

استفاده از تکنیک داده کاوی برای تجزیه و تحلیل داده ها و اطلاعات

در سیستم های کیفیت

حسین نوپچی



فوق لیسانس مهندسی صنایع

کارشناس برنامه ریزی شرکت کلر پارس تبریز

رئیس هیئت مدیره شرکت سرآمدان کیفیت آذربایجان

## چکیده

از دیر زمان مقوله داده و اطلاعات مد نظر سازمان ها و مدیران بوده است و تلاش متخصصان بر این بوده است که دانش نهفته در این داده ها را کشف و استخراج نمایند. امروزه سازمان هایی موفق هستند که دسترسی سریع و صحیح به اطلاعات داشته باشند. اخیراً فناوری اطلاعات و ارتباطات (ICT) با قابلیت های ویژه خود به کمک سازمان ها آمده است تا بر روی داده های در دسترس و موجود پایگاه داده ها تمرکز کنند و به اطلاعات و دانش مناسب برای تصمیم گیری و پشتیبانی تصمیمات دست یابند. مدیریت صحیح دانش کسب شده از تجزیه و تحلیل داده های درون و برون سازمانی و کسب اطلاعات و کشف دانش از این داده ها و تبدیل آن به دانش سازمانی، عاملی برای کسب مزیت های رقابتی و بهبود فرآیندهای سازمانی و در نهایت بالا بردن کیفیت محصولات/خدمات برای برتری بر رقبای خواهد بود. امروزه فناوری اطلاعات و ابزارهای آن نظیر تکنیک داده کاوی که نگرشی تازه به مسئله کشف (استخراج) دانش از پایگاه داده ها دارند مطرح گردیده است. در این مقاله با یک رویکرد نوین نسبت به اهمیت داده، اطلاعات و دانش در سازمان پرداخته شده است و نقش داده کاوی برای کمک به بهبود مستمر کیفیت فرآیندهای سازمان ها جهت فعالیت در یک محیط رقابتی بمنظور بالا بردن کیفیت ارائه محصولات/خدمات به مشتریان و ذینفعان پرداخته شده است.

**کلید واژه ها:** کیفیت، فناوری اطلاعات، داده کاوی (Data Mining).

### ۱- مقدمه

امروزه دانش به منزله یک منبع ارزشمند و استراتژیک مطرح بوده و دانش بمنزله دارایی سازمانی در نظر گرفته می شود. ارائه محصولات و خدمات با کیفیت مناسب و اقتصادی، بدون مدیریت و استفاده از این منبع ارزشمند، امری سخت و بعضاً غیر ممکن است [۱۳]. بسیاری از سازمان ها دریافته اند برای کسب موفقیت در دنیای پرچالش رقابتی، به نگرشی یکپارچه و سیستماتیک به اطلاعات و دانش درون سازمانی خود نیاز دارند. به همین دلیل طی سال های اخیر، تحقیقات خود را حول محور هایی از قبیل مدیریت بر جریان کسب، حفظ و استفاده از دانش متمرکز کرده اند [۷]. بشر با پیشرفت فناوری رایانه ای در ثبت و ذخیره سازی داده ها و پردازش آنها گامی بزرگ جهت کسب دانش برداشته است. در واقع داده نمایی از واقعیت ها، معلومات، مفاهیم، رویداده ها یا پدیده ها برای برقراری ارتباط، تفسیر یا پردازش، توسط انسان یا ماشین است [۱۳]. بی دلیل نیست که یکی از مدیران ارشد شرکت هیولت پاکارد می گوید: "اگر هیولت پاکارد می دانست که چه چیز هایی را می داند، سوددهی اش سه برابر می شد" [۱۴].

## ۲- کیفیت

تعریف کیفیت در سال های مختلف به تبع فرهنگ صنعتی و خدماتی، رشد تکنولوژی و توان علمی تغییرات مختلفی داشته است. وجه مشترک این تعاریف عبارت است از "سازگاری کالا و خدمات با نیازها و انتظارات مشتریان" [۹]. کیفیت به منزله صرفه جویی در وقت و پول است و از عمده ترین عوامل رقابتی محسوب می شود. امروزه تمرکز روی کیفیت از محصول به فرآیند تعمیم یافته است. گفته می شود که تولید ناخالص ملی آمریکا تنها در صورت اتخاذ یک سیاست ملی کیفیت می تواند حدود یک درصد افزایش یابد [۱۱]. کیفیت، به مواردی فراتر از کارایی استفاده از مواد مربوط می شود. کیفیت نامطلوب، نه فقط منجر به دور ریختن و رد شدن محصول می شود، بلکه نیروی زیادی را نیز صرف دوباره کاریها و تولید مجدد می کند. بدین ترتیب منابعی که به این ترتیب هدر می روند بسیار بیشتر از منابعی هستند که صرف بازرسی های ضروری خواهند شد. به علاوه، زیانی که از نارضایتی مشتری ایجاد می شود، ضرر فاحشی به شمار می آید و نیرویی نیز که صرف بازاریابی و فروش می شود، به هدر خواهد رفت [۱۰]. سیر تحول مدیریت کیفیت گویای آن است که فلسفه کیفیت جامع اولین بار در جنگ جهانی دوم توسط متخصصین و مهندسين کیفیت آمریکا طرح گردید. در واقع روش های آماری دکتر شوهارت توسط دمنینگ با توجه به عوامل فرهنگی و انسانی سازمان مورد توجه جدی قرار گرفت [۹]. سیستم های مدیریت کیفیت زیادی مطرح شده اند که مدیریت کیفیت جامع<sup>۱</sup> بعنوان یک سیستم کیفیتی مشتری گرا می باشد که دنبال بهبود مداوم فرآیند و مدیریت سازمان از طریق تکنیک های کنترل آماری، طراحی رویه ها، اصلاح سیاست ها و مدیریت منابع می باشد [۱۰].

