



داده کاوی^۱: تعمیق نگاه برای کشف دانش در پیوند با دنیای الکترونیک و کاربرد آن در کتابداری و اطلاع رسانی

*اورانوس تاج الدینی - دانشجوی کارشناسی ارشد کتابداری و اطلاع رسانی دانشگاه پیام نور مشهد
Tajedini@yahoo.com و کارشناس کتابداری دانشگاه شهید باهنر کرمان .
*علی سادات موسوی - دانشجوی کارشناسی ارشد کتابداری و اطلاع رسانی دانشگاه پیام نور مشهد
و کارشناس کتابداری دانشگاه آزاد اسلامی تهران ، واحد علوم و تحقیقات . Moosavi56@gmail.com

چکیده :

فناوری داده کاوی یکی از روشهای سودمند در یافتن اطلاعات مفید در بین حجم عظیمی از اطلاعات است . ایده ای که مبنای داده کاوی است یک فرایند با اهمیت در شناخت الگوهای بالقوه مفید ، تازه و در نهایت قابل درک در داده هاست که ارزشهای پنهانی داده ها را آشکار می سازد و با کشف این ارزشها می توان ارزش دیگر متغیرها را پیش بینی کرد و در امر تصمیم گیری از آنها بهره جست .
داده کاوی یکی از پیشرفتهای اخیر در راستای فناوری مدیریت داده هاست .
امروزه با توجه به ایجاد سیستمهای یکپارچه اطلاعاتی و رشد فناوری اطلاعات و فناوریهای نوین اطلاعاتی و الکترونیکی اهمیت استخراج داده های با ارزش از میان انبوه داده ها بیش از پیش نمایان گشته است .
مقاله حاضر ضمن معرفی داده کاوی و مراحل و عملکردهای این فناوری به پیوند آن با فناوری اطلاعات اشاره داشته و تأثیر آن در دنیای الکترونیک را بیان ساخته و در نهایت کاربرد این فناوری در کتابداری و اطلاع رسانی که جزء لاینفک فناوری اطلاعات و علم به کارگیری این فناوری برای کندوکاو در سایر علوم است را مطرح می نماید .

کلید واژه ها : فناوری اطلاعات ، داده کاوی ، دنیای الکترونیک ، کتابداری و اطلاع رسانی ، مدیریت دانش

مقدمه :

در گذر زمان و پس از افت و خیزهای فراوان دنیای دانش از بدو تولد تاکنون، امروزه به مرحله ای گام نهاده ایم که

امروزه با توجه به ایجاد سیستمهای یکپارچه اطلاعاتی و رشد فناوری اطلاعات و فناوریهای نوین اطلاعاتی و الکترونیکی اهمیت استخراج داده های با ارزش از میان

است که هر کشور به منظور توسعه علمی، فنی، اقتصادی و... خود به آن نیاز دارد و این کشف و استخراج دانش، پیوندی ناگسستنی با فناوریهای نوین اطلاعات برقرار کرده است، به طوریکه همگام با پیشرفتهای الکترونیکی برای استخراج دانش پیشرفت فناوری داده کاوی و کنکاش برای کشف داده‌ها از دل پایگاههای دانش، لشکر پیشگام پیشرفتهای علمی محسوب می‌گردد.

در کشور ما نیز سازمانها، شرکتهای، موسسات دولتی و خصوصی و... به طور فزاینده ولی لنگ لنگان در حال ایجاد یا خرید نرم افزارهای پایگاه داده‌ها و مکانیزه سازی سیستمهای اطلاعات خود هستند.

با توجه به برنامه چهارم توسعه در خصوص دادوستدهای الکترونیکی و همچنین تأکید به برخورداری کشور از فناوریهای جدید اطلاعات برای دستیابی آسان به اطلاعات داخلی و خارجی، دولت مکلف شده است امکانات لازم برای دستیابی آسان به اطلاعات، زمینه سازی برای اتصال کشور به شبکه‌های جهانی و ایجاد زیرساختهای ارتباطی و شاهراههای اطلاعاتی را فراهم سازد (فصول دهم و یازدهم برنامه سوم توسعه کشور، ۱۳۷۹) و بدیهی است که این امر باعث ایجاد پایگاه عظیم داده‌ها شده و ضرورت استفاده از فناوری داده کاوی را بیش از گذشته نمایان ساخته است.

