

بررسی استفاده از داده‌کاوی و یکپارچه‌سازی داده‌های منابع انسانی

رضا بختیاری، کارشناسی ارشد علوم کامپیوتر، Reza.Bakhtiari@shahed.ac.ir

چکیده

در این مقاله نقش داده‌کاوی و یکپارچه‌سازی داده‌های منابع انسانی ارائه شده است. داده‌کاوی یکی از راه‌های تحلیل و استفاده از اطلاعات است که با استفاده از روش‌های تخصصی آماری و منطقی به تحلیل داده‌های بزرگ‌مقیاس می‌پردازد و به موسسات در اخذ تصمیم‌های کلان کمک می‌کند. اهداف مدیریت بهتر داده‌های منابع انسانی و بررسی ارتباط بین مدیریت منابع انسانی (HRM)، داده‌کاوی و مدیریت اقتصادی است. درک عمیق دانش نهفته در داده‌های منابع انسانی برای موقعیت رقابتی شرکت و تصمیم‌گیری سازمانی حیاتی است. تجزیه و تحلیل الگوها و روابط در داده‌های منابع انسانی بسیار مهم می‌باشد. داده‌های منابع انسانی معمولاً برای پاسخگویی به سوالات پردازش می‌شوند. از آنجایی که داده‌های منابع انسانی به پردازش داده‌ها یعنی وارد کردن داده‌ها و ذخیره آن‌ها در پایگاه داده برای اهداف گوناگون نظیر گزارش‌گیری نیاز دارد، می‌بایست به داده‌های قابل اندازه‌گیری توجه کرد. مقاله حاضر به بررسی تحقیقات در مورد داده‌کاوی منابع انسانی می‌پردازد تا به طور سیستماتیک پیشرفت‌های اخیر و همچنین وظایف باقی مانده برای کارهای آینده را کشف کند. در این مقاله نشان داده می‌شود که داده‌کاوی و یکپارچه‌سازی داده‌ها، الگوهای مفیدی را از این مجموعه داده بزرگ استخراج می‌نماید. در این مقاله پیشنهادهای نیز جهت جمع‌آوری داده‌ها و توانایی داده‌کاوی جهت افزایش عملکردها ارائه می‌گردد.

واژه‌های کلیدی: داده‌کاوی، یکپارچه‌سازی، تصمیم‌گیری، جمع‌آوری داده‌ها، داده‌های منابع انسانی، مدیریت داده

۱- مقدمه

در دنیای امروز، سازمان‌ها باید توانایی استفاده هوشمندانه از داده‌های موجود و سرمایه‌های فکری خود را داشته باشند. سازمان‌ها و شرکت‌ها مهم‌ترین منابع را به عنوان کارکنان موجود در سازمان و شرکت شناسایی می‌کنند. در همین راستا برای ایجاد تمایز یک شرکت و یا سازمان، نیاز است که مهارت، سازگاری، دانش و فداکاری کارکنان افزایش یابد. افزایش تصاعدی داده‌ها، که عمدتاً ناشی از جمع‌آوری الکترونیکی داده‌ها و ذخیره آن‌ها در پایگاه‌های داده می‌باشد، تقاضا برای تجزیه و تحلیل این داده‌های تولید شده توسط سازمان‌ها و شرکت‌ها را افزایش می‌دهد تا سازمان‌ها و شرکت‌ها بتوانند به سرعت پاسخگوی نیازهای خود باشند.

داده‌کاوی که همان کشف دانش در داده‌های بزرگ می‌باشد، پس از جمع‌آوری، تجزیه و تحلیل و دسترسی به داده‌های سازمان‌ها و شرکت‌ها تصمیم‌گیری‌های مناسب اتخاذ می‌گردد. برای این تصمیم‌گیری‌ها از ابزارهایی مانند ابزارهای پرسش و گزارش، ابزارهای پردازش تحلیلی و ابزارهای سیستم پشتیبانی تصمیم استفاده می‌کند.

داده‌های مورد نیاز برای مدیریت منابع انسانی معمولاً وجود ندارد و اغلب در بسیاری از سازمان‌ها و شرکت‌ها در دسترس نیست. در همین راستا به لطف داده‌کاوی و یکپارچه بودن داده‌ها، مدیر اجرایی می‌تواند با ارزش‌ترین دارایی خود یعنی کارکنان شرکت را مدیریت نماید.

پیروی از تعریفی که اغلب استفاده می‌شود، داده‌کاوی به فرآیند غیر پیش پا افتاده شناسایی الگوهای جدید، مفید و معتبر در داده‌ها اشاره دارد [۱]. در این میان مجموعه گسترده‌ای از حوزه‌های کاربردی داده‌کاوی و حوزه‌های تحقیقاتی مربوطه وجود دارد که در میان آنها دامنه مدیریت شرکتی با زیر دامنه‌های به خوبی تثبیت شده مانند مدیریت مشتری [۲]، مدیریت تولید [۳] یا مدیریت مالی [۴] وجود دارد.

با اصلاح و توسعه مستمر نظام اقتصادی، توسعه و رقابت بنگاه‌ها نیز پیوسته در حال تغییر است. علاوه بر این، مدل رقابت نیز از مبتنی بر منابع مادی به مبتنی بر منابع انسانی تغییر کرده است. بنابراین، مدیریت منابع انسانی (HRM) به یک پیوند حیاتی برای شرکت‌ها تبدیل شده است [۵] و [۶]. سیستم مدیریت منابع انسانی برنامه ریزی منابع انسانی را به عنوان پایه و اساس می‌گیرد. می‌تواند از طریق جمع‌آوری و مدیریت یکپارچه اطلاعات کارکنان، قضاوتی کلی در مورد اطلاعات کارکنان داخلی شرکت ارائه دهد [۷] و [۸]. HRM دستی نه تنها وقت گیر و کار فشرده است، بلکه مستعد خطاهای قابل توجه است. از این رو فناوری اطلاعات کامپیوتری برای حل این مشکل معرفی شده است. فناوری داده‌کاوی در سال‌های اخیر به سرعت توسعه یافته است. این یک روش تجزیه و تحلیل داده است که با توانایی استخراج اطلاعات و دانش بالقوه مشخص می‌شود. روش‌های تحلیل آماری، روش‌های القایی، روش‌های خوشه‌بندی فضایی، روش‌های تحلیل فضایی، نظریه مجموعه‌های خشن و نظریه مجموعه‌های فازی، الگوریتم‌های داده‌کاوی معروف هستند [۹] و [۱۰]. با توجه به عملکرد عالی آن در پردازش داده‌های عظیم، فناوری داده‌کاوی دارای کاربرد گسترده‌ای است. الگوریتم درخت تصمیم یکی از روش‌های ضروری در میان تجزیه و تحلیل طبقه بندی مبتنی بر داده‌کاوی است. می‌تواند قوانین طبقه بندی مربوطه را بر اساس هر شاخه درخت بدست آورد [۱۱] و [۱۲].

ماروزو و همکاران با توجه به کاربرد فناوری داده‌کاوی در سیستم‌های مدیریتی، از فناوری نرم افزار ابری برای توصیف طراحی و اجرای چارچوب ابری داده‌کاوی استفاده کرد. آنها دریافتند که این سیستم می‌تواند زبان گردش کار بصری را با مدل‌های خدمات نرم افزاری یکپارچه کند تا حجم کار برنامه نویسی را به حداقل برساند [۱۳]. زابورسکی و همکاران از الگوریتم‌های داده‌کاوی، از جمله شبکه‌های عصبی مصنوعی و درختان رگرسیون طبقه بندی، برای تجزیه و تحلیل و تحقیق در مورد پیش بینی پارامترهای تولید مثل استفاده کرد [۱۴]. شکیل و همکاران فرآیند توسعه HRM و اطلاعات مدیریت زنجیره تامین را بررسی کرد. آنها تحقیقات مربوطه در پایگاه داده اسکوپوس را با استفاده از ابزارهای تحلیل علمی تجزیه و تحلیل کردند، نتایج را طبقه بندی و بحث کردند، و اطلاعات HRM را بیشتر مورد بررسی قرار دادند [۱۵].

برای جمع آوری داده‌های بیان شده، نیاز مند سامانه‌هایی برای مدیریت داده‌ها می‌باشیم که در این مقاله نمونه‌ای از این سامانه‌ها معرفی شده است. این سامانه‌ها می‌بایست به موازات پیشرفت‌های تکنولوژی مرتبط با کسب و کار و پاسخ‌گویی به مشکلات، پیشرفت‌های لازم را داشته باشند.

گوناسکاران و رحمان [۱۶] درک درستی از این که چگونه منابع انسانی و مدیریت فناوری اطلاعات باید مکمل یکدیگر باشند برای بهبود ارتباطات و رقابت ارائه می‌دهند. نعمتی و بارکو [۱۷] بر اهمیت تئوری سازمانی که در اجرای داده‌کاوی در یک سازمان و شرکت تفاوت ایجاد می‌کند، تاکید می‌کنند.

در دنیای امروز که دنیایی رقابتی می‌باشد، از دست دادن یک کارمند یا گروهی از کارمندان با عملکرد ضعیف ممکن است به معنای واقعی کلمه بین موفقیت و شکست تفاوت ایجاد کند. با این شرایط، جای تعجب نیست که مدیریت ارشد به داده‌های بیشتری از بخش منابع انسانی خود نیاز داشته باشد. این داده‌ها باید شامل داده‌های گردش مالی، استخدام، مقایسه‌های گرامت و مزایا، توزیع‌های قومی، جنسیتی و فرهنگی باشد.

در این مقاله سعی شده است تکنیک‌های داده‌کاوی بررسی شود تا داده‌ها به اطلاعات مفید تبدیل شوند. علاوه بر توضیح تکنیک‌ها و روش‌ها، به الگوها احتمالی استخراج شده از داده‌کاوی نیز اشاره می‌کند که در تصمیم‌گیری‌های استراتژیک بسیار مهم می‌باشند. هدف از این مقاله، نشان داده نقش داده‌کاوی در جهت بهبود کیفیت فرآیند تصمیم‌گیری در سازمان‌ها و شرکت‌ها در امور می‌باشد. در این مقاله نشان داده شده است که داده‌کاوی با تاثیرگذاری و شکل دادن به استراتژی منابع انسانی تاثیر عمیقی بر موفقیت سازمان‌ها و شرکت‌ها دارد.

در این مقاله، در بخش دوم، کارهای پیشین معرفی شده است. در بخش سوم، مفاهیم مدیریت منابع انسانی که با داده‌کاوی مرتبط هستند شرح داده شده است. در بخش چهارم داده‌کاوی، مزایا و کاربردهای آن مورد بحث قرار داده شده است. در بخش پنجم نمونه سامانه مورد نیاز برای جمع آوری داده‌ها معرفی شده است. در بخش ششم تاثیر داده‌کاوی بر تصمیم‌گیری‌ها مورد بررسی قرار گرفته شده است. در بخش هفتم فرآیند داده‌کاوی و الگوهای قابل اجرا بررسی شده است و در بخش هشتم نیز نتیجه‌گیری نهایی شرح داده شده است.