از عمده ترین مشکلات اجرای مدیریت کیفیت جامع نیاز آن به حجم عظیم داده ها است که استفاده از فناوری اطلاعات را ضروری می سازد و در صورت برنامه ریزی و اجرای صحیح می تواند کمک بزرگی به مدیریت کیفیت سازمان باشد. زیرا فناوری اطلاعات ورود و خروج و پردازش داده ها را انجام می دهد و دانش و الگوهای ناشناخته را از درون داده کشف و استخراج می کند [۱۰].

### ۲-۱- کیفیت و تجزیه تحلیل اطلاعات

از اواخر دهه ۸۰ بشر به فکر دستیابی به اطلاعات و الگوهای نهفته در پایگاه داده های حاصل از نتایج عملکرد سازمان و داده های انبارش شده مشتریان افتاد. مدیران سازمان ها می دانند ماشین آلات، تجهیزات و ساختمان ها را نمی توان اصلی ترین دارایی یک سازمان به حساب آورد. آنچه بعنوان دارایی مهم هر سازمان به شمار می رود، اطلاعات و دانش سازمانی است [۱۲].

<sup>۱</sup> : Total Quality Management(TQM)

تجزیه و تحلیل داده ها و اطلاعات همواره در فرآیندهای سیستم های کیفیت از قبیل سری استانداردهای ISO 9001, ISO 14001 و مدیریت کیفیت جامع و از جمله در مدل تعالی سازمانی جایگاه خاصی را به خود اختصاص داده است. بطوریکه اطلاعات و تجزیه و تحلیل آنها در اکثر جوایز کیفیت امتیازهای بالایی را شامل می شوند. در جدول شماره ۱- امتیازهایی که به معیار ارزیابی و تجزیه و تحلیل اطلاعات، داده شده اند ارائه گردیده است [۸].

معیارهای ارزیابی	MBNQA بالدرج	CQBEA کانادا	EQA اروپا	DP دینور	ABEA استرالیا	PMQA مالزی	IAPQA پاسفیک	JQA ژاپن	PQA فیلیپین	HKQA هنگ کنگ
اطلاعات و تجزیه و تحلیل آنها	۸۵	۰	۱۸	۱۵۰	۱۰۰	۱۳۳	۸۵	۸۰	۸۵	۸۵

جدول شماره ۱- امتیازهای اختصاصی به معیار ارزیابی تجزیه و تحلیل اطلاعات

### ۳- نقش فناوری اطلاعات (IT) در سازمان ها

ساختار سازمانی، یک مسئله درخور توجه است و ساختار غلط به طور جدی به عملکرد تجاری آسیب خواهد زد، حتی ممکن است آن را تخریب کند. سازمان های آینده مبتنی بر اطلاعات خواهند بود و عمدتاً از متخصصان عملیاتی تشکیل خواهند شد و صلاحاً مدیر میانی نخواهند داشت. تغییر در ساختار سازمانی با در نظر گرفتن تعداد سطوح سلسله مراتبی و موقعیت های مدیریت میانی، رسمی سازی و تمرکز اقتدار قابل اندازه گیری است [۱۴].

یکی از مسائل اساسی در سازمانها این است که سازمان به دنبال تغییرات بنیادین در فرآیندهایش است تا بتواند از مهلکه رقابت شدید جان سالم بدر برد و نکته حائز اهمیت در اینجا توجه به نقش مهم و حیاتی سیستم های اطلاعاتی می باشد. امروزه کاربرد IT روز به روز در حال افزایش است. پیشرفت های حاصل شده در فن آوری های ارتباطات و کامپیوتر به کارکنان یک سازمان این اجازه را می دهد که در حالی که بیرون سازمان هستند، همچنان به سازمانشان متصل باشند و برای سازمان مربوطه کار کنند و یا فاصله خود را با مشتریان به حداقل برسانند، به عبارت دیگر سازمانها با استفاده از این فن آوری ها به سمت مجازی شدن پیش می روند. نمونه هایی از این فن آوری ها عبارتند از: ارتباطات ویدئویی، پست الکترونیکی، طراحی به کمک کامپیوتر، تولید به کمک کامپیوتر و مهندسی به کمک کامپیوتر، این قبیل فن آوری ها باعث هماهنگ سازی فعالیت ها در سازمان می گردد [۱۴]. بطور کلی از دید اسمیت و استروف هر شرکت دارای سه شالوده اصلی فرآیند تصمیم گیری، جریان اطلاعات و جریان مواد است.

فناوری اطلاعات روی هر کدام از این سه شالوده می تواند تاثیر گذار باشد. به طور کلی کاربرد فناوری اطلاعات به شرکت ها کمک می کند تا از مزایای زیر بهره مند گردند:

- ۱- صرفه جویی در هزینه و بهبود دقت اطلاعاتی.
  - ۲- اجتناب از اشتباهات انسانی گاهی که وظایف تکراری یا بسیار پیچیده است.
  - ۳- صرفه جویی مالی به علت کاهش خطا در زمانی که صرف انجام وظایف می گردد.
  - ۴- ادغام و یکپارچه سازی چند وظیفه در یک وظیفه.
  - ۵- بهبود بخشیدن کارایی سازمان.
  - ۶- بهبود در مدیریت میانی و کاهش فرایندهای زائد از طریق تهیه اطلاعات مفیدتر
- محیط های امروز ایجاب می کند که شرکت ها به سرعت توسعه یابند و تولیداتی را ارائه نمایند که نیازهای خریداران را برآورده نماید. امروزه شرکت های با درجه پایین مشارکت فرایندها قادر به جوابگویی به نیازهای متنوع مشتریان نمی باشند. نقشی که فناوری اطلاعات در این محیط ها ایفا می کند تسهیل تغییرات افزایش درجه مشارکت و کاهش درجه کارهای واسطه ای می باشد. خصوصیات فناوری اطلاعات برای انجام این وظایف دارد به شرح زیر است [۱۴]:
- ۱- ایجاد پایگاه داده های اشتراکی: به این وسیله کارهای مختلفی به صورت مستقیم و با استفاده از اطلاعات موجود در پایگاه داده توسط اکثر افراد سازمان دیده می شود.
  - ۲- فناوری های تصویری: به این وسیله افراد همزمان و در حالیکه در بیرون از سازمان هستند با همدیگر ارتباط تصویری دارند. و روی گراف ها و نمودارها تجزیه و تحلیل انجام می دهند.
  - ۳- تحلیل داده های الکترونیکی و انتقال آن به افراد دیگر: فناوری اطلاعات به همه افراد متخصص اجازه می دهد که در ارزیابی اطلاعات دخیل باشند.