معرفی مفهوم داده کاوی:

به طور کلی داده‌ها در جهان واقعی آلوده هستند، این داده‌ها، بی کیفیت، ناقص، پرخطا و متناقض می‌باشند، بدیهی است که داده‌های بی کیفیت منجر به نتایج بی کیفیت می‌شوند.

داده کاوی با استفاده از رایانه به جمع آوری اطلاعات سودمند از بین حجم عظیمی از داده‌های ذخیره شده در یک پایگاه داده‌ها - که چه بسا تا حد زیادی بی کیفیت هستند - می‌پردازد تا ساختارهای باارزش، الگوهای مرتبط و اصول پنهانی داده‌ها را کشف نماید.

عبارت داده کاوی معادل یکی از عبارتهای استخراج دانش، برداشت اطلاعات، واریسی داده‌ها و حتی لایروبی کردن داده‌هاست که در حقیقت کشف دانش در پایگاه داده‌ها (KDD)^۲ را توصیف می‌کند. داده کاوی کاربرد در سطح بالای فنون و ابزار به کار برده شده برای معرفی و تحلیل داده‌های تصمیم گیرندگان است.

(کاربیس، ۲۰۰۳) داده کاوی استخراج غیر بدیهی

این شعار در همه دنیا پذیرفته شده است که اطلاعات طلاست. در دنیای کنونی، کشور و ملتی موفق تر است که قدرت علمی آن حرف اول را بزند و در کشف دانش نخستین گامها را بردارد. با وجود پدیده انفجار اطلاعات و رشد سریع دانش، نیاز به پیشرفت در روشهای کشف، سازماندهی و اشاعه دانش بیش از پیش احساس می‌شود.

از زمان ابداع رایانه برای تحلیل و ذخیره سازی داده‌ها؛ و بعد تولد اینترنت و سپس وب همه به دنبال استفاده از قابلیتهای اعجاب آور آن بودند و پس از رسوخ این دنیای عنکبوتی در سطوح مختلف زندگی انسان با وجود مزایا و معایب گوناگون این موضوع حتی کشف و شهود علمی نیز تحت تاثیر فناوری رایانه ای قرار گرفت و واژه داده به عنوان مقدمه هر قانون علمی متولد شد، به طوریکه هر کشفی برای معتبر شدن و تبدیل به قانون علمی گشتن ناگزیر به طی این فرایند است:



و خرد جمعی همان قوانین علمی و نظریه‌های پذیرفته شده در سراسر جهان است. (ابرامی، ۱۳۷۹)

با پیشرفت فناوری اطلاعات و نفوذ آن در همه رشته‌های علمی و جای جای زندگی بشر نیاز به تحلیل و استخراج اطلاعات و کشف دانش در تعامل با رایانه و فناوری اطلاعات بیش از پیش آشکار گشت و پایگاههای اطلاعاتی به عنوان منابع اصلی و معتبر داده‌های مختلف اطلاعاتی گام به عرصه وجود نهادند و این پایگاهها منشاء ظهور پیشرفتهایی چون سیستمهای یکپارچه اطلاعاتی، سیستمهای یکپارچه بانکی، تجارت الکترونیک و... گشتند و به وجود آمدن انبارهای عظیمی از داده‌ها، ضرورت کشف و استخراج سریع و دقیق دانش از پایگاههای داده‌ها را نمایان ساخت.