۲- کارهای پیشین

در این بخش به معرفی چند نمونه از کارهای پیشین و مرتبط با این مقاله پرداخته شده است. برای نمونه اول کار کمالش و کورالمانی [۱۸] سه تکنیک اصلی داده‌کاوی، یعنی ارتباط، خوشه بندی و پیش بینی داده‌های منابع انسانی را مورد بحث قرار می‌دهد. برای نمونه دوم لانگ و تروت [۱۹] اهمیت داده‌کاوی در HRIS را بیان کرده اند. برای نمونه سوم هیرجی [۲۰] به نقش برنامه‌های کاربردی داده‌کاوی برای کسب مزیت استراتژیک و رقابتی اشاره می‌کند. برای نمونه چهارم پایل [۲۱] الگوهای رفتار انسان را تا حدی مورد بحث قرار می‌دهد تا به پیش بینی چگونگی واکنش افراد به رویدادهای خاص کمک کند.

برای نمونه پنجم کودیبا وهاپتروف [۲۲] به طور خلاصه به تکنیک‌های رگرسیون، کاربردهای شبکه عصبی در عملکرد پرسنل و حفظ در منابع انسانی و انتخاب پرسنل اشاره می‌کنند. برای نمونه ششم کواچ و کاتکارت [۲۳] بر اهمیت جمع آوری اطلاعات و تبدیل چنین اطلاعاتی به الگوها برای حمایت از تصمیم‌گیری تاکید می‌کنند. برای نمونه هفتم جانتزکو [۲۴] نیاز به داده‌کاوی برای جذب دانش در جوامع، برای تجزیه و تحلیل همزمانی مفاهیم، استخراج قوانین، هستی شناسی‌ها، شبکه‌های طبقه بندی و غیره شناسایی می‌کند. منبع بالقوه مزیت رقابتی هستند. برای نمونه هشتم تاکدا [۲۵] خط مشی‌های منابع انسانی مختلفی را که توسط مدیران یک سازمان مفهوم سازی شده اند، توصیف می‌کند. برای نمونه نهم شفیلد [۲۶] مدل‌هایی را برای چارچوب دانش جامع در سازمان‌ها پیشنهاد می‌کند. برای نمونه دهم گومز و همکاران [۲۷] بر ارزش اندازه‌گیری عملکرد در سازمان‌ها تاکید

دارند. برای نمونه یازدهم راجاگوپالان و همکاران [۲۸] تاثیر آموزش بر نیروی کار را برای اثربخشی در برنامه‌های کاربردی سازمانی پیشنهاد می‌کنند.

۳- مفاهیم مدیریت منابع انسانی

واکر و ریف [۲۹] بیان می‌کنند که متخصصان منابع انسانی اکنون می‌توانند زمان بیشتری را به تصمیمات استراتژیک تجاری اختصاص دهند، زیرا توسعه فناوری امکان اتوماسیون بسیاری از فرآیندهای HR معاملاتی را فراهم کرده است. داده‌کاوی به دلیل دسترسی گسترده به حجم عظیمی از داده‌ها و نیاز حتمی برای تبدیل چنین داده‌هایی به اطلاعات مفید، توجه زیادی را در جامعه تجاری به خود جلب کرده است [۳۰].

فراتر از انجام تجزیه و تحلیل داده‌ها بر روی مجموعه داده‌های بزرگ، تکنیک‌های داده‌کاوی می‌تواند سازمان‌هایی را که هزاران کارمند و تعداد زیادی اطلاعات مرتبط با شغل دارند، دریایی کند. آن‌ها همچنین می‌توانند الگوهای موجود در پایگاه داده‌های منابع انسانی را با ارائه برخی بینش‌ها مانند جزئیات حفظ کارکنان، برنامه ریزی جبران خسارت، پیش بینی ریسک‌ها و شناسایی فرصت‌ها و ارزیابی‌ها استخراج کنند.

بارنی [۳۱] عقیده دارد که شرکت باید منابع و قابلیت‌های سازمانی را ارائه دهد که معمولاً کمیاب، ارزشمند، غیرقابل تعویض و به طور ناقص قابل تقلید هستند و این منابع مبنای مزیت رقابتی پایدار یک شرکت را تشکیل می‌دهند. باکسال [۳۲] اشاره می‌کند که HRMS با استخدام و توسعه کارمندان مستعد و هم‌افزایی مشارکت آن‌ها در بسته منابع شرکت، یک مزیت رقابتی پایدار برای تصمیم‌گیری ایجاد می‌کند.

تجزیه و تحلیل‌های مختلف با استفاده از داده‌کاوی شامل حرکت و عملکرد کارکنان، فرسایش نیروی کار به فروشگاه، عملکرد نیروی کار به فروشگاه، جبران خسارت، ساییدگی و سایر تحلیل‌ها و گزارش‌های سفارشی شده است. زمانی که یک سازمان و شرکت استراتژی تجاری خود را توسعه می‌دهد، داده‌کاوی بسیار حیاتی و بسیار مهم است. همانطور که یک شرکت استراتژی‌های تجاری خود را شناسایی می‌کند، تکنیک‌های داده‌کاوی موفقیت استراتژی را با ارائه منابع انسانی مناسب و تصمیم‌گیری‌های صحیح با جمع آوری، ذخیره و تجزیه و تحلیل داده‌های مرتبط با وظیفه مناسب برای پیشبرد در فرآیند برنامه ریزی تضمین می‌کند. همه شرکت‌ها اطلاعات مربوط به کارکنان خود را جمع آوری می‌کنند. اما اقدام بر روی داده‌های جمع آوری شده متفاوت است. کوچ و کانتارت [۲۳] پیشنهاد می‌کنند که برخی ممکن است از داده‌ها فقط برای اهداف اداری استفاده کنند، برخی ممکن است این داده‌ها را به اطلاعات مفید تبدیل کنند و برخی ممکن است الگوهای موجود در اطلاعات به دست آمده را برای پشتیبانی در تصمیم‌گیری تجزیه و تحلیل کنند.

۴- داده‌کاوی، مزایا و کاربردها

داده‌کاوی فرآیند استخراج اطلاعات از مجموعه داده‌های بزرگ از طریق استفاده از الگوریتم‌ها و تکنیک‌های استخراج شده از حوزه آمار، یادگیری ماشین و سیستم‌های مدیریت پایگاه داده است [۳۳]. روش‌های سنتی تجزیه و تحلیل داده‌ها اغلب شامل کار دستی و تفسیر داده‌ها می‌شود و بنابراین کند، پرهزینه و بسیار ذهنی هستند [۳۴].

در سامانه‌های منابع انسانی، برای اهداف تصمیم‌گیری، تعداد زیادی از یکپارچگی داده‌ها باید تجزیه و تحلیل شوند. داده‌کاوی تاثیر مستقیمی بر نتایج تحلیلی دارد که تصمیمات تجاری را هدایت می‌کند [۳۵]. مثلاً، ممکن است از طریق ارائه‌های ساده مانند زبانه‌های متقاطع، نمودارهای دو بعدی استاندارد، داده‌ها در الگوها و تعامل با داده‌ها و الگو تا زمانی که از کیفیت پیش‌بینی‌ها مطمئن شویم، با داده‌های منابع انسانی موجود شروع کنیم. در درک کارمند، به دلیل ظهور بسیاری از بخش‌ها، بسیاری از کارمندان و بسیاری از ابزارهای پیچیده، پیچیدگی به طور چشمگیری افزایش یافته است. اما با داده‌کاوی، می‌توانیم بگوییم که چقدر احتمال

دارد که شخصی کاری را که به آن علاقه مندیم، در زمان واقعی انجام دهد. از اطلاعات مفید استخراج شده از پایگاه‌های داده بزرگ برای اتخاذ تصمیمات تجاری حیاتی استفاده می‌کند [۳۶].

کاربردهای HRMS فرصت‌هایی را برای تولید گزارش‌های موقت یا استاندارد از مجموعه‌های خاص داده‌ها فراهم می‌کند. محققان ثابت کرده‌اند که [۲۰] برنامه‌های داده‌کاوی، زمانی که به طور موثر کاشته شوند، می‌توانند منجر به برنامه‌ریزی استراتژیک و مزیت رقابتی شوند.

داده‌کاوی شامل پرس و جویها، گزارش‌ها و ابزارهای پشتیبانی تصمیم‌گیری مختلف بر اساس داده‌های تاریخی سازمان است و حتی می‌تواند چشم انداز آینده را پیش بینی کند. داده‌کاوی [۳۷] را می‌توان بر روی تعدادی پایگاه داده مختلف انجام داد که شامل داده‌های رابطه‌ای، داده‌های تراکنش، انبار داده، داده‌های وب، جریان‌های داده، فایل‌های مسطح، داده‌های سری زمانی و غیره است.

۴-۱- انبارهای داده

انبار داده مخزنی از اطلاعات جمع آوری شده از منابع متعددی است که در یک ذخیره سازی فیزیکی ذخیره شده و در یک سایت قرار دارند. اینمون [۳۸] ویژگی‌های انبار داده را موضوع محور، یکپارچه، غیر فرار و تاریخی تعریف می‌کند. فرض کنید یک سازمان بین‌المللی، با شعبه‌هایی در سراسر جهان، هر شعبه مجموعه‌ای از پایگاه‌های اطلاعاتی رابطه‌ای خاص خود را دارد. رئیس سازمان و شرکت می‌خواهد تحلیل عملکرد کارکنان سازمان و شرکت را برای مدتی بداند. اگر سازمان و شرکت یک انبار داده داشته باشد، این کار آسان خواهد بود.

واتسون و هیلی [۳۹] انبار داده را به عنوان یک مخزن جداگانه و سفارشی داده برای برنامه‌های کاربردی تصمیم‌گیری تعریف می‌کنند. هزینه‌های ساخت انبار داده شامل سخت‌افزار، نرم‌افزار، هزینه‌های پرسنل داخلی، هزینه‌های مشاوران و پیمانکاران، هزینه شبکه‌ها و هزینه آموزش برای میز کمک کاربران فناوری اطلاعات و تجاری، هزینه‌های عملیات و مدیریت سیستم است. از این رو سرمایه گذاری ممکن است برای همه سازمان‌ها مناسب نباشد. انبارهای داده در نظر گرفته شده و به طور گسترده به عنوان منابع مفیدی برای ادغام اطلاعات شرکت، برای تهیه گزارش‌های خلاصه و چند بعدی، برای اطمینان از استقلال کاربر نهایی با انعطاف پذیری بالا در نظر گرفته می‌شوند. بسیاری از سازمان‌ها از اطلاعات انبار داده برای حمایت از فعالیت‌های تصمیم‌گیری تجاری مانند تجزیه و تحلیل عملیات و جستجوی منابع سود، مدیریت روابط کارکنان و افزایش روابط با مشتری استفاده می‌کنند. آگوستا [۴۰] و مالاخ [۴۱] توضیح می‌دهند که وقتی اطلاعات انبار داده فقط یک حوزه عملکردی مانند منابع انسانی را پوشش می‌دهد، به آنها داده‌های مارت می‌گویند. دیتا مارت گیرنده اطلاعاتی است که توسط فرآیند تحویل داده جذب شده و حاوی داده‌هایی از انبار داده است. دیتا مارت از دیدگاه مشترک شرکتی از داده‌های استراتژیک ایجاد شده در انبار داده توسط فرآیند یکپارچه سازی و تبدیل استفاده می‌کند [۴۲].