### ۳-۱- تأثیر فناوری اطلاعات روی فرآیند های تجاری

به موازات ورود شرکت ها به دهه ۱۹۹۰، دانش به یکی از مهمترین منابع استراتژیک تبدیل گردید. تولید دانش برای پایداری مزیت رقابتی و موفقیت سازمانی بسیار اساسی است. بدین منظور فناوری اطلاعات به کمک سازمان ها آمده است تا آنها شکل های نوآورانه ای از مزیت های رقابتی را برای خود ایجاد نمایند. فناوری اطلاعات می تواند سه اثر جداگانه روی فرآیند های تجاری داشته باشد [۱۴].

۱- تأثیرات اتوماسیون: با اینکار فناوری اطلاعات جانشین نیروی کار شده و تأثیراتی از قبیل بهبود

بهره وری و کاهش هزینه را دربر دارد.



۲- تأثیرات اطلاعاتی: که از توانایی فناوری اطلاعات در جمع آوری، ذخیره، بازیابی اطلاعات ناشی می شود، که بهبود کیفیت تصمیم گیری، افزایش قدرت کارکنان، کاهش استفاده از منابع، افزایش تأثیرات سازمانی از نتایج آن می باشد.

۳- تأثیرات تغییر: که ناشی از محرک های ارزشی در توانایی فناوری اطلاعات در نوآوری فرآیند است.

### ۲-۳- تأثیر فناوری اطلاعات بر کیفیت

تکنولوژی اطلاعات (IT)، ابزار قدرتمندی است که به عنوان اهرمی برای طراحی دوباره فرآیندهای کسب و کار در سازمان ها به کار می رود و این امکان را به سازمانها می دهد که به جای بهبود مقطعی و تدریجی وضعیت موجود، سیستم های جدید و یکپارچه ایجاد نمایند. مطالعات مختلف نشان می دهد که فناوری اطلاعات باعث تغییرات بنیادین در مدیریت و هماهنگی تولید می گردد فناوری اطلاعات توان یکپارچه سازی وظایف در تمام سطوح سازمان و قسمت های مختلف آن را دارا می باشد. ارتباطات و اتصالات الکترونیکی می تواند اختیارات و قدرت تصمیم گیری کارکنان را به وسیله گسترش توانایی ها و قابلیت های محوری سازمان دگرگون نماید [۱۴].

رقابتی شدن محیط های تجاری با تأکید بر روی کارایی عملیات موجب اصلاح کیفیت محصول و خدمت و پاسخگویی می شود، که بسیاری از سازمان ها برای حصول این اهداف از فناوری اطلاعات استفاده می کنند. در نتیجه این تفکر، سازمان ها سرمایه گذاری قابل توجهی را روی فناوری اطلاعات انجام داده اند که بالغ بر ۵۰٪ سرمایه گذاری های سالیانه می باشد [۱۴].

فناوری های نو ظهور و پیشرفته برای حل مسائل مدیریتی در حال تغییر دادن شیوه های ساخت سازمان های مهندسی مجدد و مدیریت سازمان ها هستند. مدیریت برای تصمیم گیری نیاز به اطلاعات مناسب دارد که دانش او را افزایش و عدم اطمینان وی را کاهش دهد تا فعالیت مؤثری داشته باشد. لذا توجه به IT و کیفیت در پیرو بهبود کل سیستم ضروری بنظر می رسد [۱۰].

### ۴- داده، اطلاعات و دانش

بین داده ها و اطلاعات همانند خبر و اطلاع رابطه وجود دارد. خبری که دریافت می شود، پس از ارزیابی به اطلاع تبدیل می شود. داده ها نیز پردازش می شوند تا اطلاعات را پدید آورند. به بیان دیگر اطلاع حاصل تکامل داده ها است. به این ترتیب بین داده ها و اطلاعات یک شکاف وجود دارد که اندازه



این شکاف با حجم داده‌ها ارتباط مستقیم دارد. هر چه داده‌ها حجیم‌تر باشند، این شکاف بیشتر خواهد بود. امروزه افزایش سریع حجم پایگاه داده‌ها به شکلی است که توانایی انسان برای درک این داده‌ها بدون ابزارهای پر قدرت و فناوری اطلاعات میسر نمی‌باشد. در این وضعیت، تصمیم‌گیری‌ها به جای تکیه بر اطلاعات بر درک مدیران و کاربران تکیه دارند، چرا که تصمیم‌گیرندگان ابزار قوی برای استخراج اطلاعات با ارزش را در دست ندارند. بدون تردید تصمیم‌گیری در این وضعیت به احتمال زیاد روی کیفیت محصولات/خدمات و همچنین کیفیت فرآیندهای سازمان تأثیر خواهد گذاشت. در واقع شرایط فعلی توصیف‌کننده حالتی است که ما از لحاظ داده غنی، اما از لحاظ اطلاعات ضعیف هستیم [۱۳].

رشد اطلاعات و دانش در سال‌های اخیر بسیار سریع و نمایی بوده است و عصر حاضر را عصر انفجار اطلاعات نامیده‌اند. به طوری که ۸۰٪ یافته‌های فن آوری و دانش و نیز ۹۰٪ تمام دانش‌ها و اطلاعات فنی در جهان در قرن ۲۰ تولید شده است. این در حالی است که در هر پنج سال و نیم حجم آن دو برابر می‌گردد [۱۳]. جوامع کنونی به دو گروه ثروتمندان اطلاعاتی و فقرا اطلاعاتی تقسیم می‌شوند، که گروه اول به ابزار پیشرفت و توسعه مسلح‌اند و پیش‌تاز میدان تلقی می‌شوند و گروه دوم نیز دنباله روی آن‌ها خواهند بود [۱۰].