هم اکنون در هر کشور، سازمان، شرکت و... برای امور بازرگانی، پزشکی، آموزشی، آماری و... آن ارگان، پایگاه داده‌ها ایجاد یا خریداری شده است به طوریکه تصمیم گیریهای راهبردی، تهیه گزارشهای مختلف، توصیف وضعیت جاری و... با استفاده از پایگاه داده‌ها انجام می‌گیرد. (سعیدی، ۱۳۸۴)

داده کاوی یا استخراج و کشف سریع و دقیق اطلاعات باارزش و پنهان از پایگاه داده‌ها از جمله اموری

دانش از پایگاه داده ها باید با طی مراحل منظمی صورت گیرد که اگر با دیده اغماض به آنها بنگریم شامل مراحل زیر می باشد : (سعیدی ، ۱۳۸۴)

۱. درک قلمرو .
۲. آماده کردن مجموعه داده ها .
۳. کشف الگوها .
۴. پردازش بعد از کشف الگوها .
۵. استفاده از نتایج .

در مرحله درک قلمرو، هدف از استخراج داده ها مشخص می شود و این که چه چیزی و در چه سطحی مدنظر است . داده های خام از منابع مختلفی جمع آوری می شوند و از طریق استخراج ، ترجمه و فرایندهای بازخوانی به انبار داده ها وارد می شوند و اصولاً داده هایی انتخاب می شوند که با اهداف ما منطبق و هماهنگ باشند .

در بخش آماده کردن مجموعه داده ها و مهیاسازی ، داده ها از انبار خارج شده و به صورت یک فرمت مناسب برای داده کاوی در می آیند . هدف از این مرحله تولید یک مجموعه آماده از داده های کاوش شده است ؛ شناسایی متغیرهای زاید در یک مجموعه داده ها و رفع این متغیرها نیز از اهداف این مرحله است .

در بخش کشف الگوها، با روشهای داده کاوی برای پاسخ به سئوالات خاصی که به ذهن می رسند الگوریتمهایی را استخراج می کنند و از این الگوریتمها برای ساخت الگوها استفاده می شود .

در پردازش الگوها باید درمورد نتایج به دست آمده بحث شود و مورد تجزیه و تحلیل قرار بگیرند تا مطمئن شویم که یافته ها صحیح و منطبق با اهداف ما هستند و همچنین بررسی شود که نتایج قابل دسترسی می باشند .

به عبارت بهتر در تجزیه و تحلیل، الگوهای مورد نظر به دانش مفید و قابل استفاده تبدیل می شوند و پس از بهبود الگوهایی که کاراً محسوب می شوند، در یک سیستم اجرایی به کار گرفته خواهد شد .

نتیجه و گزارش نهایی ارائه شده برای کاربر تهیه می شود و باید به استناد به کل فرایند داده کاوی باشد ، بدیهی است که هدف نهایی داده کاوی استفاده از نتایج کشف شده برای ایجاد یک موقعیت جدید و بهتر است .

اگر مراحل داده کاوی را به صورت نمودار نشان دهیم این نمودار حاصل می گردد :

اطلاعات بالقوه مفید از روی داده هایی است که قبلاً ناشناخته مانده اند . این مطلب برخی از روشهای فنی مانند خوشه بندی داده ها ، خلاصه سازی داده ها ، فراگیری قاعده های رده بندی ، یافتن ارتباط شبکه ها ، تحلیل تغییرات و کشف بی قاعدگی را شامل می شود . (سپانتسکی ، ۱۹۹۱)

امروزه داده کاوی توسط آمارشناسان ، تحلیل گران داده ها ، مدیریت سیستمهای اطلاعات، پژوهشگران یادگیری ماشین و هوش مصنوعی به کار برده می شود که با توجه به رشد فناوری اطلاعات این استفاده روزافزون گشته است .

داده کاوی و استخراج دانش از پایگاه داده ها از موضوعاتی است که همزمان با تولد پایگاههای اطلاعاتی ایجاد شد . پژوهشگران و متخصصان رایانه ، هوش مصنوعی ، یادگیری ماشینی و... از اوایل دهه ۸۰ به پژوهش در این زمینه پرداختند اما شاید اولین گزارش رسمی در مورد داده کاوی را بتوان به لوول در سال ۱۹۸۳ نسبت داد .

صرف نظر از گامهای نخستین در این زمینه، پژوهشهای اساسی و جدی در مورد داده کاوی از اوایل دهه ۹۰ آغاز شد و سمینارها، دوره های آموزشی ، کارگاهها و کنفرانسهای متعددی در این زمینه برگزار شد و نتایج آنها به صورت مکتوب منتشر گشت .

