی می‌توانند کیفیت تعامل بین خود و کاربران را بهبود بخشند. داده‌های گردآوری شده حاوی اطلاعات ارزنده‌ای است که می‌توانند با استراتژی کتابخانه یکی شده و برای بهبود تصمیم‌گیری‌های کتابخانه شمر ثمر واقع گردند [۱۸]. داده کاوی ماهیت چرخه مانند دارد

بررسی تعداد مدارک امانتی توسط گروه‌های مختلف اعضا نشان داد که اعضای کتابخانه‌های دانشگاه علوم پزشکی بیرجند در طول مدت عضویشان به‌طور متوسط ۳۵ مدرک به امانت گرفته‌اند. همان‌طور که در شکل چهار نیز نشان داده شده است، مقطع تحصیلی اعضا نسبت به سایر عوامل، تأثیر بیشتری در تعداد مدارک امانتی داشته است. نتایج این بخش می‌تواند بر تعیین مدارک مجاز قابل امانت اعضا تأثیر داشته باشد. بر این اساس پیشنهاد می‌شود در کتابخانه‌های دانشگاه علوم پزشکی بیرجند تعداد مدارک مجاز برای اعضا بر اساس مقطع تحصیلی تعیین شود. بررسی گردش منابع کتابخانه مرکزی دانشگاه کیوشو ژاپن توسط مینامی نشان داد که بیش از نیمی از دانشجویان حداقل پنج کتاب در سال امانت می‌گیرند. دانشجویان مقطع کارشناسی ارشد ۲۳ درصد و دانشجویان دکتری تخصصی ۱۶ درصد حجم گردش منابع را به خود اختصاص داده‌اند [۱۲]. البته روش بررسی مطالعه مینامی با مطالعه حاضر متفاوت بوده است اما همان‌طور که شکل چهار نشان می‌دهد دانشجویان مقطع کارشناسی به‌طور متوسط در طی دوره عضویشان ۳۲ کتاب به امانت گرفته‌اند که متوسط استفاده سالانه دانشجویان این مقطع به نتایج مطالعه مینامی نزدیک است.

تحلیل سبد خرید، توانایی فهم این مطلب است که مشتری‌ها به‌طور معمول چه کالاهایی را سفارش می‌دهند، یا سبد خرید را چه اقلامی تشکیل می‌دهد [۱۶]. با تحلیل منابع امانت گرفته شده در کتابخانه‌ها با استفاده از الگوریتم قوانین وابستگی می‌توان ترجیحات اعضا در استفاده از منابع را شناسایی و مدیریت بهتری بر منابع کتابخانه اعمال کرد [۱۷]. این بررسی نشان داد (جدول دو) ترکیب دو گانه امانت کتاب‌های فارسی کلیات فیزیولوژی انسانی و کالبدشناسی بیشترین تکرار را داشته است. همین‌طور امانت گیری اطلس‌های لاتین کلیات پزشکی که عمدتاً مجموعه اطلس‌های CIBA هستند به همراه چکیده‌ها و سؤالات جراحی و سؤالات آمادگی برای آزمون دستیاری و پره‌انترنی بیشترین تکرار را ترکیب‌های سه گانه داشته‌اند. همچنین در شکل ۵ وابستگی موضوعی کتاب‌ها به‌منظور شناسایی تقدم و تأخر در

References

1. Dwivedi RK, Bajpai R, editors. Use of Data Mining in the field of Library and Information Science: An Overview. 2nd International CALIBER-2004; 2004 Feb 11-13; New Delhi, India. Ahmedabad: INFLIBNET Centre; 2004
2. Alizadeh S, Ghazanfari M, Teimorpour B. [Data mining and knowledge discovery]. 2nd ed. Tehran: Iran University of Science and Technology; 2011. [Persian].
3. Chen SY, Liu X. The contribution of data mining to information science. Journal of Information Science 2004; 30(6): 550-8.
4. Shahrabi J. [Data mining]. Tehran: Jahad Daneshgahi; Amirkabir University of Technology; 2011. [Persian].
5. Patkar VN. Data mining applications in library and academic institutions. Information Studies 2005; 11(3): 145-56.
6. Nicholson S, Stanton JM. Gaining strategic advantage through bibliomining: Data mining for management decisions in corporate, special, digital, and traditional libraries. In: Nemati H, Barko C, editors. Organizational data mining: Leveraging enterprise data resources for optimal performance. Hershey, PA: Idea Group Publishing; 2003. p. 247-62.
7. Kao S-C, Chang H-C, Lin C-H. Decision support for the academic library acquisition budget allocation via circulation database mining. Information Processing & Management 2003; 39(1): 133-47.