#### ۴-۱- کیفیت و مدیریت دانش و اطلاعات

یکی از عوامل توانمند ساز در مدل تعالی سازمانی حوزه شراکت‌ها و منابع می‌باشد، که سازمان چگونگی هماهنگی و اداره امکانات خارجی و منابع داخلی خود را جهت حمایت از خطی مشی، راهبرد‌ها و افزایش کارایی عملیات فرآیندهایش شناسایی و مورد توجه قرار می‌دهد. این حوزه شامل ۵ معیار است، که یکی از معیارها مدیریت اطلاعات و دانش می‌باشد. این معیار نیز شامل نکات زیر است [۸]:

۱. جمع‌آوری، سازماندهی و مدیریت اطلاعات و دانش با هدف پشتیبانی از خطی مشی و راهبرد سازمانی.
۲. تلاش برای بدست آوردن، افزایش استفاده موثر از دانش.
۳. ایجاد امکان دسترسی برای استفاده‌کنندگان داخلی و خارجی به اطلاعات و دانش مرتبط.
۴. تضمین و بهبود، صحت یکپارچگی و امنیت اطلاعات.
۵. پرورش، توسعه و حفاظت مناسب‌داری‌های فکری جهت پیشینه‌کردن ارزش افزوده برای مشتریان سازمان.
۶. ایجاد تفکر خلاق و نوآورانه در سازمان با استفاده از منابع اطلاعاتی و دانش مرتبط.



هریک موارد بالا زمانی در یک سازمان تحقق خواهند یافت که فناوری اطلاعات بعنوان یک ابزار اساسی در نظر گرفته شود.

## ۵- داده کاوی

با رشد فناوری اطلاعات و تکنیک های تولید و جمع آوری داده ها، پایگاه داده های مربوط به داده های تبادلات تجاری، کشاورزی، اینترنت، جزئیات مکالمات تلفنی، داده های پزشکی و غیره سریعتر از هر روز جمع آوری و انبارش می شوند. از اواخر دهه ۸۰ میلادی بشر به فکر دستیابی به اطلاعات نهفته در این پایگاه داده های حجیم افتاد زیرا سیستم های سنتی قادر به این کار نبودند [۱۳].

بدلیل رقابت در عرصه های سیاسی، نظامی، اقتصادی و علمی و اهمیت دستیابی به اطلاعات در کمترین زمان بدون دخالت انسان علم و تجزیه و تحلیل داده ها یا داده کاوی پا به عرصه گذاشت. داده-کاوی فرآیندی است که در آغاز دهه ۹۰ مطرح شد و با نگرشی نو، به مسئله استخراج اطلاعات از پایگاه داده ها می پردازد. از سال ۱۹۹۵ داده کاوی بصورت جدی وارد مباحث آمار شد و در سال ۱۹۹۶، اولین شماره مجله کشف دانش و معرفت از پایگاه داده ها<sup>۱</sup> منتشر شد. در حال حاضر، داده کاوی مهمترین فناوری جهت بهره برداری موثر از داده های حجیم است و اهمیت آن رو به فزونی است. بطوریکه تخمین زده شده است که مقدار داده ها در جهان هر ۲۰ ماه به حدود دو برابر می رسد. در یک تحقیق که بر روی گروه های تجاری بسیار بزرگ در جمع آوری داده ها صورت گرفت مشخص گردید که ۱۹ درصد از این گروه ها دارای پایگاه داده هایی با سطح بیشتر از ۵۰ گیگا بایت می باشند و ۵۹ درصد از آنها انتظار دارند که در آینده ای نزدیک در چنین سطحی قرار گیرند [۲].

در سال ۱۹۸۹ و ۱۹۹۱ کارگاه های کشف دانش و معرفت از پایگاه داده ها، توسط پیاتسکی<sup>۲</sup> و همکارانش برگزار شد. در فواصل سالهای ۱۹۹۱ تا ۱۹۹۴ کارگاه های کشف دانش و معرفت از پایگاه داده ها توسط فیاد<sup>۳</sup> و پیاتسکی و دیگران برگزار شد. بطور رسمی اصطلاح داده کاوی برای اولین بار توسط فیاد در اولین کنفرانس بین المللی کشف معرفت و داده کاوی، در سال ۱۹۹۵ مطرح شد [۱۳].

افزایش داده های بسیار باعث پیدایش فرصت های تازه برای کار در علوم مهندسی و کسب و کار شده است. زمینه داده کاوی بعنوان یک رشته علمی جدید در مهندسی و علوم کامپیوتر ظهور کرده است و مهندسی صنایع با حوزه های گوناگون و در بر داشتن فرصت های بی نظیر برای کاربرد داده کاوی و کشف دانش از پایگاه داده ها و توسعه مفاهیم و تکنیک های تازه در این زمینه آماده است.

<sup>1</sup> : Knowledge Discovery in Database(KDD)

<sup>2</sup> : Piatetsky

<sup>3</sup> : Fayyad



فرآیند های صنعتی زیادی اکنون برای مطمئن شدن از کیفیت سفارشات محصول و کاهش هزینه های محصول بطور اتوماتیک از فناوری اطلاعات استفاده می کنند [۱].

#### ۵-۱- روش های گرد آوری اطلاعات

با توجه به اطلاعات مورد نیاز هر شرکت، اطلاعات سودمند را چگونه می توان به صورت موثر و کارآمد گرد آوری نمود. اطلاعات با خود سود مندی و هزینه سازی را به همراه می آورد و ممکن است در گرد آوری داده ها زیاده روی شود. شرکت هایی هستند که در اطلاعات غرق شده ولی همچنان تشنه دانش و آگاهی هستند، در واقع میان داده ها، اطلاعات، دانش و خرد تفاوت گسترده و آشکاری است. تا هنگامی که داده به اطلاعات پرورده، و سپس به دانش دست آخر به خرد و شناخت از بازار تبدیل نشوند، هدر رفته و بی مصرف هستند. مدیریت بازار برای گردآوری اطلاعات سه راه مشاهده، داده های دست اول و داده های دست دوم را پیش روی دارد [۶].