مراحل داده کاوی :

کشف دانش در جهت کشف اطلاعات مفید از مجموعه بزرگ داده ها می باشد . دانش کشف شده می تواند قاعده ای باشد تا ویژگیهای داده ها، الگوهای متناسب و خوشه بندی موضوعهایی که درون پایگاه رخ می دهد و... را توصیف نماید .

اساس فرایند داده کاوی بر مبنای سه فعالیت اصلی است که عبارتند از :

۱. حذف داده ها : در این عمل داده های بی ارزش و عوامل بیرونی حذف می شوند .
۲. فشردن داده ها : این عمل به وسیله کدگذاری داده ها صورت می گیرد .
۳. کشف الگوها : در این عمل الگوهای موجود در پایگاه داده ها از قبیل طبقه بندی ، الگوهای زنجیری و... کشف می شوند .

سیستم KDD باید ابزارهایی با اثر تعاملی داشته باشد نه سیستمهای تجزیه و تحلیل خودکار؛ لذا کشف



اطلاعاتی به امری حیاتی برای بهره گیری از پیشرفتهای انجام گرفته در علم تبدیل گشته است. هدف اصلی داده کاوی پیش بینی است و برای ثمربخشی دقیق و اصولی این فرایند نیاز به تسلط بر پیشرفتهای صورت گرفته در فناوری اطلاعات بیش از پیش ضروری است.

داده کاوی همراه با پیشرفتهای دنیای الکترونیک :

داده کاوی اصولاً برای کاوش داده ها صورت می گیرد و یافته ها با به کارگیری الگوهایی احراز اعتبار می شوند.

امروزه با پیشرفتهای فناوری به کارگیری فناوری اطلاعات به جزء لاینفکی از زندگی افراد تبدیل گشته است و در یک دگردیسی پیچیده دنیای نامتناهی مجازی پیوندی ناگسستنی با زندگی افراد جوامع برقرار نموده است.

برای به دست آوردن اطلاعات مفید با توجه به فناوری اطلاعات، نیاز به عملکردی اصولی برای استخراج دانش مورد نیاز از دل انبارهای عظیم داده ها احساس می شود تا اطلاعات به دست آمده برای ایجاد یک موقعیت جدید و بهتر کاربرد داشته باشد.

داده کاوی برای استخراج دانش از پایگاه داده چه بزرگ و چه کوچک کاربرد دارد و از نتایج و الگوهای تولیدشده به وسیله آن در تصمیم گیری استراتژیک در این دنیای الکترونیک می توان بهره جست، به عبارتی داده کاوی اطلاعاتی می دهد که برای گرفتن تصمیم های هوشمندانه به آنها نیاز داریم.

نفوذ الکترونیک در عرصه های مختلف زندگی اتفاق افتاده است، از آن جمله شرکتهای تجاری، بانکداری، بیمه، آموزش عالی، کتابداری و اطلاع رسانی، کتابخانه های دیجیتالی و... بخشهایی هستند که به طور اساسی با پیشرفتهای فناوری و فناوریهای اطلاعاتی و خدمات الکترونیکی دست به گریبانند.

داده کاوی در حوزه بازاریابی به دلیل پیوستگی غیر قابل انتظاری که بین پروفایل یک مشتری و الگوی خرید ایجاد می کند، اهمیت خاصی دارد. (باری و لینف، ۱۹۹۷) در یک شرکت تجاری یا فروشگاههای زنجیره ای و یا شرکتهای بیمه پس از داده کاوی است

البته شایان ذکر است که در کنار همه این مراحل، فرایندهای جنبی از قبیل انتخاب، ابزار، قالب بندی، پاسخگویی، مدل سازی، اعتبارسازی و... نیز باید مورد توجه قرار گیرد.

