تاریخ دریافت: ۱۳۹۴/۰۴/۱۰
تاریخ پذیرش: ۱۳۹۴/۰۹/۱۳

فصلنامه علوم و فنون نظامی/ سال یازدهم/ شماره ۳۱، بهار ۱۳۹۴
صص ۴۷-۲۹

کاربردهای داده‌کاوی در تصمیم‌گیری فرماندهان و مدیران نظامی

محسن مرادیان[۱]
حسین مؤمنی فرد[۲]
محمدتقی پرتوی[۳]

چکیده:

اطلاعات در نیروهای نظامی نقشی اساسی دارند بطوریکه مبنای فعالیت همه ارکان نظامی اطلاعات محسوب می‌شود، بر همین اساس جمع‌آوری اطلاعات به‌صورت مستمر و ایجاد بانک‌های اطلاعاتی برای نیروهای نظامی هر کشوری امری انکارناپذیر است. از زمان بهره‌برداری سازمان‌های مختلف از رایانه؛ یکی از کاربردهای اصلی آن ذخیره داده‌ها و اطلاعات گوناگون درون سازمانی و محیطی است. وجود پایگاه داده‌های مختلف در سازمان‌ها امکان داده‌کاوی و کشف دانش پنهان و استخراج دانش جدید را فراهم نموده است. در این مقاله با توجه به اهمیت اطلاعات در ارتش و نیروهای نظامی، جایگاه و کاربردهای داده‌کاوی تشریح گردیده است. در ادامه زمینه‌های کاربردی وسیع آن در حیطه‌های مختلف امور نظامی بیان‌شده و مقایسه تطبیقی از به‌کارگیری این دانش در ارتش‌های مدرن صورت گرفته است و درنهایت با توجه به تجزیه‌وتحلیل و نتایج حاصله از تحقیق مبنی بر ضرورت بهره‌مندی از داده‌کاوی پیشنهادهای کاربردی برای بهره‌گیری فرماندهان و مدیران عالی نیروهای نظامی کشور در تصمیم‌گیری‌های سطوح راهبردی و عملیاتی به‌عمل‌آمده است.

واژگان کلیدی:

داده‌کاوی، تصمیم‌گیری، فرماندهان نظامی

۱- استادیار و عضو هیئت‌علمی دانشگاه عالی دفاع ملی
۲- دانشجوی دوره دکتری عملیات مشترک و مرکب و عضو هیئت‌علمی دانشگاه فرماندهی و ستاد آجا
۳ - دانشجوی دکتری مدیریت (تحقیق در عملیات) دانشگاه تهران و عضو هیئت‌علمی دانشگاه فرماندهی و ستاد آجا

مقدمه

دانش داده‌کاوی یکی از ده دانش درحال‌توسعه‌ای است که دهه آینده را با انقلاب فناورانه مواجه خواهد ساخت لذا در سال‌های اخیر در دنیا گسترش فوق‌العاده سریعی داشته است. امروزه در دنیای توسعه‌یافته، هیچ موضوعی بدون بهره‌گیری از دانش داده‌کاوی یافت نمی‌شود، به‌گونه‌ای که این دانش در تمامی زمینه‌های مختلف در سازمان‌ها و کشورها نقش دارد. دانش داده‌کاوی فرآیند کشف دانش پنهان درون داده‌ها است، که با برخورداری از دامنه وسیع در زمینه‌های تخصصی با توصیف، تشریح، پیش‌بینی و کنترل پدیده‌های گوناگون پیرامونی، امروزه دارای کاربرد بسیار وسیع در حوزه‌های مختلف ازجمله صنعتی، پزشکی، ارتباطات، کشاورزی، انرژی، علوم اجتماعی، فرهنگی، سیاسی، اقتصادی، بازرگانی، نظامی و آموزشی و غیره است به‌گونه‌ای که امروزه مرز و محدودیتی برای کاربرد دانش داده‌کاوی و کشف دانش در نظر گرفته نشده و زمینه‌های کاری این دانش را در تمامی عرصه‌های برخوردار از داده می‌دانند.

داده‌کاوی[1] پل ارتباطی میان علم آمار، علم کامپیوتر، هوش مصنوعی، الگو شناسی، فراگیری ماشین و بازنمایی بصری داده می‌باشد. داده‌کاوی فرآیندی پیچیده جهت شناسایی الـــگوها و مـدل‌های صحیح، جدید در حجم وسیعی از داده می‌باشد، به طریقی که این الگوها و مدل‌ها برای انسان‌ها قابل‌درک باشند. داده‌کاوی به‌صورت یک محصول قابل خریداری نمی‌باشد، بلکه یک‌رشته علمی و فرآیندی است که بایستی به‌صورت یک پروژه در سازمانهای مختلف پیاده‌سازی شود. (غضنفری، علیزاده، ۱۳۸۷ :۳).