#### ۵-۲- داده کاوی چیست؟

ترجمه تحت اللفظی داده کاوی، به ما در درک بهتر این واژه کمک می کند. Mine به معنای استخراج از منابع نهفته و با ارزش زمین اطلاق می شود. پیوند این کلمه با کلمه داده، جستجوی عمیق جهت پیدا کردن اطلاعات اضافی مفید که قبلاً نهفته بودند، از داده های قابل دسترس حجیم، را پیشنهاد می کند. داده- کاوی یک رشته نسبتاً جدید علمی می باشد که از انجام تحقیقات در رشته های آمار، یادگیری ماشینی، علوم کامپیوتر خصوصاً مدیریت پایگاه داده ها شکل گرفته است [۴]. تعریفی که در اکثر مراجع به اشتراک ذکر شده عبارت است از "استخراج اطلاعات و دانش و کشف الگوهای پنهان از یک پایگاه داده های بسیار بزرگ و پیچیده" [۵]. داده کاوی یک متدولوژی بسیار قوی و با پتانسیل بالا می باشد که به سازمانها کمک می کند که بر روی مهمترین اطلاعات از مخزن داده های خود تمرکز نمایند [۲].

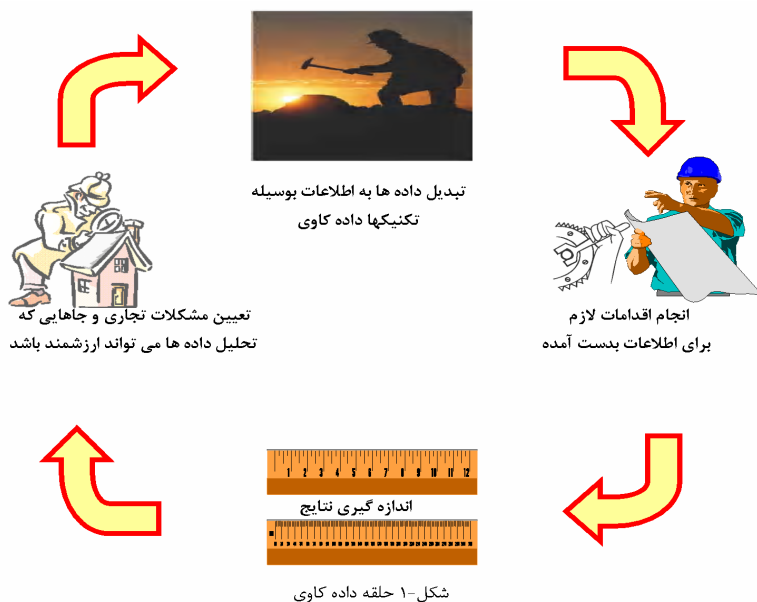
داده کاوی فرآیندی است که از ابزارهای تحلیلی گوناگونی برای کشف الگوها و روابط بین داده ها استفاده می کند که ممکن است برای اعتبار بخشیدن به پیش بینی استفاده شود [۵]. داده کاوی کمک می کند تا سازمانها با کاوش بر روی داده های یک سیستم، الگوها، روندها و رفتارهای آینده را کشف و پیش بینی کرده و بهتر تصمیم بگیرند. داده کاوی با استفاده از تحلیل وقایع گذشته یک تحلیل اتوماتیک و پیش بینانه ارایه می نماید و به سوالاتی جواب می دهد که پاسخ آنها در گذشته ممکن نبوده و یا به زمان زیادی نیاز داشت. ابزارهای داده کاوی الگوهای پنهانی را کشف و پیش بینی می کنند که متخصصان ممکن است به دلیل اینکه این اطلاعات و الگوها خارج از انتظار آنها باشد، آنها را مدنظر قرار ندهند و به آنها دست نیابند [۱۳].



### ۳-۵- حلقه داده کاوی چیست؟

کلید موفقیت داده کاوی در آمیختن آن با فرآیند تجارت و کسب و کار می باشد. این مراحل بسیار به هم وابسته می باشند و نتایج هر مرحله مانند یک فرآیند تولید چند مرحله ای یک ورودی برای مرحله بعد می باشد. چهار قدم اصلی حلقه داده کاوی عبارتند از [۱۳]:

- ۱- صورت مساله و مشکل تجاری را تعریف کنیم.
- ۲- از داده کاوی استفاده کنیم تا داده ها را به اطلاعات و دانش تبدیل کنیم.
- ۳- اقدامات لازم را در مورد اطلاعات کسب شده انجام دهیم.
- ۴- نتایج را اندازه گیری کنیم.



### ۱-۳-۵- شناسایی فرصتهای تجاری

شناسایی فرصتهای تجاری یک قدم بسیار مهم می باشد که هنگامی در سازمان انجام می شود که افزایش اطلاعات بتواند سبب انجام بهتر کارها شود و هدف آن تعیین زمینه ها و نواحی است که داده ها می توانند باعث افزایش ارزش شوند و که الگوهای درون داده ها پتانسیل افزایش ارزش را داشته باشد. بعنوان مثال شرکت فرنچ (French)، سازنده کامیون های دیزلی تصمیم گرفت از داده های ثبت شده در طول تست موتور های دیزلی تولید شده بهره برداری کند. هدف از این مطالعه بهینه کردن یک فرآیند صنعتی که شامل چک و تنظیم کنترل شده الکترونیکی موتورهای دیزلی بود. در واقع هدف کشف دانش از مجموعه داده های فرآیند تست موتور به معنی کاهش زمان فرآیند سفارشات (در حدود ۲۵٪) بود. این شرکت روش کشف دانش از پایگاه داده ها را مورد استفاده قرار داد. این مطالعه به شرکت اجازه

داد تا دو سیستم کاربرد کشف برای بدست آوردن مدل تکمیل پیش بینی (محاسبه) زمان واقعی (با یک کاهش حقیقی ۲۸٪) و کشف پشتیبان محیط را را توسعه دهد [۱۰].