پیوند داده کاوی با مدیریت دانش و فناوری اطلاعات :

داده کاوی یا کشف دانش در پایگاه داده ها (KDD) علم نسبتاً تازه ای است با توجه به پیشرفت کشور در زمینه IT و نگاههای ویژه به دولت الکترونیک و نفوذ استفاده از سیستمهای رایانه ای در صنعت و ایجاد بانکهای اطلاعاتی بزرگ توسط ادارات دولتی، بانکها و بخش خصوصی نیاز به استفاده از آن به طور عمیقی احساس می شود.

دانش به طور انحصاری محصول فناوری اطلاعات نیست ولی این فناوری به طور کلی در ایجاد دانش و فرایند مدیریت دانش از ابتدا مشارکت داشته است. امروزه مدیریت دانش از زیر مجموعه های اصلی فرایند فناوری اطلاعات به شمار می رود، زیرا در جمع آوری دانش، تبدیل دانش، انتقال داده ها و اطلاعات دانش نقش کلیدی دارد. (هند، ۱۹۹۸) هنگامی که از دیدگاه دانش و مدیریت آن به داده کاوی می نگریم هدف آن کشف دانش سازمانی ولی نهفته در اطلاعات خام است، البته به این صورت نیست که هر بینش حاصل از داده کاوی، دانش را بسازد بلکه بسیاری از نتایج به دست آمده، اطلاعات برای مدیریت یا هوش سازمانی است. داده کاوی در مدیریت دانش، کشف را با تحلیل تلفیق می کند، تلفیقی که اغلب منجر به ایجاد دانش می گردد.

داده کاوی اصولاً در تعامل با پایگاههای اطلاعاتی که زاینده پیشرفت در فناوری و فناوری اطلاعات هستند معنا می یابد، به بیان بهتر داده کاوی تجزیه و تحلیل ماشینی داده ها برای پیدا کردن الگوهای مفید و قابل استناد در پایگاههای اطلاعاتی می باشد. (زعفریان، ۱۳۸۰)

امروزه با توجه به پیشرفتهای چشم گیر در فناوری اطلاعات، استخراج اطلاعات از درون پایگاههای

غلبه بر رشد مستمر و تنوع بانکهای اطلاعاتی طلب می کند .

به کار گیری فناوری اطلاعات توسط هر سازمان در عمل یک فرایند هموار نیست و کتابخانه ها نیز از این قضیه مستثنی نیستند ؛ اما تجربه نشان داده است که یک برنامه نظام مند می تواند ظهور و نگهداری فناوری داده کاوی در محیط کتابخانه را تسهیل نماید . داده کاوی می تواند در بخشهای مختلفی از کتابداری از قبیل بانکهای اطلاعاتی و اطلاع رسانی ، گردآوری منابع ، استفاده از مجموعه ، امانت بین کتابخانه ای ، داده های بخش امانت ، داده های هزینه و... کاربرد داشته باشد . (پاتکار ، ۲۰۰۴)

ظهور کتابخانه های دیجیتالی در عصر حاضر مشکلات و مسائل عدیده ای را برای مدیران کتابخانه ها پدید آورده است .

مشکلاتی از قبیل شناخت مراجعه کنندگان و نیازهای آنها ، بهبود خدمت رسانی ، نحوه دسترسی کاربران به این کتابخانه ها ، نحوه دسترسی کاربران به صفحات وب و... رخ نمایانده است .

یکی از روشهای حل این مسائل داده کاوی است که مدیران با استفاده از الگوهای کشف شده توسط داده کاوی در امر تصمیم گیری موفق تر عمل کنند . با استفاده از گردآوری و تحلیل داده های مربوط به آمار شمارش مراجعین ، میزان امانت ، امانت بین کتابخانه ای ، گسترش مجموعه ، تهیه مواد ، کاربرد منابع الکترونیکی و روندهای استفاده از وب و... از داده کاوی برای تحکیم تصمیم گیریهای اداری کتابخانه استفاده می شود . (مهریزی ، ۱۳۸۲)