برای ساده‌سازی و خودکارسازی تراکنش‌ها، می‌توان از داده‌های مارت منابع انسانی استفاده کرد. با در دسترس قرار دادن داده‌ها برای تصمیم‌گیری مؤثر، به ادغام همه برنامه‌های منابع انسانی کمک می‌کند. این داده‌ها کارمند را قادر می‌سازد تا به خودکفایی برسد و با مقامات مربوطه ارتباط مستقیم داشته باشد. این ارتباط کلامی و افقی در سازمان را با خودکارسازی بخش عمده‌ای از کارهای معاملاتی بهبود می‌بخشد. بکارگیری تکنیک‌های داده‌کاوی در سیستم‌های منابع انسانی به اندازه سازمان بستگی دارد. برخی از سازمان‌ها ممکن است کوچک باشند و از این رو می‌توانند سیستم‌های منابع انسانی را به عنوان بخشی از پروژه‌های بزرگ بگنجانند. برخی از سازمان‌ها ممکن است برنامه ریزی منابع سازمانی (ERP) را پیاده سازی کرده باشند که چندین عملکرد سازمان را یکپارچه می‌کند.

۴-۲- داده‌کاوی و آمار

توانایی ساخت یک مدل پیش بینی موفق به داده‌های گذشته بستگی دارد. داده‌کاوی به گونه‌ای طراحی شده است که از موفقیت‌ها و شکست‌های گذشته درس بگیرد و بتواند اتفاقات بعدی را پیش‌بینی کند (پیش‌بینی‌اینده). ممکن است کسی فکر کند که "چرا از داده‌کاوی در سازمان‌ها استفاده کنیم، زمانی که تجزیه و تحلیل آماری در حال انجام است؟" ابزار داده‌کاوی اهمیت آماری الگوها و گزارش‌های پیش بینی شده را بررسی می‌کند.

داده‌کاوی نشان می‌دهد که احتمال دارد اتفاقی بعید رخ دهد. اگر ابزار داده‌کاوی متوجه شود که ۱۰٪ از کارکنان یک گروه بزرگ خاص برای تجزیه و تحلیل عملکرد گنجانده شده‌اند، اما در بین آنها تنها ۱۰ کارمند دارای ویژگی‌های رتبه بندی عملکرد بالا هستند، آنگاه این ابزار می‌تواند هشدار دهد که به احتمال زیاد به جای یک الگوی پیشگویانه معمول، یک ویژگی خاص از پایگاه داده باشد [۴۳].

تفاوت بین داده‌کاوی و آمار در این است که داده‌کاوی فرآیند آماری را که به چندین ابزار نیاز دارد خودکار می‌کند. استنتاج آماری مبتنی بر فرض است به این معنا که یک فرضیه شکل می‌گیرد و در برابر داده‌ها آزمایش می‌شود. در مقابل، داده‌کاوی مبتنی بر کشف است. یعنی فرضیه به طور خودکار از داده‌های داده شده استخراج می‌شود. دلیل دیگر این است که تکنیک‌های داده‌کاوی برای داده‌های آشفته دنیای واقعی قوی‌تر هستند و همچنین توسط کاربران متخصص کمتر مورد استفاده قرار می‌گیرند [۴۴]. داده‌کاوی می‌تواند به سوالات تحلیلی پاسخ دهد: مسائل مربوط به عملکرد کارکنان چیست؟ چه عوامل یا ترکیباتی مستقیماً بر کارکنان تأثیر می‌گذارد؟ بهترین و برجسته‌ترین کارمندان چه کسانی هستند؟ نحوه استخدام کارمندان به چه صورت باشد تا میزان ورود کارمندان به سازمان و شرکت متناسب با امور مالی باشد؟ آیا طرح‌های تشویقی باید به گروه خاصی از کارمندان تعمیم داده شود؟ کدام کارمندان احتمالاً سازمان‌ها را ترک می‌کنند یا احتمالاً ارتقا می‌یابند؟ چگونه می‌توان منابع را به طور بهینه برای اطمینان از اثربخشی و کارایی تخصیص داد؟ و غیره.

روش‌های تحلیل روند و پیش‌بینی در آمار به داده‌های تاریخی نیاز دارند که بدون پایگاه‌های اطلاعاتی جامع به دست آوردن آن‌ها دشوار خواهد بود. تاریخچه‌های فردی کارکنان را می‌توان به عنوان بخشی از تجزیه و تحلیل عملکرد در نظر گرفت. این به ایجاد روابط و مهمتر از همه در هدایت سازمان منابع انسانی از تمرکز اداری به مدیریت استراتژیک کمک می‌کند.

۵- سامانه مدیریت منابع انسانی

HRMS در سازمان‌ها داده‌های اولیه شخصی و مرتبط با کار را به منظور انجام معاملات اولیه منابع انسانی جمع‌آوری می‌کند [۲۳]. هانتر و همکاران [۴۵] عناصر اصلی HRMS شامل کاربرد منابع انسانی، استخدام، حقوق و دستمزد، مدیریت جریان کار، گزارش و تجزیه و تحلیل را ارائه می‌دهد. HRMS می‌تواند بهترین روش‌ها را برای برآوردن اهداف تجاری آتی و نیازهای یادگیری سازمان با کمک راه‌حل‌های پشتیبانی تصمیم [۴۶] ارزیابی کند و کارمندان و مدیران را قادر می‌سازد تا سریع، دقیق و آگاهانه شخصی و تجاری را انجام دهند. تصمیمات فیشر و همکاران [۴۷] اطلاعاتی را در مورد گزارش‌های مربوط به هر فعالیت منابع انسانی مانند استخدام و جابجایی، روابط کار و مزایای جبرانی که به طور منظم ایجاد می‌شود، ارائه می‌دهد. HRIS ستون فقرات عملکرد HRM معاصر است [۴۸]. بارترام [۴۹] دریافت که شرکت‌های کوچک نسبت به شرکت‌های متوسط یا بزرگ احتمال کمتری برای اتخاذ شیوه‌های مدیریت منابع انسانی دارند. جنا و همکاران [۵۰] اهمیت منابع انسانی را در سیستم‌های سازمانی توصیف می‌کنند.

مطالعات همچنین گزارش داده‌اند که برای سازمان‌هایی که HRMS را پذیرفته‌اند، سرمایه‌گذاری‌ها کوچک هستند و معمولاً به جای استراتژی، تاکتیک‌ها و مدیریت را مورد توجه قرار می‌دهند. با این حال، همانطور که سازمان‌ها به دنبال رشد و رقابت در

اقتصاد مبتنی بر دانش به سرعت در حال تحول هستند، فشار برای ایفای نقش استراتژیک‌تری در سازمان منابع انسانی همچنان افزایش می‌یابد [۵۱].

توابع منابع انسانی از اطلاعات برای پشتیبانی و توسعه برنامه‌های منابع انسانی استفاده می‌کنند و همچنین به مدیران نشان می‌دهند که چگونه منابع انسانی می‌تواند به تأثیرگذاری بر نتایج نهایی کمک کند. فیتز-انز [۵۲] نمونه‌هایی را ارائه می‌دهد که شامل حقوق و مزایا به عنوان درصدی از هزینه‌های عملیاتی، هزینه هر استخدام، بازگشت آموزش و هزینه گردش مالی است. بیشتر، داده‌های به دست آمده از این سیستم‌ها فرصتی برای جستجوی روابط و الگوهای اضافی را فراهم نمی‌کند. به عنوان مثال، برای ارزیابی ارزش کارمندی که برای مدت طولانی در سازمان می‌ماند، وظایف منابع انسانی باید به جای ارزیابی هزینه استخدام و غیره، اطلاعاتی در مورد ارزش استخدام ارائه دهند. محققان [۵۳] و [۵۴] و [۵۵] که بر اجرای HRMS در هر سازمانی تأثیر می‌گذارند و در نتیجه، نوآوری‌ها از پیشرفت‌های فناوری در تصمیم‌گیری خبر می‌دهند، موفق نمی‌شوند.

نقش داده کاوی در HRMS ادغام برنامه‌های مختلف منابع انسانی، جمع آوری داده‌ها در نقاط مختلف پایگاه‌های داده منابع انسانی و استفاده از ابزارهای موثر برای در دسترس قرار دادن اطلاعات برای تصمیم‌گیری موثر است. نقش دیگر می‌تواند توسعه و بهبود برنامه‌های کاربردی مبتنی بر اینترنتان برای ساده سازی فرآیندها، برای بهبود پیش بینی مدیریت عملکرد کارکنان، هماهنگی و مدیریت گردش کار باشد. به عنوان مثال، می‌توان یک فرآیند جداگانه برای استخدام و انتخاب، یک فرآیند جداگانه برای سیستم‌های ردیابی برنامه، استخدام‌های داخلی، فرآیند دیگری برای آموزش و توسعه، مدیریت عملکرد و یک فرآیند جداگانه برای توسعه مدیریت ارتباط با کارکنان (ERM) وجود داشته باشد. این سیستم‌ها به ادغام اطلاعات در سراسر سازمان کمک می‌کنند تا همه چیز را از گزارش‌دهی موقت گرفته تا تحلیل روابط پیچیده و روند نیروی کار انجام دهند. این سیستم‌ها در صورت پیاده‌سازی می‌توانند شاخص‌های عملکرد کلیدی از پیش تعریف‌شده‌ای را ارائه دهند که نشانگر داشبوردی از نحوه عملکرد نیروی کار است. سپس، سازمان می‌تواند تأثیر نیروی کار را بر اهداف شرکت ارزیابی کند.

در شکل ۱ و ۲، بخشی از قسمت استخدام HRMS قابل مشاهده می‌باشد که به دریافت بخش‌هایی از اطلاعات فردی افراد سازمان و شرکت می‌پردازد.

The screenshot shows a web-based form for entering personal information. The form is organized into several sections with input fields:

- Identification:** نام پرسنلی (Employee Name), پیش شماره (Prefix Number), تاریخ صدور (Issue Date), آدرس (Address).
- Personal Details:** نام (Name), شماره تلفن (Phone Number), استان تولد (Birth State), جنسیت (Gender).
- Employment/Service:** نام خانوادگی (Family Name), شماره همکار (Employee ID), کد ملی (National ID), وضعیت استخدام (Employment Status), تاریخ تولد (Date of Birth), آخرین شماره سمت (Last Position Number), وضعیت استخدام (Employment Status).
- Accounting/Finance:** شماره حساب بانکی (Bank Account Number), شماره حساب بانکی (Bank Account Number), کد شهر (City Code), نام شهر (City Name).
- Other:** نام پدر (Father's Name), شماره شناسنامه (National ID Number), وضعیت ازدواج (Marital Status), تاریخ تولد (Date of Birth), آخرین شماره سمت (Last Position Number), وضعیت فعلی (Current Status), کد شهر (City Code), نام شهر (City Name).

At the bottom of the form, there is a button labeled "بخش بعدی" (Next Section).