بعضی فرآیندهای تجاری از پیش به تحلیل حجم زیادی از داده ها برای انجام عملیات خود تکیه دارند و بعنوان مرحله شناسایی چرخه تعیین می گردند. در این موارد فرصتهای تجاری بسادگی شناسایی شده و داده کاوی باید به صورتی یکپارچه و توأم با این فرآیندها انجام شود [۴]. از جمله این فرآیندهای تجاری عبارتند از:

- بخش بندی بازار و مشتریان و شناسایی بازار هدف
- برنامه ریزی بازاریابی برای محصولات جدید
- قیمت گذاری محصولات و خدمات جدید
- شناخت نارضایتی های مشتریان

#### ۵-۳-۲- داده کاوی

داده کاوی تکنیکی است که فرصتهای تجاری سازمان را دریافت نموده و به دانش و اطلاعات برای انجام اقدامات لازم تبدیل می کند.

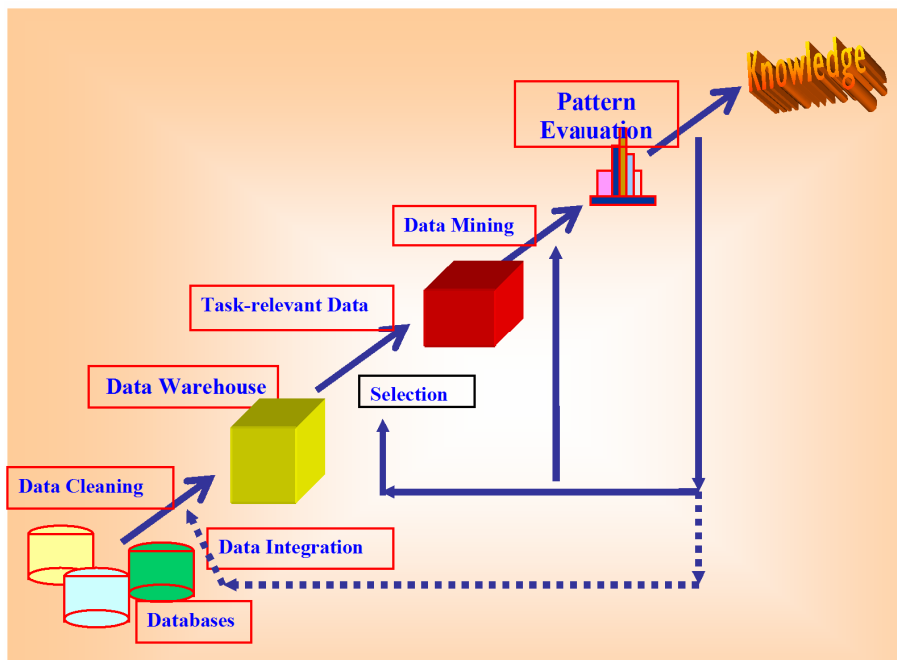
#### ۵-۳-۲-۱- مراحل داده کاوی

قدم داده کاوی در این چرخه خود شامل مراحل زیر است که در شکل ۲- نشان داده شده است [۱۳].

۱. تعیین اطلاعات گذشته.
۲. تمیز کردن داده ها و پردازش اولیه [۲].
۳. یکپارچه سازی داده ها.
۴. انتخاب مجموعه داده های هدف.
۵. یافتن ویژگیهای مورد استفاده و تعیین ویژگیهای جدید.
۶. نمایش داده ها بصورتیکه بتوان برای داده کاوی استفاده نمود.
۷. انتخاب عملیات داده کاوی (دسته بندی، خوشه بندی، پیش بینی و ...).
۸. انتخاب روش داده کاوی (شبکه های عصبی، درخت تصمیم، خوشه بندی و ...).
۹. داده کاوی و جستجو برای یافتن الگوی مناسب.
۱۰. ارزیابی و تحلیل الگوی بدست آمده و حذف الگوهای نامناسب.
۱۱. تفسیر نتایج داده ها و استنتاج از اطلاعات با ارزش.

### ۵-۳-۳- انجام اقدامات

در این مرحله نتایج و اطلاعات حاصل از داده کاوی گرفته شده و اقدامات لازم انجام می شود و خروجی های این اقدامات در مرحله بعد اندازه گیری می شود.



شکل ۲- مراحل داده کاوی

### ۵-۳-۴- اندازه گیری نتایج

اندازه گیری در واقع باعث فراهم آمدن بازخوردهایی جهت بهبود مستمر می شود. اندازه گیری توجیهی بر علت انجام داده کاوی و سندی بر موفقیت یا عدم موفقیت آن می باشد. با وجود این مرحله از داده کاوی معمولاً در سازمانها نادیده و کم اهمیت در نظر گرفته می شوند.

### ۵-۴- آمار و داده کاوی [۱۳]

داده کاوی اساساً یک رشته کاربردی است و یک داده کاو باید از روشهای آماری درک خوبی داشته باشد. در داده کاوی تلاش می شود بین آمار و علوم رایانه ای رابطه ای برقرار گردد. روابط در داده کاوی غالباً بصورت الگوها و مدلهایی از قبیل معادلات رگرسیونی، سریهای زمانی، خوشه ها، رده بندی ها، گرافها و غیره ارائه می شوند. در داده کاوی نیز همانند آمار غالباً داده هایی که تحلیل می شوند، نمونه ای از جامعه هستند که به تبع بزرگ بودن جامعه با نمونه ای حجیم مواجه هستیم. در هنگام کار با مجموعه داده های حجیم مشکلات تازه ای بروز می کند. آمار و داده کاوی هر دو با روشهای تحلیل و

مدل‌بندی داده‌ها مرتبط می‌باشند. بنابر این اشتراک زیادی بین این دو رشته وجود دارد. به عنوان یک شوخی، یکی از نویسندگان در پاسخ سوال اینکه داده‌کاوی چیست؟ بیان می‌کند که همان آمار است، اما با یک نام خیلی بهتر. آماردان‌ها با داده‌های دست اول، که برای تحقیق درستی فرضهای خاصی جمع آوری و تولید شده‌اند، کار می‌کنند. اما داده‌کاوها با تحلیل داده‌های دست دوم و یا داده‌های مشاهده‌ای که اغلب از منابع مختلفی گردآوری شده‌اند، کار می‌کنند. منظور پیدا کردن وقایع مورد علاقه و اطلاعات مفیدی است، که در داده‌ها مخفی شده‌اند و اغلب با اهداف ابتدایی که داده‌ها بخاطر آن جمع‌آوری شده‌اند، رابطه‌ای ندارند. از طرفی داده‌کاوی با حجم وسیعی از داده‌ها سروکار دارد و همچنین بعضی از پایگاه داده‌ها ساختار مناسب داده‌های آماری را ندارند.