در بخش امانت کتابخانه از داده کاوی به منظور افزایش میزان امانت گرفتن منابع و یا به عبارتی جذب مراجعه کنندگان بیشتر برای امانت گرفتن منابع استفاده می شود ، در حقیقت داده کاوی نقش پشتیبان در پیدا کردن اطلاعات مورد نیاز مراجعه کنندگان را ایفا می کند . ارتباط بین داده کاوی و مدیریت کتابخانه را می توان به صورت نمودار ساده زیر در نظر گرفت :



که مشخص می شود میزان استفاده از کالاها و خدمات جدید چه مقدار بوده است ، میزان استقبال مشتریان و آینده نگری برای وضعیت آینده شرکت یا مؤسسه نیز معلوم می گردد و با استفاده از نتایج داده کاوی است که اغلب تجار آنها با تصمیم گیریهای استراتژیک و اتخاذ خط مشیهای جدید برای خدمات رسانی بهتر مواجه می گردند . (شاه سمندی ، ۱۳۸۴)

در بانکداری الکترونیک نیز با استفاده از نتایج داده کاوی نمونه های زیر قابل پیش بینی می گردد :

• کشف پول شویی که از جالب توجه ترین کاربردهای داده کاوی است .

• تشخیص مشتریان ثابت و همیشگی .

• تعیین مشتریان استفاده کننده از یک سرویس خاص .

آموزش عالی نیز همواره با داده ها و اطلاعات بسیار زیادی در مورد دانشگاهها ، دانشجویان ، اعضاء هیات علمی ، پرسنل ، منابع مادی و... روبروست و در اکثر مواقع این داده ها می تواند حامل اطلاعات و الگوهای ارزشمند باشند . پیدا کردن الگوها و دانش نهفته در این اطلاعات می توانند به تصمیم گیرندگان آموزش عالی کمک شایانی نماید . (سعیدی ، ۱۳۸۴)

داده کاوی یک رشته جدید با کاربردهای وسیع و گوناگون است که با توجه به فناوریهای جدید اطلاعاتی چشم نواز جلوه می کند و علاوه بر مسائل ذکر شده می تواند در امور پزشکی ، زیست پزشکی ، تجزیه و تحلیل های مربوط به DNA ، کشف ناهنجاریهای جوامع ، اسناد جعلی ، ارتباطات از راه دور ، ورزش ، سرگرمی و... بسیار راهگشای باشد . (رویجر ، ۲۰۰۳)

کاربرد داده کاوی در کتابداری و اطلاع رسانی :

با توجه به پیوند عمیق فناوری اطلاعات و پیشرفتهای فناوری اطلاعات با علم کتابداری و اطلاع رسانی و نیز تأثیر غیر قابل انکار علم میان رشته ای کتابداری و اطلاع رسانی با جنبه های متفاوت پیشرفت علم و فناوری دانش و نیز استفاده شایان توجه متخصصان علوم مختلف از این علم به عنوان راهکار و راهگشای پیشرفتهای علمی و سرلوحه ای برای استفاده از فناوری اطلاعات ، به کاربرد داده کاوی در کتابداری و اطلاع رسانی به طور جداگانه ای اشاره می شود .

پیشرفتهای حاصله در علم کتابداری و اطلاع رسانی و فناوری اطلاعات ، فنون و ابزارهای جدیدی برای



• کسب اطلاع از پایگاههای مختلف داده .
 انتظار می رود داده کاوی در گسترش سازمان کتابخانه مشارکت داشته باشد .
 کشف انتخابهای نوین با بهره گیری از داده کاوی اطمینان بخش بهترین کاربرد ممکن منابع موجود است .
 داده کاوی ماهیت چرخه مانند دارد ، برای این که در پی کشف الگوها سئوالات جدیدی پدید خواهد آمد که دور بعد فرایند را شکل می دهد .
 بهره برداری از فناوری داده کاوی مطمئناً برای متخصصان کتابداری یک چالش دائمی خواهد بود چرا که آنها خلاقیت طلبند و برای نوآوری تلاش می کنند .