شکل ۱- دریافت اطلاعات فردی در HRMS بخش اطلاعات شخصی

The screenshot shows a web interface for the HRMS system. The main content area is titled 'اطلاعات شخصی' (Personal Information). It contains several input fields and dropdown menus for the following categories:

- اطلاعات شخصی (Personal Information):** نام (Name), نام خانوادگی (Last Name), تاریخ تولد (Date of Birth), شماره شناسنامه (National ID Number), شماره کارت ملی (National ID Card Number), شماره کارت پایان خدمت (Discharge Card Number), شماره کارت بازنشستگی (Retirement Card Number), شماره کارت بازنشستگی (Retirement Card Number), شماره کارت بازنشستگی (Retirement Card Number).
- اطلاعات تحصیلی (Educational Information):** مقطع تحصیلی (Level of Education), گرایش تحصیلی (Specialization), نام دانشگاه (University Name), تاریخ فراغت از تحصیل (Date of Graduation), نام مدرک (Degree Name), نوع مدرک (Degree Type).
- اطلاعات شغلی (Job Information):** تاریخ استخدام (Date of Hire), تاریخ انقضای قرارداد (Contract Expiry Date), نام واحد (Department Name), نام پست (Job Title), نام واحد (Department Name), نام پست (Job Title).

At the bottom of the form, there are buttons for 'بازرسی مجدد' (Review) and 'ثبت مجدد' (Save).

شکل ۲- دریافت اطلاعات فردی در HRMS بخش تحصیلات

در شکل ۳، بخش دیگری از قسمت استخدام HRMS قابل مشاهده می‌باشد که به دریافت بخشی از اطلاعات استخدامی افراد سازمان و شرکت می‌پردازد.

The screenshot shows a web interface for the HRMS system. The main content area is titled 'اطلاعات استخدامی' (Employment Information). It contains several input fields and dropdown menus for the following categories:

- اطلاعات استخدامی (Employment Information):** تاریخ شروع به کار (Start Date), تاریخ شروع به کار (Start Date), تاریخ شروع به کار (Start Date), تاریخ شروع به کار (Start Date), تاریخ شروع به کار (Start Date), تاریخ شروع به کار (Start Date), تاریخ شروع به کار (Start Date), تاریخ شروع به کار (Start Date).
- اطلاعات قرارداد (Contract Information):** تاریخ شروع به کار (Start Date), تاریخ شروع به کار (Start Date), تاریخ شروع به کار (Start Date), تاریخ شروع به کار (Start Date), تاریخ شروع به کار (Start Date), تاریخ شروع به کار (Start Date), تاریخ شروع به کار (Start Date), تاریخ شروع به کار (Start Date).
- اطلاعات شغلی (Job Information):** تاریخ شروع به کار (Start Date), تاریخ شروع به کار (Start Date), تاریخ شروع به کار (Start Date), تاریخ شروع به کار (Start Date), تاریخ شروع به کار (Start Date), تاریخ شروع به کار (Start Date), تاریخ شروع به کار (Start Date), تاریخ شروع به کار (Start Date).

At the bottom of the form, there are buttons for 'بازرسی مجدد' (Review) and 'ثبت مجدد' (Save).

شکل ۳- دریافت بخشی از اطلاعات استخدامی در HRMS

در شکل ۴، بخش دیگری از قسمت استخدام HRMS قابل مشاهده می‌باشد که به دریافت اطلاعات مربوط به انتقالات افراد سازمان و شرکت می‌پردازد.

شکل ۴- دریافت بخشی از اطلاعات انتقالات در HRMS

در شکل ۵، بخش دیگری از قسمت استخدام HRMS قابل مشاهده می‌باشد که به دریافت بخشی از اطلاعات مربوط به کاربایی و تخصیص افراد سازمان و شرکت می‌پردازد.

شکل ۵- دریافت بخشی از اطلاعات کاربایی و تخصیص در HRMS

۶- تاثیر داده‌کاوی بر تصمیم‌گیری‌ها

از دیدگاه کسب و کار، برای بقا و موفقیت در محیط جهانی بسیار رقابتی امروزی، کاربران تجاری پاسخ‌های تجاری را در قالب تصمیماتی که باید به سرعت و با استفاده صحیح از تمام داده‌های موجود اتخاذ شوند، طلب می‌کنند. زمانی که حجم داده‌ها هر سال دو برابر می‌شود، حفظ مزیت رقابتی و حفظ دامنه کسب و کار بسیار دشوار است. از منظر فناوری، واضح است که پهنای باند شبکه در حال افزایش و قیمت پهنای باند در حال کاهش است [۴۳]. قیمت سرعت پردازش رایانه رو به کاهش است و فضای کار از نظر سخت افزار و نرم افزار ناهمگن تر می‌شود، سیستم‌های قدیمی بیشتری در انتظار ادغام هستند [۴۳]. توانایی توانمندسازی سازمان‌ها برای تصمیم‌گیری نیازمند اطلاعات منطقی، خودکار و یکپارچه در سطح سازمانی است که صحیح، مرتبط و کامل باشد. اگر سازمان‌ها داده‌ها را از منظر خرد و کلان درک کنند و از هرگونه اشتباهی که روی داده‌های ناقص و نادرست صورت می‌گیرد اجتناب کنند، می‌توانند چندین ساعت در نیروی انسانی صرفه‌جویی کنند.

زمان آن فرا رسیده است که داده‌ها را با انجام تحلیل گذشته نگر و تحلیل پیش بینی به طور کارآمد تجزیه و تحلیل کنیم. برای انجام این کار، سازمان‌ها باید تمام منابع داده (اعم از خارجی و داخلی) را جمع‌آوری، جمع‌آوری و ادغام کنند، آنها را استخراج کرده و دانش به دست آمده را به گونه‌ای سازماندهی کنند که تصمیم‌گیری پیچیده را تسریع کند. توانایی تمرکز بر فرآیندهای تجاری، انجام تجزیه و تحلیل و تصمیم‌گیری، درک کل سیستم‌ها و سازماندهی داده‌ها در همه سطوح نیاز به پردازش تحلیلی دارد که فراتر از پردازش پرس و جو است. ایجاد پاسخ‌های سریع به پرسش‌های موقت برای آزمایش فرضیه، همبستگی داده‌های فعلی و تاریخی برای تجزیه و تحلیل گزارش‌های تولید شده از طریق سیستم‌های اطلاعات اجرایی (EIS)، همچنین به جای روش‌های پرس‌ومحور معمول، به پردازش اطلاعات نیاز دارد. به عنوان مثال، یک سازمان می‌خواهد دو گروه مختلف کارمند خود را با هم مقایسه کند. این می‌تواند با رویکرد پرس و جو محور برای استخراج چنین کارمندان خاصی که به آن گروه‌ها اختصاص داده شده‌اند انجام شود. اما اگر در حال ارزیابی تحلیل عملکرد باشیم و بخواهیم روندها و توالی الگوها را در کارکنان بدانیم و بر اساس جزئیات تاریخی کارکنان، پیش بینی و پیش بینی انجام دهیم، نیاز به پردازش تحلیلی (کشف دانش) دارد. رویکرد متداول مبتنی بر پرس و جو داده‌ها را واضح‌تر و یا به طور کامل نشان نمی‌دهد.

افزایم و آرونسون [۵۶] اشاره می‌کنند که سرعت سریع فناوری‌های اطلاعاتی جدید مانند سخت‌افزار رایانه‌های شخصی، مجموعه‌های نرم‌افزاری، شبکه‌های کلاینت/سرور، و نسخه‌های رایانه شخصی شبکه‌شده سیستم‌های پشتیبانی تصمیم (DSS) و سیستم‌های اطلاعات اجرایی (EIS) نرم‌افزار پشتیبانی تصمیم‌گیری را برای سطوح پایین‌تر مدیریت، و همچنین برای همه افراد و تیم‌های حرفه‌ای کسب‌وکار خود هدایت می‌کند. DSS شامل فرآیند مدل‌سازی تحلیلی تعاملی است که شامل تجزیه و تحلیل چه می‌شود، تحلیل حساسیت، تجزیه و تحلیل جستجوی هدف و تجزیه و تحلیل بهینه‌سازی است [۵۷].

ما می‌توانیم برخی از مدل‌سازی‌های تحلیلی تعاملی را روی داده‌های منابع انسانی انجام دهیم. ما می‌توانیم روندها و الگوها را در داده‌های کارکنان شناسایی کنیم، که به بهبود بیشتر فرهنگ سازمان و تصمیمات سازمان کمک می‌کند. DSS می‌تواند اطلاعات سایر حوزه‌های عملکردی درون شرکت مانند خدمات مشتری و بازاریابی را برای کمک به حمایت از تصمیمات HRMS ترکیب کند [۵۸].

اگرچه سازمان‌ها روی راه‌حل‌های سازمانی و اندازه‌گیری عملکرد تجاری سرمایه‌گذاری کرده‌اند، این راه‌حل‌ها بر روی دسترسی، تحویل و تجزیه و تحلیل اطلاعات تمرکز دارند. سازمان‌ها به راهی برای جفت کردن راه‌حل‌های تحلیلی پیشرفته، مدل‌های پیش بینی و قوانین تجاری با اتوماسیون فرآیند، سیستم‌های عملیاتی گردش کار و دانش سازمانی جامع نیاز دارند.

اطلاعات تحلیلی بازاریابی شده از طریق داده‌کاوی مفید خواهد بود و ابعاد مختلفی مانند فناوری، بستر دسترسی به داده‌ها، یکپارچه سازی و گزارش دهی و همچنین تصمیم‌گیری را در بر می‌گیرد. همچنین شامل تخصص انسانی است که در قوانین حوزه و کسب و کار گرفته شده است. پس از استفاده از مدل می‌توان آن را اندازه‌گیری کرد.

ما می‌توانیم مدلی بسازیم که کارکنان و ویژگی‌های آنها را که اغلب در تکالیف و وظایف کاری خاص درگیر هستند، شناسایی کند. برخی از تصمیمات شامل مدلسازی و تحلیل ریاضی است که مدیریت باید آنها را اتخاذ کند، به عنوان مثال، اندازه مناسب نیروی کار در یک سازمان تولیدی بزرگ چقدر است؟ چند نفر باید برای رسیدگی به یک هجوم جدید تقاضا یا یک قرارداد بزرگ جدید استخدام شوند؟ این امر مستلزم اعتقاد دقیق ریاضی است زیرا عوامل مؤثر بر بهره‌وری معمولاً چندان شناخته شده نیستند. یک مدل تحلیلی و ریاضی به درستی طراحی شده، با استفاده از داده‌های کافی مناسب می‌تواند به مدیران کمک کند تا تصمیمات بهتری در مورد سطح منابع انسانی مورد نیاز در بازاری که با تغییرات قابل توجه در تقاضا مشخص می‌شود، اتخاذ کنند. به طور کلی، الگوهای تولید شده از طریق داده‌کاوی باید درک شوند، باید روی داده‌های جدید یا آزمایشی با درجه‌ای از اطمینان معتبر باشند و باید بالقوه مفید و جدید باشند. اگر کسی به اندازه کافی در هر مجموعه داده جستجو کند، می‌توان الگوهای را یافت که به نظر می‌آیند اما لزوماً از نظر آماری معنی‌دار یا مفید نیستند [۵۹]. اگر یک الگوی فرضیه‌ای را که کاربر تأیید می‌کند، جالب خواهد بود. داده‌کاوی با رتبه بندی الگوهای کشف شده بر اساس جذابیت آنها و فیلتر کردن الگوهای غیر جالب مرتبط است [۶۰].