#### ۵-۵- داده‌کاوی چه کاری انجام می‌دهد؟

داده‌کاوی نام خود را از شباهت بین جستجو برای یافتن اطلاعات با ارزش تجاری در یک پایگاه اطلاعات بزرگ و کاویدن کوه برای یافتن رگه‌های اصلی سنگ معدن گرفته است [۳]. هر دو فرآیند شامل حرکت و یا جستجوی هوشمندانه در یک حجم بسیار بزرگ بمنظور یافتن ارزش می‌باشد. در جدول شماره ۲- عملیات، برخی از کاربردها و روشهای داده‌کاوی به تفکیک آورده شده است.

تکنیکها	کاربردها	عملیات
۱- شبکه‌های عصبی	۱- دسته‌بندی مشتریان	۱- دسته‌بندی
۲- درخت تصمیم‌گیری	۲- نگهداری مشتریان	۲- پیش‌بینی
۳- نزدیکترین همسایگی	۳- شناسایی تخلفات	۳- خوشه‌بندی
۴- تحلیل خوشه	۴- تحلیل سایتهای اینترنتی	۴- تخمین
۵- استنتاج قانون	۵- تحلیل داده‌ها در بازاریابی	۵- تعیین ارتباط متغیرها
۶- الگوریتم ژنتیک	۶- تحلیل و پیش‌بینی روندهای بازار بورس	۶- .....

جدول شماره ۲- عملیات، کاربردها و روش‌های داده‌کاوی

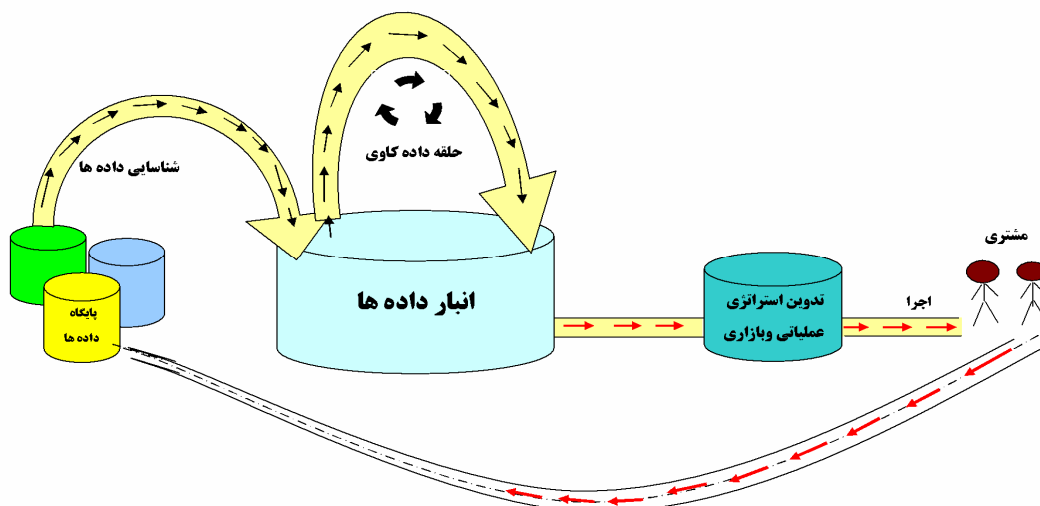
#### ۶- متدولوژی: کاربرد داده‌کاوی در سیستم‌های کیفیت

امروزه حجم زیادی از داده‌های مربوط به سازمان‌ها و مشتریان در پایگاه داده‌ها جمع‌آوری و انبارش می‌شوند. این پایگاه داده‌ها همانند یک معدن با ارزش می‌باشند که باید منابع با ارزشی نظیر اطلاعات و دانش از آن کشف و استخراج گردد. استخراج دانش از این پایگاه داده‌ها مستلزم استفاده از یک تکنیک و فناوری می‌باشد زیرا روشهای سنتی و آماری قادر به پاسخ‌گویی به این حجم داده‌های

انبارش شده نیستند. لذا استفاده از فناوری اطلاعات و ابزارهای آن از جمله تکنیک داده کاوی برای حل مشکلات تجاری و خصوصاً برای کمک به نحوه مدیریت ارتباط با مشتری (CRM) ضروری بنظر می رسد.

### ۶-۱- بهبود، نوآوری و یادگیری مستمر

عملکرد سازمان زمانی به تعالی خواهد رسید که بر پایه استفاده از دانش ناشی از بهبود، نوآوری و یادگیری مستمر در سازمان باشد. روزانه وقایع و اتفاقات مهمی در سازمان ها رخ می دهد و شاید داده ها و اطلاعات مربوط به آنها ثبت می گردد که شاید برخورد علمی و تحلیلی با داده ها صورت نمی گیرد و نتایج حاصله از آنها به نحوه مناسب ثبت و ذخیره نمی شود. دانش چیزی نیست که در تعالی سازمانی، کمتر مورد توجه قرار گیرد، لذا مدیریت دانش یکی از ارکان تعالی سازمانی است [۸]. در متدولوژی پیشنهاد شده مطابق شکل ۳- ابتدا سازمان بامروری بر پایگاه داده های خود اقدام به شناسایی داده های لازم می کند. بدلیل اینکه داده ها از منابع مختلفی جمع آوری می شوند و معمولاً دارای قالب های متفاوتی هستند باید یک پردازش اولیه روی این داده ها صورت گیرد و بعضی از داده های نامناسب پاکسازی شوند. در مرحله بعد این داده ها به یک انبار داده ها انتقال می یابند. آنگاه برای مسأله مورد نظر سازمان داده های مورد نیاز، از انبار داده ها انتخاب می گردند تا عملیات داده کاوی روی آنها انجام گیرد.