زیرا در پی کشف الگوها، سؤالات بیشتری پیش خواهد آمد که دور بعدی فرایند را شکل می‌دهند. بهره‌برداری از فناوری‌های پیشرفته مثل داده‌کاوی، مطمئناً برای متخصصان کتابداری و مدیران مؤسسات آموزشی یک چالش دائمی خواهد بود؛ چراکه آن‌ها خلاقیت طلبند و برای نوآوری تلاش می‌کنند [۵].

آماده‌سازی داده‌ها برای استفاده در فرایند داده‌کاوی بسیار وقت‌گیر و طاقت‌فرسا است. ضرورت دارد متناسب با هر کارکرد مورد انتظار، داده‌ها و فیلدهای اطلاعاتی مناسبی تهیه و مورد استفاده قرار گیرند. شرایط و قوانین حاکم بر کتابخانه و تنوع اعضا بر داده‌های تراکنشی گردش منابع تأثیر دارد، لذا هنگام تفسیر الگوها و روابط بین داده‌ها، لازم است به شرایط و امکانات کتابخانه و شناخت صحیح اعضا و نیازهای اطلاعاتی آن‌ها توجه شود.

اگرچه داده‌کاوی پیشرفت شگرفی در نوع ابزارهای تحلیلی به وجود آورده است، اما محدودیت‌هایی نیز در کاربردپذیری آن وجود دارد. یکی از این محدودیت‌ها این است که ابزارهای داده‌کاوی هنوز استانداردسازی نشده‌اند [۱۹]. اخیراً تعداد، تنوع و پیچیدگی طرح‌های مربوط به داده‌کاوی یا کشف دانش در پایگاه‌های اطلاعاتی شتاب بیشتری گرفته و لازم است استانداردسازی جنبه‌های مربوط به فرایند توسعه آن‌ها مورد توجه قرار گیرد تا بتوان نتایج مربوط را به کار بست و در آینده، آن‌ها را مجدد مورد استفاده قرارداد [۲۰]. از این رو پیشنهاد می‌شود محققان این حوزه، استانداردسازی استفاده از فنون داده‌کاوی و کشف دانش در خدمات کتابداری و اطلاع‌رسانی را مورد توجه قرار دهند.

تشکر و قدردانی

نویسندگان مقاله بدین وسیله مراتب سپاس و قدردانی خود را از آقای مهندس مهدی قاضی‌زاده به خاطر راهنمایی‌های ارزنده ایشان در به‌کارگیری نرم‌افزار SQL اعلام می‌دارند.