داده‌ها اغلب حجیم می‌باشند و به‌تنهایی قابل‌استفاده نیستند، بلکه دانش نهفته در داده‌ها قابل‌استفاده می‌باشد. بنابراین بهره‌گیری از قدرت فرآیند داده‌کاوی جهت شناسایی الگوها و مدل‌ها و نیز ارتباط عناصر مختلف در پایگاه داده جهت کشف دانش نهفته در داده‌ها و نهایتاً تبدیل داده به اطلاعات، روزبه‌روز ضروری‌تر می‌شود.

ریشه‌های داده‌کاوی در میان سه حوزه از علوم، قابل‌پیگیری می‌باشد. مهم‌ترین این خانواده‌ها، آمار کلاسیک[2] می‌باشد. بدون آمار، هیچ داده‌کاوی وجود نخواهد داشت، به‌طوری‌که آمار، اساس اغلب تکنولوژی‌هایی می‌باشد که داده‌کاوی بر روی آن‌ها بنا می‌شود. آمار کلاسیک مفاهیمی مانند تحلیل رگرسیون، توزیع استاندارد، انحراف استاندارد، واریانس،

1. Data Mining
2. Classic Statistics

تحلیل خوشه، و فاصله‌های اطمینان را که همه این موارد برای مطالعه داده و ارتباط بین داده‌ها می‌باشد، را در برمی‌گیرد. مطمئناً تحلیل آماری کلاسیک نقش اساسی در تکنیک‌های داده‌کاوی ایفا می‌کنند (ملک‌محمدی، ۱۳۸۵: ۸).

دومین حوزه‌ای که داده‌کاوی به آن تعلق دارد هوش مصنوعی[1] می‌باشد. هوش مصنوعی که بر پایه روش‌های ابتکاری می‌باشد و با در تقابل با آمار قرار دارد، تلاش دارد تا فرایندی مانند فکر انسان، را برای حل مسائل آماری بکار بندد. چون این رویکرد نیاز به توان محاسباتی بالایی دارد، تا اوایل دهه ۱۹۸۰ عملی نشد. هوش مصنوعی کاربردهای اندکی را در حوزه‌های علمی، حکومتی و حتی نظامی پیدا کرد، اما نیاز به استفاده از کامپیوترهای بزرگ باعث شد همه افراد نتوانند از تکنیک‌های ارائه‌شده استفاده کنند.

سومین حوزه داده‌کاوی، یادگیری ماشین[2] می‌باشد، که به مفهوم دقیق‌تر، اجتماع آمار و هوش مصنوعی می‌باشد. درحالی‌که هوش مصنوعی نتوانست موفقیت کاربردی عمومی کسب کند، یادگیری ماشین در بسیاری از موارد جایگزین آن گردید. از یادگیری ماشین به‌عنوان تحول هوش مصنوعی یاد شد، چون مخلوطی از روش‌های ابتکاری هوش مصنوعی به همراه تحلیل آماری پیشرفته می‌باشد. یادگیری ماشین اجازه می‌دهد تا برنامه‌های کامپیوتری در مورد داده‌ای که آن‌ها مطالعه می‌کنند، مانند برنامه‌هایی که تصمیم‌های مختلف را بر مبنای کیفیت داده مطالعه شده می‌گیرند، یادگیری داشته باشند و برای مفاهیم پایه‌ای آن از آمار استفاده می‌کنند و الگوریتم‌ها و روش‌های ابتکاری هوش مصنوعی را برای رسیدن به هدف بهره می‌گیرند. (ملک‌محمدی، ۱۳۸۵: :۲۶)

مبانی نظری تحقیق:

قبل از اینکه در مورد داده‌کاوی بیشتر صحبت شود، ابتدا مفاهیم بنیادی مرتبط با این حوزه علمی به صورت اجمالی توضیح داده می‌شود.