۷- فرآیند داده‌کاوی و الگوهای قابل اجرا

سازمان‌ها بر مدل‌های انتخاب سیستمی و تصمیم‌محور تکیه می‌کنند که تصمیم‌گیرنده را قادر می‌سازد تا بازدهی را که انتظار می‌رود از اجرای یک برنامه انتخاب پیشنهادی حاصل شود، ارزیابی کند. چنین مدل‌هایی فراتر از بررسی اندازه ضریب اعتبار هستند و مجموعه‌ای از مسائل مانند بودجه بندی سرمایه و نتایج استراتژیک در سطوح گروهی و سازمانی را در بر می‌گیرند. بسیاری از سازمان‌ها کوه‌هایی از داده‌ها را درباره اطلاعات کارکنان، سیاست‌ها، گزارش‌های عملکرد، گزارش‌های منابع انسانی، داده‌های حقوق و دستمزد، مزایای جبرانی، تسهیلات بازنشستگی، روش‌های مشاوره و غیره تولید می‌کنند. این داده‌ها یک منبع استراتژیک برای سیستم‌های منابع انسانی است. در حال حاضر، استفاده از اکثر این منابع استراتژیک منجر به بهبود کیفیت سیستم‌های منابع انسانی می‌شود.

فلدرز و همکاران [۳۳] شش مرحله مهم در فرآیند داده‌کاوی را بیان کرده‌اند که در جدول ۱ قابل مشاهده می‌باشد.

جدول ۱- شش مرحله مهم در فرآیند داده‌کاوی

شماره مرحله	نام مرحله
مرحله اول	تعریف مشکل
مرحله دوم	کسب دانش
مرحله سوم	انتخاب داده‌ها
مرحله چهارم	پیش پردازش داده‌ها
مرحله پنجم	تحلیل و تفسیر
مرحله ششم	گزارش و استفاده

برتولد و هند [۶۱] فرآیند داده‌کاوی را شناسایی کرده‌اند که در جدول ۲ قابل مشاهده می‌باشد.

جدول ۲- فرآیند داده‌کاوی

شماره مرحله	نام مرحله
مرحله اول	تعریف اهداف تحلیل
مرحله دوم	انتخاب و پیش تصفیه داده‌ها
مرحله سوم	مشخصات روش‌های آماری
مرحله چهارم	تحلیل توضیحی
مرحله پنجم	تجزیه و تحلیل داده‌ها
مرحله ششم	ارزیابی و مقایسه روشها
مرحله هفتم	تفسیر مدل انتخاب شده.

تکنیک‌ها و روش‌ها و یا الگوها در داده‌کاوی برای درک بهتر نیاز به توضیح مختصری دارند که در ادامه مقاله توضیح داده شده است.

۷-۱- انجمن‌ها، استخراج الگوهای مکرر

این روش‌ها قواعد قرابت را در بین مجموعه‌ها مشخص می‌کنند. هند و همکاران [۶۲] اشاره می‌کند که الگوها اغلب در طول فرآیند داده‌کاوی رخ می‌دهند. عواقب بعدی مانند حرکت کارکنان و نگرش کاری که اغلب رخ می‌دهد، در عملکرد آنها الگوهایی دارد. مثال دیگر در HRMS می‌تواند فرم‌های ادعایی ارسال شده توسط کارکنان به بیمه درمانی سازمان باشد. هر ادعا شامل مجموعه‌ای از اقدامات پزشکی است که در طی یک بازدید بر روی کارمند معین انجام شده است. با تعریف مجموعه‌ای از موارد به عنوان "مجموعه تمام اقدامات پزشکی قابل ارسال" و سوابق مربوط به هر فرم ادعا، برنامه می‌تواند با استفاده از تابع ارتباط، روابط بین روش‌های پزشکی را که اغلب با هم ترکیب می‌شوند پیدا کند. کاربردهای قوانین انجمن شامل تجزیه و تحلیل سبد بازار، ارسال پستی پیوست در بازاریابی مستقیم، کشف تقلب، برنامه ریزی طبقه/قفسه فروشگاه و غیره است.

۷-۲- طبقه‌بندی و پیش‌بینی

مدل‌های طبقه‌بندی و پیش‌بینی دو تکنیک تجزیه و تحلیل داده‌ها هستند که برای توصیف کلاس‌های داده و پیش‌بینی کلاس‌های داده‌ایند استفاده می‌شوند. طبقه‌بندی یکی از روش‌های داده‌کاوی است که برای پیش‌بینی چگونگی عضویت نمونه‌های داده در گروه‌های گوناگون استفاده می‌شود. در طبقه‌بندی هر نمونه از چند ویژگی تشکیل شده است که یکی از آنها با عنوان ویژگی دسته شناخته می‌شود. در روش‌های طبقه‌بندی الگوریتم در ابتدا در مرحله آموزش یک مدل یاد می‌گیرد و سپس در مرحله ارزیابی،

کارایی مدل فرا گرفته شده بررسی می‌شود. یک شرکت کارت اعتباری که سابقه اعتباری مشتری آن مشخص است، می‌تواند سابقه مشتری خود را به عنوان خوب، متوسط یا ضعیف طبقه بندی کند. به طور مشابه، سطوح درآمد مشتری را می‌توان به عنوان بالا، پایین و متوسط طبقه بندی کرد. پیتر و دولف [۶۳] توضیح می‌دهند که اگر سوابقی حاوی رفتار مشتری داشته باشیم و بخواهیم داده‌ها را طبقه بندی کنیم یا پیش‌بینی کنیم، متوجه می‌شویم که وظایف طبقه بندی و پیش‌بینی بسیار نزدیک به هم مرتبط هستند. پیش‌بینی اینکه آیا یک کارمند خاص نوع خاصی از رفتار را نشان می‌دهد یا خیر، مستلزم این فرض است که کارمند به نوع خاصی از گروه کارمند تعلق دارد و بنابراین نوع خاصی از رفتار را نشان می‌دهد.

مدل‌های درخت تصمیم، طرح‌های طبقه بندی مبتنی بر شبکه‌های عصبی برای تصمیم‌گیری‌های تجاری بسیار مفید هستند. طبقه بندی روی داده‌های گسسته و نامرتب کار می‌کند، در حالی که پیش‌بینی روی داده‌های پیوسته کار می‌کند. رگرسیون اغلب استفاده می‌شود زیرا یک روش آماری است که برای پیش‌بینی عددی استفاده می‌شود. هنگام تصمیم‌گیری برای انتخاب کارمند، انتخاب افراد برای پر کردن مشاغل موجود تنها زمانی واضح و معنادار می‌شود که متقاضیان بیش از مشاغل باشند. تاکید اولیه باید بر دقت اندازه‌گیری انتخاب و کارایی پیش‌بینی شود. رگرسیون ساده یا چندگانه مدل پیش‌بینی اساسی است که تصمیم‌گیرنده را قادر می‌سازد تا وضعیت معیار هر فرد را بر اساس اطلاعات پیش‌بینی کننده پیش‌بینی کند.

رویه‌هایی برای انتخاب امتیازهای برش مناسب در فرآیندهای انتخاب کارکنان وجود دارد. جیمز [۶۴] و پرل [۶۵] مقدمه‌ای بر رگرسیون و مدل‌های خطی لگ ارائه می‌دهند. اسمیت و گوپتا [۶۶] از طریق مطالعات موردی نشان می‌دهند که چگونه فناوری شبکه‌های عصبی در حوزه‌های مختلف کسب و کار مفید است. ما بحث خود را در مورد الگوریتم‌ها و اثبات در اینجا محدود کردیم.

۷-۳ - خوشه بندی

این روشی است که توسط آن رکوردهای مشابه با هم گروه بندی می‌شوند. خوشه بندی معمولاً به معنای تقسیم بندی استفاده می‌شود. الگوریتم خوشه بندی اطلاعاتی را که ویژگی‌های نزدیک به هم و مشابه دارند را در دسته‌های جداگانه که به آن خوشه گفته می‌شود قرار می‌دهد. به بیان دیگر خوشه بندی همان دسته بندی‌های ساده‌ای است که در کارهای روزانه انجام می‌دهیم. وقتی با یک مجموعه کوچک از صفات روبرو باشیم دسته بندی به سادگی قابل اجرا است. یک سازمان می‌تواند سلسله مراتب کلاس‌هایی را که رویدادهای مشابه را گروه بندی می‌کنند، انتخاب کند. با استفاده از خوشه بندی، کارکنان را می‌توان بر اساس درآمد، سن، شغل، مسکن و غیره گروه بندی کرد. در تجارت، خوشه بندی به شناسایی گروه‌هایی از شباهت‌ها، مشخص کردن گروه‌های مشتریان بر اساس الگوهای خرید و غیره کمک می‌کند. یک سازمان ممکن است کارکنان را تقسیم بندی کند و خوشه‌های بالقوه را کشف کند. کارکنان بر اساس ویژگی‌ها (نوع سمینارها، برنامه‌های آموزشی شرکت کرده و تعداد مرخصی‌های استفاده شده و غیره).

۷-۴ - اهداف

برای فرآیند داده‌کاوی در منابع انسانی، داده‌های جامع منابع انسانی باید در دسترس باشد. تعیین دقیق سؤالات، دامنه، هدف، داده‌های مربوط به کار و توانایی تفسیر نتایج مورد نیاز است. بوردلی و همکاران [۶۷] پیشنهاد می‌کنند که انتخاب الگوریتم داده‌کاوی به استفاده مورد نظر از دانش استخراج شده بستگی دارد. تخصیص نیروی انسانی بر اساس پیش‌بینی تقاضا مستلزم داده‌های دقیق مربوط به وظایف است. با توجه به تغییرات فصلی تقاضا، می‌توان نیروی انسانی موقت را برای حفظ سطح خدمات استخدام کرد. سطوح تقاضا نیز در یک روز کاری متفاوت است که می‌توان از آن برای تخصیص منابع استفاده کرد.

یکی دیگر از عوامل مهم در فرآیند داده‌کاوی، تفسیر پرسشنامه‌های مربوط به داده‌های کیفی کارکنان است. به عنوان مثال، با استفاده از ابزار داده‌کاوی، ممکن است الگویی مربوط به کارمندی که بالاترین سطح حقوق و طول خدمت را دارد، شناسایی شود. تفسیر ممکن است متفاوت باشد. آیا می‌توان نتیجه گرفت که اگر کارمندان طولانی مدت بمانند حقوق بیشتری خواهند داشت یا فقط برای بسته بالاتر بیشتر می‌مانند؟ دانش دامنه همراه با تخصص برای تفسیر علل و اثرات روابط ایجاد شده توسط این الگوها

مورد نیاز است. از آنجایی که هر شرکت داده‌های منابع انسانی خود را دارد، روابط جالب در الگوها به رقابت شرکت در بازار، مرحله چرخه رشد، ادغام برنامه‌های مختلف منابع انسانی و غیره بستگی دارد.