8. Huang CM, Kang SH, Chang CC, Lu SH. Apply Data Mining Techniques to Library Circulation Records and Usage Patterns Analysis. [cited 2015 Feb 20]. Available from: <http://140.125.84.58:8080/TeacherWeb/fileDownload/55.pdf>.
9. Goswami S, Verma S, Krishnan NR. Relationship visualization between books and users based on mining library circulation data. Proceedings of the First International Conference on Intelligent Interactive Technologies and Multimedia; 2010 Dec. 28-30; Allahabad, India: ACM; 2010. p. 109-13.
10. Yu P. Data Mining in Library Reader Management. 2011 International Conference on Network Computing and Information Security (NCIS); 2011 May 14-15; Guilin, China: Institute of Electrical and Electronics Engineers (IEEE); 2011. p. 54-7.
11. Tsuji K, Kuroo E, Sato S, Ikeuchi U, Ikeuchi A, Yoshikane F, et al. Use of Library Loan Records for Book Recommendation. Proceedings of the 2012 IIAI International Conference on Advanced Applied Informatics; 2012 Sept. 20-22; Fukuoka: IEEE Computer Society; 2012. p. 30-5.
12. Minami T. Potentials of Circulation Data Analysis for Library Marketing: A Case Study in a University Library. In: Kim T-h, Adeli H, Cuzzocrea A, Arslan T, Zhang Y, Ma J, et al. , editors. Database Theory and Application, Bio-Science and Bio-Technology. Berlin Heidelberg: Springer; 2011. p. 90-9.
13. Le X, Yan L, Ying J. Scientific management in classical books circulation using the time sequence mining method. 2011 International Conference on E-Business and E-Government (ICEE); 2011 May 6-8; Shanghai, China: Institute of Electrical and Electronics Engineers (IEEE); 2011. p. 1-3.
14. Wang R, Tang Y, Li L. Application of BP neural network to prediction of library circulation. In: Sugawara K, Wang Y, Hattori F, Nishida T, Fujita S, Kinsner W, et al. , editors. 11th International Conference on Cognitive Informatics & Cognitive Computing (ICCI*CC); 2012 Aug. 22-24; Kyoto: IEEE; 2012. p. 420-3.
15. Zhang Q, Li X, Wang P. Association Analysis of Library Borrowing Information Based on Data Mining Technology. In: Liu X, Ye Y, editors. Proceedings of the 9th International Symposium on Linear Drives for Industry Applications. Berlin Heidelberg: Springer; 2014. p. 721-8.
16. ShahrabiJ, Shakoorniyaz V. [Data mining in SQL Server]. Tehran: Jahad Daneshgahi Amirkabir; 2009. [Persian].
17. Bussaban K, Kularbphettong K. Analysis of Users Behavior on Book Loan Log Based on Association Rule Mining. International Journal of Computer, Control, Quantum and Information Engineering 2014; 8(1): 18-20.
18. Chang C-C, Chen R-S. Using data mining technology to solve classification problems: A case study of campus digital library. The Electronic Library 2006; 24(3): 307-21.
19. Seifert JW. Data mining: An overview. In: Pegarkov DD, editor. National security issues. New York: Nova Science Publishers; 2006. p. 201-17.
20. Marbán O, Segovia J, Menasalvas E, Fernández-Baizán C. Toward data mining engineering: A software engineering approach. Information Systems 2009; 34(1): 87-107.



The Analysis of interlibrary Loan Transactions in Birjand University of Medical Sciences using Data Mining Algorithms

Beydokhti H¹/ Riahinia N²

Abstract

Introduction: Data mining is a process for discovering meaningful relationships and patterns from data. Identify behavior patterns of libraries users can helps improve decision-making in libraries. This study aimed to analyze the interlibrary loan transactions in Birjand University of Medical Sciences using data mining algorithms.

Methods: In this descriptive study, knowledge discovery and data mining were used to determine patterns of interlibrary transactional and circulation data in Birjand University of Medical Sciences. Information on the members and circulation transactions was extracted from the library database, and saved in Excel file format after adding some computational fields. Data were processed using Microsoft SQL Server 2008 Data Mining Add-Ins, Office 2007. Classification Matrix was used to evaluate the accuracy of the models. The findings were reported in tables and graphs.

Results: 394,011 records of library circulation transactions from 5600 members were collected up to 20th of June 2013. Flow of transactions per semester was of a regular sequence. About 87% of the borrowed books were returned less than 46 days and 96% of students fail to return books on due date and the delay time is less than 18 days. The members borrowed 35 books on average during their membership. Related subject of interest to members was discovered through association rule algorithm.

Conclusion: Data mining can be used for identifying behavior patterns of members, classifying and identifying factors affecting the duration of book loan and delay. Revealed patterns are suggested for decision-making in human resources management and development of regulations on library circulation services.

Keywords: Data Mining, Library Administration, Library Surveys, Library Automation, Information Management

• Received: 3/April/2015 • Modified: 14/Sep/2015 • Accepted: 27/Oct/2015

-
1. Instructor of Medical Information Department, Faculty of Medicine, Birjand University of Medical Sciences, Birjand, Iran; PhD Student of Knowledge and Information Science, Kharazmi University, Tehran, Iran; Corresponding Author (beydokhti54@bums.ac.ir)
 2. Professor of Knowledge & Information Science Department, Faculty of Psychology and Education, Kharazmi University, Tehran, Iran