شکل ۳- چرخه داده کاوی در سازمان

با انجام عملیات داده کاوی روی الگوهای ناشناخته را کشف و استخراج می شوند که این الگوها و اطلاعات در مرحله بعد با تجزیه و تحلیل تبدیل به دانش سازمانی می شوند. که از جمله این اعمال دسته

بندی مشتریان، خوشه بندی (بخش بندی) مشتریان، یا بخش بندی بازار شناسایی تخلفات، پیش بینی و تخمین روند های آتی می باشد. با شناخت الگو ها و دانش از انبار داده ها، سازمان طبق این اطلاعات و دانش اقدام به تهیه و تدوین استراتژی های بازاری و برنامه ریزی برای فعالیت های عملیاتی و بازاری خود می کند. سپس با اجرای این برنامه ریزی مطابق خطی و مشی و استراتژی های تدوین شده با مشتریان ارتباط برقرار می کند. داده های حاصله از عملکرد های عملیاتی و بازاری و ارتباط با مشتریان و نظرات مشتریان دوباره به پایگاه داده ها باز خورد داده می شود و دوباره ثبت و ذخیره و انبارش می شود. در واقع در مدل پیشنهاد شده دو حلقه همواره برقرار است، اولی حلقه مدیریت ارتباط با مشتری و دیگری حلقه داده کاوی می باشد. سازمان با تجزیه و تحلیل اطلاعات حاصل از این مدل می تواند اقدامات زیر را جهت بهبود در فرآیند های سازمانی انجام دهد.

۱- برنامه ریزی و اقدام اصلاحی

۲- آموزش کارکنان

نتایج حاصل از این چرخه تجزیه و تحلیل داده ها و اطلاعات منجر به ارتقاء بهره وری سازمان از جمله کارکنان، فرآیند های تولید و فروش خواهد شد. ارتقاء بهره وری عاملی برای بالا بردن کیفیت، رضایت مشتری مشتریان خواهد بود، که حد اعلاای رضایت مشتری منجر به وفاداری آنان خواهد شد.

## ۷- نتیجه

در این مقاله، نقش و اهمیت داده و اطلاعات در سازمان و افزایش و کسب دانش سازمانی برای فعالیت در محیط رقابتی و همچنین پشتیبانی تصمیمات نشان داده شده است. سازمان ها برای استقرار سیستم های کیفیت نیاز به تجزیه و تحلیل داده ها و اطلاعات درون و برون سازمانی دارند. فناوری اطلاعات بعنوان ابزاری که می تواند به سازمان ها کمک کند تا داده های تولید و انبارش شده حاصل از فعالیت و عملکرد سازمان را تجزیه و تحلیل کند، مطرح گردیده است. بدلیل تعمیم و انتقال کیفیت از محصول به فرآیند های سازمانی، بنگاه ها ناگزیر به تجزیه و تحلیل داده های درون سازمانی خود جهت ارتقاء کیفیت محصولات و خدمات برای رضایت مشتریان هستند.

## منابع

- [1] C. Gertisio and A. Dussauchoy, "Knowledge Discovery from Industrial Data base", Journal of Intelligent Manufacturing, 15, 29-37, 2004.

- [2] [http://www.thearling.com/an introduction to data mining.htm](http://www.thearling.com/an%20introduction%20to%20data%20mining.htm)
- [3] [http://www.thearling.com/-an overview of data mining techniques.htm](http://www.thearling.com/-an%20overview%20of%20data%20mining%20techniques.htm)
- [4] B. Alex, S. Stephen, and T. Kurt "Building Data Mining Applications for CRM", 2004, Tata McGraw-Hill, New York.
- [5] B. Michael and L. Gordon "Data Mining Techniques: For Marketing, Sales, and Customer Support" New York: John Wiley and Sons, 1997.
- [۶] کاتلر، فیلیپ، "کاتلر در مدیریت بازار"، ترجمه عبدالرضا رضائی نژاد، تهران: انتشارات فرا، ۱۳۷۹.
- [۷] دانپورت، تامس و لارنس، پروساک "مدیریت دانش"، ترجمه حسین رحمان سرشت، نشر ساپکو، ۱۳۷۹.
- [۸] امیران، حمید "راهنمای گام به گام سرآمد سازمانی بر اساس مدل های EFQM/INQA"، چاپ سوم، انتشارات شرکت مشاورین کیفیت سلز، تهران، ۱۳۸۳.
- [۹] کزازی، ابوالفضل "مدیریت کنترل کیفیت فراگیر"، نشر مرکز آموزش مدیریت دولتی، تهران، ۱۳۷۸.
- [۱۰] شارعی، علی "بررسی تأثیر فناوری اطلاعات و مدیریت کیفیت جامع بر کارایی و اثربخشی سازمان های صنعتی و خدماتی".
- [۱۱] فصلنامه مدیریت فرا/ ویژه نامه فناوری اطلاعات و ارتباطات و مدیریت آن، ش ۱، انتشارات تولید دانش، تهران، ۱۳۸۳.
- [۱۲] ماهنامه تدبیر/ ناکامی مدیران دانش در سازمان ها، ش ۱۶۱ ص ۲۶ تا ۲۸، تهران، ۱۳۸۴.
- [۱۳] توپچی، حسین "استفاده از داده کاوی برای ارایه محصول به مشتریان بر اساس وفاداری مشتریان" پایاننامه کارشناسی ارشد مهندسی صنایع، به راهنمایی محسن اکبرپور شیرازی و ایرج مهدوی، دانشگاه علوم و فنون مازندران، ۱۳۸۴.
- [14] [http://www. Iranian Information Center of Industries & Mines.com](http://www.Iranian%20Information%20Center%20of%20Industries%20&%20Mines.com)