۷-۵- پیش پردازش داده‌ها

پیش پردازش داده‌ها یکی از چالش‌های اصلی در داده‌کاوی است. سازمان دارای انبار داده از ابزارهای Extract, Transform, Load (ETL) برای پیش پردازش داده‌ها استفاده می‌کند. در طی این فرآیند، داده‌ها از پایگاه‌های داده عملیاتی استخراج شده، به انبار داده تبدیل شده و سپس در انبار داده بارگذاری می‌شوند. ETL اغلب ترکیبی پیچیده از فرآیند و فناوری است که بخش قابل توجهی از تلاش‌های توسعه انبار را مصرف می‌کند و به مهارت‌های تحلیلگران تجاری، طراحان پایگاه داده و توسعه‌دهندگان اپلیکیشن نیاز دارد.

اگر داده‌ها بدون هیچ تحلیل خاصی در ذهن جمع‌آوری شوند، ممکن است کیفیت داده‌ها ناکافی باشد [۳۳]. یک فرآیند دقیق با توجه به داده‌های گم شده یا پر سر و صدا مورد نیاز است. معمولاً داده‌های منابع انسانی برای اهداف اداری (مثلاً پردازش حقوق و دستمزد) جمع‌آوری می‌شوند، اما اگر داده‌های منابع انسانی را برای تصمیم‌گیری استراتژیک جمع‌آوری و پردازش کنیم، داده‌ها باید مرتبط، دقیق، بدون هیچ ارزش گمشده، غیر زائد، کامل و سازگار باشند. برای نشان دادن اینکه آیا برنامه آموزشی جدید عالی و برتر از روش‌های قبلی یا موجود برای دستیابی به اهداف سازمان است، شرکت‌ها به داده‌های سیستماتیک و کامل نیاز دارند تا اثرات داده‌های مستقل را ارزیابی کنند که احتمالاً بر نتایج آموزشی تأثیر می‌گذارد، مانند روش‌های مختلف آموزشی، اعماق مختلف آموزش، یا انواع مختلف رسانه برای ارائه آموزش. اگر کارمندان انواع خاصی از داده‌ها را وارد نکرده باشند و اگر کارمندان فکر کنند که ممکن است به آنها مربوط نباشد، نادیده گرفته شوند، آنگاه منجر به گم شدن داده‌ها می‌شود و در نتیجه الگوهای پنهان جالبی که تولید می‌شوند ممکن است نادرست باشند.

داده‌های ناقص، پر سر و صدا و متناقض زمانی که اندازه پایگاه داده بزرگ است رایج است. دلایل متعددی برای این امر وجود دارد. داده‌های مورد علاقه کارکنان ممکن است در دسترس نباشد، یا کارمندان ممکن است آن‌ها را صرفاً به این دلیل که در زمان ورود مهم تلقی نمی‌شود، لحاظ نکرده باشند. داده‌های مربوطه ممکن است به دلیل سوء تفاهم ثبت نشوند. خطاهای انسانی و رایانه‌ای در ورود داده‌ها نیز وجود داشته است. کارکنانی که مایل به ارائه جزئیات نیستند، آدرس‌های قدیمی، گزینه‌های زیاد برای سؤالات، مشکل شناسایی موجودیت از منابع متعدد همه نمونه‌هایی از تشخیص مغایرت داده‌ها هستند.

روش‌های جمع‌آوری داده‌ها ممکن است معیوب باشند، بنابراین مقادیر نادرست نوین‌دار در داده‌ها ایجاد می‌شود. لیتل [۶۸]، دیگل و کنوارد [۶۹] و هارتلی و هاکینگ [۷۰] چندین مدل برای تجزیه و تحلیل آماری داده‌های از دست رفته ارائه می‌دهند. رامان و هلرستین [۷۱] توصیف روشی از ابزار پاکسازی داده‌های در دسترس عموم ارائه می‌دهند. همه سازمان‌ها نیازی به ساخت انبار داده ندارند زیرا ممکن است مقرون به صرفه نباشد. مرحله پاکسازی و پیش پردازش داده‌ها شامل حذف نویز و تصمیم‌گیری در مورد استراتژی‌هایی برای مدیریت داده‌های از دست رفته است [۳۴].

۷-۶- استفاده داده‌کاوی در HRMS

همه می‌دانند که داده‌های نرم روی میز برنده نمی‌شوند، اما داده‌های سخت برنده هستند. حرکتی در حال انجام است تا HRMS استراتژیک‌تر شود و بیشتر به داده‌های قابل اندازه‌گیری توجه کند.

هر یک از کارکنان دارای مهارت‌ها و دانش‌هایی هستند که ارزش بزرگ و منحصر به فردی دارند - ارزشی که شرکت باید آن را حفظ کند. در بازار کار فشرده امروزی، هزینه‌های سخت جایگزینی یک کارمند در حال خروج به اندازه کافی بالاست، زیرا تبلیغات، استخدام و آموزش گران هستند. در دنیای رقابتی جهانی، از دست دادن یک کارمند یا تیمی از کارمندان با عملکرد ضعیف ممکن است به معنای واقعی کلمه بین موفقیت و شکست تفاوت ایجاد کند. با این شرایط، جای تعجب نیست که مدیریت ارشد به اطلاعات

بیشتر و بیشتری از بخش منابع انسانی خود نیاز داشته باشد. آنها خواستار داده‌های سخت، اطلاعات در مورد گردش مالی، استخدام، مقایسه غرامت و مزایا، توزیع‌های قومیتی، جنسیتی و فرهنگی هستند [۷۲].

می‌توان به گزارش‌های انبوه در مورد موقعیت‌های خالی، الگوهای حقوق، جزئیات عملکرد و غیره دسترسی داشت. اگر سازمانی بخواهد بداند چه نوع تأثیری بر شرکت دارد، می‌تواند بودجه آموزشی کلی را شناسایی کند. داده‌کاوی می‌تواند الگوهای کارکنان موفق را که در برنامه‌های آموزشی شرکت کرده‌اند، پیشرفت بیشتر آنها پس از برنامه‌های آموزشی، عملکرد کاری آنها، ارتقاء و پیشرفت آنها پس از برنامه‌های آموزشی و غیره را شناسایی کند. کارمندان هانتر و همکاران [۴۵] مفهوم تحلیل منابع انسانی را ذکر کرد. ابزارهای داده‌کاوی بر روی داده‌های تاریخی تمرکز می‌کنند [۴۵] و رویدادهای گذشته را گزارش می‌کنند، تلاش می‌کنند بر پیش‌بینی روند تمرکز کنند تا مناطقی را که نیاز به توجه مدیریت دارند برجسته کنند. چنین رویکردی زمانی که با یک رویکرد کارت امتیازی متوازن برای تعریف و نظارت بر شاخص‌های عملکرد کلیدی در منابع انسانی ترکیب می‌شود، می‌تواند بسیار قدرتمند باشد، اگرچه هزینه‌های زیادی در راه‌اندازی و نگهداری چنین ابزارهایی وجود دارد.

پژوهشگران نشان داده‌اند که بین نگرش کارکنان و عملکرد کارکنان رابطه معناداری وجود دارد [۷۳] و [۷۴]. صدیق [۷۵] اشاره می‌کند که در محیط رقابتی امروزی که لازم است سازمان‌ها شیوه‌های موثر مدیریت منابع انسانی داشته باشند، HRIS می‌تواند از تلاش‌های استخدام، انتخاب و آموزش سازمان‌ها پشتیبانی کند و همچنین نظارت کند که چه کسی و به چه دلیل سازمان را ترک می‌کند.

داده‌کاوی می‌تواند تشخیص دهد که کدام گروه‌ها موفق‌ترین گروه‌ها هستند، کدام دفاتر و مکان‌ها به آموزش خوب پاسخ داده‌اند، کدام کارمند به چه آموزشی پاسخ می‌دهد، و غیره. حفظ یک کارمند موجود مقرون به صرفه‌تر از به دست آوردن یک کارمند جدید است. برای توسعه برنامه‌های موثر حفظ کارکنان، تجزیه و تحلیل دلایل فرسایش کارمندان ضروری است.

از طریق روش‌های داده‌کاوی بر روی داده‌های منابع انسانی، می‌توان تراکنش‌های فردی و عوامل موثر بر یک کارمند را اندازه‌گیری کرد. در سازمان‌ها، داده‌کاوی دانش عمیقی را در بین فرآیندهای منابع انسانی مانند ارزیابی کارکنان، تکنیک‌های مشاوره و مدیریت عملکرد برای تصمیم‌گیری‌های استراتژیک موثر و کارآمد فراهم می‌کند.

به عنوان مثال، معمولاً تأثیرات یک تغییر، یک برنامه آموزشی یا افزایش دستمزد برای گروهی از کارمندان از لحاظ تاریخی بر اساس نظرات تحصیل کرده و نه بر اساس معیارهای آماری پیش‌بینی شده است. اغلب، این نظرات نادرست به نظر می‌رسد. هنگامی که این اتفاق می‌افتد، سرمایه‌گذاری قابل توجهی انجام شده است که هیچ منفعتی برای شرکت ندارد. کارگران تازه آموزش دیده ممکن است برای پیوستن به مسابقه ترک کنند. افزایش حقوق ممکن است تأثیری بر گردش مالی نداشته باشد. برای پاسخگویی به تقاضاهای جدید و افزایش دید مدیریت منابع انسانی، HRMS با روی آوردن به روش‌های داده‌کاوی به دنبال نقشی استراتژیک است.

جدول ۳ الگوهایی را نشان می‌دهد که می‌توانند با استفاده از داده‌کاوی در HRMS تولید شوند.