نتیجه گیری :

فناوری داده کاوی یکی از روشهای سودمند در

داده کاوی در بخش مرجع کتابخانه و خدمات اطلاع رسانی نیز کاربرد دارد با استفاده از داده کاوی می توان روندهای استفاده از وب و خدمات الکترونیکی را ردیابی کرد و در نتیجه سفارشگاههای بزرگ منابع اطلاعاتی ایجاد نمود ؛ ایجاد چنین پایگاههایی باعث تسهیل انجام تحقیق برای کاربران کتابخانه می شود .
 ارتباط داده کاوی که مقوله ای با کاربرد بسیار گسترده در حوزه های مختلف می باشد - با علم کتابداری و اطلاع رسانی وسیع تر از آنچه ذکر شد ، است اما اگر به یک حرف از هزاران بسنده کنیم ؛ این موضوع می تواند به صورت زیر بیان شود :

- کشف ارتباط بین آثار مختلف .
- سنجش رفتار اطلاع یابی کاربران .
- کمک به مدیران در تصمیم گیری .
- جذب مراجعه کنندگان به پایگاه .

یابد منجر به استفاده از داده های موجود در تصمیم گیریهای استراتژیک می شود. داده کاوی در پایگاه داده ها با توجه به پیشرفت کشور در زمینه فناوری اطلاعات و نگاه ویژه دولت الکترونیک به استفاده از فناوری و نفوذ استفاده از سیستمهای رایانه ای در صنعت و ایجاد بانکهای اطلاعاتی بزرگ توسط ادارات دولتی، بانکها و بخش خصوصی نیاز به استفاده از این فناوری به طور عمیقی احساس می شود و در یک کلام داده کاوی در پیوند و تعامل با فناوری اطلاعات و فناوری دانش عبارت است از یافتن طلای اطلاعات از سلسله کوههای عظیم داده ها.

یافتن اطلاعات مفید از میان حجم عظیمی از اطلاعات است. داده کاوی فرایند کشف رابطه ها، الگوها و روشهای جدید معنی داری است که به بررسی حجم وسیعی از اطلاعات ذخیره شده در انبارهای داده با فناوریهای تشخیص الگو می پردازد. شرکتها، سازمانها، مؤسسات و... غرق در انبوه داده ها و اطلاعاتی هستند که استفاده از آنها در بیشتر موارد، محدود به انجام کارهای جاری می باشند و هنوز از این داده ها در تصمیم گیریهای استراتژیک استفاده نمی شود. داده کاوی که استفاده از آن روز به روز توسعه می

پی نوشت:

۱- Data Mining

۲- Knowledge Discovery Database

منابع و ماخذ

۱. ابرامی، هوشنگ (۱۳۷۹). شناختی بر دانش شناسی. تهران: کتابدار.
۲. سعیدی، احمد (۱۳۸۴). «داده کاوی، مفهوم و کاربرد آن در آموزش عالی». نامه آموزش عالی، شماره ۱۸.
۳. هند، دیوید (۱۹۹۸). «داده کاوی چیست؟». مترجم حمید معظمی گودرزی، گزیده مطالب آماری، شماره ۵۲.
۴. شاه سمندی، پرستو. «داده کاوی در ارتباط با مشتری». تدبیر، شماره ۱۵۶.
۵. زعفریان، رضا؛ زعفریان، قاسم (۱۳۸۰). «مروری بر داده کاوی». صنایع، شماره ۱۹.
۶. مهریزی حائری، علی اصغر (۱۳۸۲). «داده کاوی: مفاهیم، روشها و کاربردها». پایان نامه کارشناسی ارشد رشته آمار اقتصادی و اجتماعی دانشگاه علامه طباطبایی.
- 7- Rygielski, Ch (2002). «Data mining techniques for customer relationship management, technology in society».
- 8- Karypis, G (2003). «Data mining». USA: university of Minnesota
- 9- Roiger, R (2003). «Data mining: A tutorial based primer». Boston: Addison Wesley.
- 10- Adriaanse, P (2003). «Data mining: 9th India reprint». Dehli: person Education.
- 11- Barry, M; Linoff, G (1997). «Data mining techniques for marketing», sals and customer report. New York: John Wiely & sons.
- 12- Patkar, V (2004). «Determining priorities for computerization in library». Annual of library science and documentation. vol. 53(3)