جدول ۳- الگوهایی که می‌توان با استفاده از داده‌کاوی تولید کرد

تکنیک‌های اصلی داده‌کاوی	الگوها
خوشه بندی	کارکنانی که خصوصیات مشابهی دارند گروه بندی کارکنان برتر تحلیل فرسایش نیروی کار گروه‌هایی از کارمندان به احتمال زیاد ترک می‌کنند

<p>بخش‌های مربوط به تسهیلات بازنشستگی، غرامت، مزایا و کمک هزینه‌ها</p>	
<p>پیش بینی نتیجه یادگیری کارکنان در یک شرکت پیش بینی احتمال موفقیت در فرآیند منابع انسانی پیش بینی درصد دقت در عملکرد کارکنان تجزیه و تحلیل اطلاعات پیش بینی و مدل برای تعیین کمیت دارایی‌های سرمایه انسانی طبقه بندی فرآیند استخدام پیش بینی نرخ فرسودگی کارکنان پیش‌بینی اینکه چه نوع کارمندانی بیشتر به احتمال زیاد ترک می‌کنند، احتمالاً چه کار می‌کنند و احتمالاً وظایف خاصی را انجام می‌دهند. پیش بینی رفتار و نگرش کارکنان پیش بینی پیش بینی تقاضا بر اساس تغییرات فصلی پیش بینی پیشرفت عملکرد در طول دوره عملکرد شناسایی بهترین مشخصات برای کارمندان مختلف پیش بینی برای یافتن عواملی که باعث جذب کارمندان جدید می‌شود طبقه بندی روند حرکت در سازمان برای مجموعه کارکنان موفق / ناموفق (مفید در مدیریت استخدام) اطلاعات مربوط به کارمندانی که از منبع خاصی استخدام شده‌اند دارای ارزش بالاتری هستند پیش‌بینی سریع سوابق کارمندان شرکت‌های بزرگ دسته بندی افراد ترک بر اساس نوع شغل نمرات کارکنان فردی در گروه خطر پیش‌بینی می‌شود که داوطلبانه آن را ترک کنند</p>	<p>طبقه‌بندی و پیش‌بینی</p>
<p>انجمن آموزش انجام شده با انواع مختلف کارکنان و نمرات عملکرد، به صورت فردی و تیمی. ارتباط پروفایل‌های کاری کارمندان به مناسب‌ترین برنامه‌ها/شغل انجمن رویکردهای تیم سازی و رهبری کارکنان ارتباط نگرش کارکنان با عملکرد انجمن و تجزیه و تحلیل حرکت کارکنان کاربرد مکانیزم ردیابی در فرآیند استخدام شناسایی الگوها در انتخاب مزایای کارکنان ارتباط اهداف استراتژیک شرکت با نتایج عملکرد کارکنان</p>	<p>انجمن</p>
<p>استانداردسازی روش‌های آموزشی، نظارت بر عملکرد وظایف در مدیریت شغلی، نظارت بر وظایف مسیر گردش کار. ارائه اطلاعات خلاصه: گزارش‌های خلاصه چند بعدی مختلف. اطلاعات خلاصه آماری (گرایش و تغییرات مرکزی داده‌ها) استفاده از داده‌های تاریخی برای ساخت مدل‌های رفتار متقلبانه و استفاده از داده‌کاوی برای کمک به شناسایی موارد مشابه. خلاصه‌ای از جمعیت کارکنان، غرامت، نرخ گردش مالی و غیره عناصر HRM</p>	<p>داده‌کاوی با استفاده از روش‌های میان رشته‌ای دیگر</p>

۷-۷- محدودیت‌ها

همه HRMS ویژگی‌های تحلیلی یکپارچه مورد نیاز برای برنامه ریزی جبران خسارت و پشتیبانی تصمیم را ارائه نمی‌دهند، مانند توانایی مشاهده حقوق و دستمزدهای بازار رقابتی در زمان واقعی مرتبط [۷۶]. در عوض، آنها بر مدیریت پایگاه داده و نگهداری سوابق مربوط به فعالیت‌های جبرانی مانند حقوق و دستمزد، افزایش حقوق شایستگی و ثبت نام مزایا تمرکز می‌کنند. مهاجرت و استفاده مؤثر از قابلیت‌های طراحی و تجزیه و تحلیل جبران یکپارچه مبتنی بر وب هنوز در مراحل اولیه است [۷۶]. استفاده از نرم‌افزار تحلیلی منابع انسانی در تمام شرکت‌ها در سراسر جهان بسیار کم است. بررسی دقیق سازماندهی داده‌های منابع انسانی مورد نیاز است زیرا داده‌های ناکافی و ناقص ممکن است منجر به الگوهای اشتباه و پوچ در داده‌کاوی شود. هیچ ابزار قدرتمندی در داده‌کاوی برای HRIS با مجموعه کاملی از ویژگی‌های تحلیلی یکپارچه مورد نیاز برای پشتیبانی تصمیم به آسانی در دسترس نیست. متخصصان منابع انسانی از طریق برنامه‌های نرم افزاری اضافی یا برنامه‌های مستقل می‌توانند عملکردهای سطح بالاتری را انجام دهند که نشان دهنده فعالیت‌های استراتژیک تری است.

۸- نتیجه

در این مقاله، ابزارهای فعال داده‌کاوی برای کسب داده برای تبدیل اطلاعات به دانش عملی در HRMS را بررسی کردیم. ما بحث کردیم که چگونه داده‌کاوی سازمانها را قادر می‌سازد تا به طور موثرتری داده‌های منابع انسانی را بدست آورند، کدگذاری کنند، تجمیع و تخصیص دهند. از طریق این بحث، نقش HRMS و نقش داده‌کاوی در تصمیم‌گیری را بررسی کردیم. از این مبنای نظری، ما پیشنهاداتی را در مورد اینکه آیا قابلیت‌های داده‌کاوی باید منجر به افزایش عملکرد یا کاهش متناظر در توانایی شرکت برای حفظ مزیت رقابتی شود، ایجاد کردیم. برای تسهیل این امر، مقاله مروری کوتاه بر تحقیقات در مورد داده‌کاوی در زمینه HRMS ارائه کرد.

مطالعه محدودی از ادبیات رشته‌های مختلف دانشگاهی تلاش کرده است تا مسائل مربوط به داده‌کاوی در HRMS را بررسی و به آن بپردازد. ما پیشنهاد می‌کنیم که یک رویکرد راه‌حل ممکن برای درک بهتر اینکه چگونه داده‌کاوی می‌تواند قابلیت‌های HRMS را تقویت یا فعال کند، اتخاذ روش ارائه شده توسط کمالش و کورالمانی [۱۸] و لانگ و تروت [۱۹] است. علیرغم کار پیش رونده خوب روی موضوعات داده‌کاوی و HRMS، مطالعات کمی به طور خاص یا موثر به انتقال داده‌های منابع انسانی با استفاده از داده‌کاوی به اطلاعات مفید پرداخته اند.

با همسویی اهداف مدیریت دارایی‌های انسانی با اهداف کلی شرکت، تاکید بیشتری بر حوزه منابع انسانی بر روی نتایج نهایی مانند افزایش سود و کاهش هزینه‌های عملیاتی وجود دارد. داده‌کاوی به سرعت به عنوان ابزاری ظاهر شده است که می‌تواند به سازمانها امکان بهره برداری از دارایی‌های اطلاعاتی خود را بدهد.

در این مقاله، ما دیدگاهی تفسیری از فرآیندی اتخاذ کرده‌ایم که توسط آن سازمانها داده‌های حیاتی منابع انسانی خود را معنا می‌کنند. ما فرآیند داده‌کاوی شناسایی، رسمی‌سازی، آزمایش و ارزیابی HRMS را با استفاده از تکنیک‌های رایج داده‌کاوی پیشنهاد کرده‌ایم و آن را با برخی الگوهای نمونه نشان داده‌ایم. ما رویکردها را مورد بحث قرار داده‌ایم و به چالش‌ها و پیامدهای پیش‌پردازش داده‌های منابع انسانی اشاره کرده‌ایم.

سازمانها باید از داده‌های منابع انسانی برای تصمیم‌گیری استراتژیک استفاده کنند. با اجرای روش‌های داده‌کاوی، سازمانها می‌توانند اطلاعات و الگوهای مفید را برای دستیابی به مزیت رقابتی تغییر دهند. ما سعی کردیم الگوهای مختلفی را توضیح دهیم که امکان تولید در HRMS با استفاده از داده‌کاوی وجود دارد. بررسی دقیق و مناسب و تهیه پایگاه داده منابع انسانی و به کارگیری تکنیک‌های داده‌کاوی فرصتی را برای شرکت‌ها ایجاد می‌کند تا رقابت پذیری و تصمیم‌گیری خود را توسعه دهند. داده‌کاوی فرصتی را در HRMS برای مدیریت و کاوش مؤثر و کارآمد داده‌ها فراهم می‌کند.

با درک الگوها می‌توان پیش‌بینی کرد که کدام فعالیت منابع انسانی احتمالاً اتفاق می‌افتد، چه فعالیتی احتمالاً موفق خواهد بود، و غیره. و رویه‌های تصمیم‌گیری پیشرفته تری دارند. نتیجه نهایی، بهبود استراتژی‌های کلی تصمیم‌گیری در سازمان خواهد بود. نتیجه می‌گیریم که داده‌کاوی در HRMS درک بهتری از روابط کارکنان و بهبود رضایت کارکنان و احتمال بیشتری برای دستیابی به مزیت رقابتی می‌دهد.

مراجع

- [1] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, "From data mining to knowledge discovery in databases" *AI magazine*, vol. 17, Sep. 1996, pp. 37-54, doi: 10.1145/240455.240463.
- [2] E. W. T Ngai, L. Xiu, and D. C. K. Chau, "Application of data mining techniques in customer relationship management: a literature review and classification," *Expert Systems with Applications*, vol. 36, Mar. 2009, pp. 2592-2602, doi: 10.1016/j.eswa.2008.02.02.
- [3] A. K. Choudhary, J. A. Harding, and M. K. Tiwari, "Data mining in manufacturing: a review based on the kind of knowledge," *Journal of Intelligent Manufacturing*, vol. 20, Oct. 2009, pp. 501-521, doi: 10.1007/s10845-008-0145-x.
- [4] E. W. T. Ngai, Y. Hu, Y. H. Wong, Y. Chen, and X. Sun, "The application of data mining techniques in financial fraud detection: a classification framework and an academic review of literature" *Decision Support Systems*, vol. 50, Feb. 2011, pp. 559-569, doi: 10.1016/j.dss.2010.08.006.
- [5] Piwowar-Sulej K, Bk-Grabowska D. Enterprise Boundaries in The Area of Human Resources. *Argum. Oecon.*, 2018; 2(41):357-389.
- [6] Huo W, Li X, Zheng M, Liu Y, Yan J. Commitment to Human Resource Management of the Top Management Team for Green Creativity. *Sustainability*, 2020; 12(3):1008.
- [7] Baral S, Subedi H, Paudel P, Chand PB, Shrestha MP, McCullough A, et al. Implementation research to assess a health workers performance—based management system in Nepal. *Acta Paediatr.*, 2018; 107:24-34. pmid:30570792
- [8] Singer G, Cohen I. An Objective-Based Entropy Approach for Interpretable Decision Tree Models in Support of Human Resource Management: The Case of Absenteeism at Work. *Entropy*, 2020; 22(8):821. pmid:33286593
- [9] Zaeem RN, German RL, Barber KS. PrivacyCheck: Automatic Summarization of Privacy Policies Using Data Mining. *ACM T. Internet Techn.*, 2018; 18(4):1-18.
- [10] Ozkaya G, Timor M, Erdin C. Science, Technology and Innovation Policy Indicators and Comparisons of Countries through a Hybrid Model of Data Mining and Operation Research Methods. *Sustainability*, 2021; 13(2):1-52.
- [11] Park S, Hamm SY, Kim J. Performance Evaluation of the GIS-Based Data-Mining Techniques Decision Tree, Random Forest, and Rotation Forest for Landslide Susceptibility Modeling. *Sustainability*, 2019; 11(20):5659.
- [12] Shaheen M, Zafar T, Khan SA. Decision tree classification: Ranking journals using IGIDI. *J. Inf. Sci.*, 2020; 46(3):325-339.

- [13] Marozzo F, Talia D, Trunfio P. A Workflow Management System for Scalable Data Mining on Clouds. *IEEE T. Serv. Comput.*, 2018; 11(3):480–492.
- [14] Zaborski D, Ali M, Eyduran E, Grzesiak W, Tirink C. Prediction of Selected Reproductive Traits of Indigenous Harnai Sheep under the Farm Management System via various Data Mining Algorithms. *Pak. J. Zool.*, 2019; 51(2):421–431.
- [15] Shakil R M, Mollah M A, Rahman S T, Habib M. A Bibliometric Review of Global Research on Human Resources Management and Supply Chain Management. *J Supply Chain Manag.* 2020; 9(4):173–184.
- [16] Gunasekaran, A. and Rahman, S.M. (2003) *Knowledge and Information Technology Management: Human and Social Perspectives*, IGI Publishing.
- [17] Nemati, H.R. and Barko, C.D. (2001) 'Issues in organizational data mining: a survey of current practices', *Journal of Data Warehousing*, Vol. 6, No. 1, pp.25–36.
- [18] Kamalesh, V. and Kuralmani, V. (2002) 'On applications of data mining to human resource data', *Proceedings of the International Conference on Scientific and Engineering Computation*, Singapore: Imperial College Press, pp.814–815.
- [19] Long, K.L. and Troutt, D.M. (2003) 'Data mining for human resource information systems', in J. Wang (Ed.) *Data Mining: Opportunities and Challenges*, Idea Group Inc., pp.367–379.
- [20] Hierji, K. (2001) 'Exploring data mining implementation', *Communications of the ACM*, pp.87–93.
- [21] Pyle, D. (1999) *Business Modeling and Data Mining*, Morgan Kaufmann, pp.35–66.
- [22] Kudyba, S. and Hoptroff, R. (2001) *Data Mining and Business Intelligence: A Guide to Productivity*, Idea Group Inc., pp.90–91.
- [23] Kovach, K.A. and Cathcart, C.E., Jr. (1999) 'Human Resource Information Systems (HRIS): providing business with rapid data access, information exchange and strategic advantage', *Public Personnel Management*, Vol. 28, No. 2, pp.275–282.
- [24] Janetzko, D. (2001) 'Knowledge creating communities', *Proceedings of the 3rd International Workshop on Advances in Learning Software Organizations*, Springer LNCS 2176, pp.19–20.
- [25] Takeda, M. (2005) 'How do managers conceptualise international human resource policy? Presenting and validating a typology of human resource management identity', *International Journal of Human Resources Development and Management*, Vol. 5, No. 1, pp.16–40.
- [26] Sheffield, J. (2005) 'Systemic knowledge and the V-model', *International Journal of Business Information Systems*, Vol. 1, Nos. 1–2, pp.83–101.
- [27] Gomes, F.C., Yasin, M.M. and Lisboa, V.J. (2007) 'An investigation of information availability and sharability for organizational performance measures', *International Journal of Business Information Systems*, Vol. 2, No. 1, pp.1–20.
- [28] Rajagopalan, B., York, M., Doane, P.D. and Tanniru, M. (2007) 'Enablers of enterprise systems training success – an exploratory investigation', *International Journal of Business Information Systems*, Vol. 2, No. 3, pp.250–265.

- [29] Walker, J.W. and Reif, W.E. (1999) 'Human resource leaders: capability, strengths and gaps', Human Resource Planning, Vol. 22, No. 4, pp.21-32.
- [30] Han, J. and Micheline, K. (2006) Data Mining Concepts and Techniques, Morgan Kaufmann, pp.4-27.
- [31] Barney, J. (1991) 'Firm resources and sustained competitive advantage', Journal of Management, Vol. 35, No. 1, pp.99-120.
- [32] Boxall, P. (1993) 'The strategic HRM debate and the resource-based view of the firm', Human Resource Management Journal, Vol. 6, No. 3, pp.59-75.
- [33] Feelders, A., Daniels, H. and Holsheimer, M. (2000) 'Methodological and practical aspects of data mining', Information and Management, pp.271-281.
- [34] Fayyad, U.M., Piatsky Shapiro, G. and Smyth, P. (1996a) 'From data mining to knowledge discovery in data base', AI Magazine, pp.37-54.
- [35] Berson, A., Smith, S. and Thearling, K. (1999) Building Data Mining Applications for CRM, McGraw-Hill Professional.
- [36] Cabena, P., Hadjinian, P., Stadler, R., Verhees, J. and Zanasi, A. (1998) Discovering Data Mining: From Concept to Implementation, Upper Saddle River, NJ: Prentice Hall PTR.
- [37] Westphal, C. and Blaxon, T. (1998) Data Mining Solutions: Methods and Tools for Solving Real World Problems, John Wiley Sons.
- [38] Inmon, W.H. (1996) Building the Data Warehouse, John Wiley & Sons.
- [39] Watson, H.J. and Haley, B.J. (1998) 'Managerial considerations', Communications of the ACM, pp.32-37.
- [40] Agostha, L. (2000) The Essential Guide to Data Warehousing, Upper Saddle River, NJ: Prentice Hall PTR.
- [41] Mallach, E.G. (2000) Decision Support and Data Warehouse System, Boston, MA: McGraw-Hill.
- [42] Kimball, R. and Ross, M. (2002) The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling, Wiley Computer Publishing, John Wiley & Sons.
- [43] Berson, A. and Smith, S.J. (2005) Data Warehousing, Data Mining & OLAP, Tata McGraw-Hill Edition, pp.5-6.
- [44] Berson, A., Smith, S. and Thearling, K. (1999) Building Data Mining Applications for CRM, McGraw-Hill Professional.
- [45] Hunter, I., Saunders, J., Boroughs, A. and Constance, S. (2006) HR Business Partners: Emerging Service Delivery Models for the HR Function, Gower Publishing Limited.
- [46] Ensher, E.A., Nielson, T. and Grant-Vallone, E.J. (2002) 'Tales from the hiring line: effects of the internet and technology on HR processes', Organizational Dynamics, pp.224-244.
- [47] Fisher, C.D., Schoenfeldt, L.F. and Shaw, J.B. (1999) Human Resource Management, Boston, MA: Houghton Mifflin Company.

- [48] Hendrickson, A.R. (2003) 'Human resource information systems: backbone technology of contemporary human resources', *Journal of Labor Research*, Vol. 24, No. 3, pp.381-394.
- [49] Bartram, T. (2005) 'Small firms, big ideas: the adoption of human resource management in Australian small firms', *Asia Pacific Journal of Human Resources*, Vol. 43, No. 1, pp.137-154.
- [50] Jena, K., Sanjib, B.S. and Patnaik, S. (2007) 'Development of HR module in ERP for IB thermal power station: a case study', *International Journal of Business Information Systems*, Vol. 2, No. 1, pp.99-125.
- [51] Sheehan, C., Holland, P. and De Cieri, H. (2006) 'Current developments in HRM in Australian organizations', *Asia Pacific Journal of Human Resources*, Vol. 44, No. 2, pp.132-152.
- [52] Fitz-Enz, J. (1998) 'Top 10 calculations for your HRIS', *HR Focus*, Vol. 75, No. 4, p.53.
- [53] Wessel, D.K. (2002) 'Decision making in the digital age', *DM Review, Resource Guide*, pp.16-18.
- [54] Merx-Chermin, M. and Nijhof, W.J. (2005) 'Factors influencing knowledge creation and innovation in an organization', *Journal of European Industrial Training*, Vol. 29, pp.135-147.
- [55] Celia, T., Romm, N.P. and Weber, Y. (1995) 'The relevance of organizational culture to the implementation of human resources information systems', *Asia Pacific Journal of Human Resources*, Vol. 33, pp.63-80.
- [56] Efraim, T. and Aronson, J. (1998) *Decision Support Systems and Intelligent Systems*, Upper Saddle River, NJ: Prentice Hall.
- [57] O'Brien, J.A. and Marakas, G.M. (2007) *Management Information Systems*, New Delhi: Tata McGraw-Hill Publishing Company Limited, pp.326-328.
- [58] Broderick, R. and Boudreau, J.W. (1992) 'Human resource management, information technology and the competitive edge', *Academy of Management Executive*, Vol. 6, No. 2, pp.7-17.
- [59] Fayyad, U.M., Piatsky Shapiro, G., Smyth, P. and Uthurusamy, R. (1996b) *Advances in Knowledge Discovery and Data Mining*, Menlo Park, CA: AAAI Press.
- [60] Duha, R.O., Hart, P.E. and Stork, D.G. (2003) *Pattern Classification*, John Wiley & Sons.
- [61] Berthold, M. and Hand, D.J. (1999) *Intelligent Data Analysis: An Introduction*, Springer, pp.3-10.
- [62] Hand, D., Mannila, H. and Smyth, P. (2001) *Principles of Data Mining*, The MIT Press, pp.9-10.
- [63] Peiter, A. and Dolf, Z. (2005) *Data Mining*, Pearson Education, pp.69-71.
- [64] James, M. (1985) *Classification Algorithms*, John Wiley & Sons.
- [65] Pearl, J. (1988) *Probabilistic Reasoning in Intelligent Systems*, Morgan Kauffman.
- [66] Smith, K. and Gupta, J. (2002) *Neural Networks in Business: Techniques and Applications*, IGI Publishing.
- [67] Bordley, C.E., Lane, T. and Stough, T.M. (1999) 'Knowledge discovery and data mining', *American Scientist*, pp.54-61.

- [68] Little, R. (1982) 'Models for non-response in sample surveys', *Journal of American Statistical Association*, pp.237–250.
- [69] Diggle, P. and Kenward, M. (1994) 'Informative dropout in longitudinal data analysis', *Applied Statistics*, Vol. 43, pp.49–94.
- [70] Hartley, H. and Hocking, R. (1971) 'The analysis of incomplete data', *Biometrics*, pp.783–808.
- [71] Raman, V. and Hellerstein, J.M. (2001) 'Potter's wheel: an interactive data cleaning system', *Proceedings of the Int. Conf. on Very Large Data Bases (VLDB'01)*, Rome, Italy, pp.381–390, <http://control.cs.berkeley.edu/abc> (retrieved 16 March 2007).
- [72] Cokins, G. (2004) *Performance Management: Finding the Missing Pieces (To Close the Intelligence Gap)*, John Wiley & Sons, Inc.
- [73] Dulebohn, J.H. (2003) 'Work redesign and technology implementation: the need for compensation system congruency', in D. Stone (Ed.) *Advances in Human Performance and Cognitive Engineering Research*, Greenwich, CT: JAI Press, Vol. 3.
- [74] Rynes, S. and Gerhart, B. (2001) 'Bringing compensation into I/O psychology (and vice versa)', in S. Rynes and B. Gerhart (Eds.) *Compensation in Organizations: Current Research and Practice*, San Francisco: New Lexington Press, pp.351–384.
- [75] Siddique, C.M. (2004) 'Job analysis: a strategic human resource management practice', *International Journal of Human Resource Management*, Vol. 15, No. 1, pp.219–244.
- [76] Gueutal, G.H. and Stone, L.D. (Eds.) (2005) *The Brave New World of eHR: Human Resources in the Digital Age*, Jossey-Bass Publishers.