مفهوم داده، اطلاعات و دانش[3]

داده: داده‌ها حقیقت، اعداد، یا متونی می‌باشند که با کامپیوتر بتوان آن‌ها را پردازش نمود. امروزه سازمان‌ها حجم عظیمی از فرمت‌های مختلف داده و بانک‌های اطلاعاتی متفاوت را جمع‌آوری و ذخیره می‌کنند. این داده‌ها شامل داده‌هایی از قبیل موجودی انواع مواد اولیه و

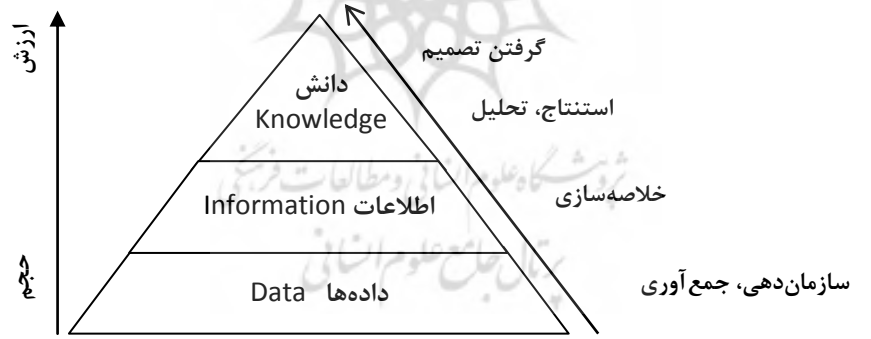
1.Artificial Intelligence
2.Machine Learning
3. Data, Information, Knowledge

کالاها، اطلاعات کارکنان، اطلاعات تجهیزات، اطلاعات مالی و حسابداری، داده‌های پیش‌بینی‌شده، داده‌های اقتصادی و داده‌هایی درباره خود داده‌ها مانند طراحی منطقی بانک‌های اطلاعاتی و غیره می‌باشد. (صادقی، ۱۳۷۷: ۴۸)

اطلاعات: الگو، وابستگی، یا ارتباط میان تمام داده‌ها و همچنین داده‌های پردازش شده را اطلاعات گویند. برای مثال تحلیل اطلاعات نیروی انسانی می‌تواند اطلاعاتی در مورد تعداد، نوع، مهارت و توانمندی، تخصص کارکنان و آخرین وضعیت آنها را فراهم آورد.

دانش: اطلاعات می‌تواند تبدیل به دانش شوند که این دانش در مورد الگوهای گذشته و گرایش‌های آینده می‌باشد. از دانش کسب‌شده می‌توان برای اتخاذ تصمیم استفاده کرد. برای مثال، خلاصه‌سازی اطلاعات می‌تواند به ما بگوید که رفتار خرید در یک فروشگاه چگونه است و به کمک شناخت این رفتار فروش کالا را افزایش داد (غضنفری، ۱۳۸۵: ۶).

در شکل ۱ جایگاه و ارتباط بین مفاهیم بیان شده نشان‌داده شده است، می‌باشد، همانطوریکه ملاحظه می‌گردد در هر سازمانی داده‌های زیادی وجود دارند که در صورت پردازش، سازماندهی و جمع‌آوری اطلاعات به اطلاعات کاربردی باارزش‌تری برای آن سازمان تبدیل می‌شوند و در صورت خلاصه سازی، ایجاد مدلهای تصمیم‌گیری مناسب و



استنتاج و تحلیل به دانش باارزش‌تری تبدیل می گردند (صادقی، ۱۳۷۷ : ۴۹).

شکل ۱- جایگاه داده، اطلاعات و دانش در مراحل تصمیم‌گیری و تصمیم‌سازی سازمان

تعریف داده‌کاوی[1]

در متون دانشگاهی تعاریف گوناگونی برای داده‌کاوی ارائه‌شده که برخی از مهمترین تعاریف ارائه شده بشرح زیر است:

1. Data Mining

- داده‌کاوی پل ارتباطی میان علم آمار، علم کامپیوتر، هوش مصنوعی، الگو شناسی، فراگیری ماشین و بازنمایی بصری داده می‌باشد. داده‌کاوی به‌صورت یک محصول قابل خریداری نمی‌باشد، بلکه یک فرآیندی است که بایستی به‌صورت یک پروژه پیاده‌سازی شود.
- داده‌کاوی فرآیندی پیچیده، جهت شناسایی الگوها و مدل‌های صحیح، جدید و به‌صورت بالقوه مفید، در حجم وسیعی از داده می‌باشد، به طریقی که این الگوها و مدل‌ها برای انسان‌ها قابل‌درک باشند. (غضنفری، ۱۳۸۷: ۸)
- داده‌کاوی به معنای کشف دانش درون داده‌هاست، کشف دانش دروندادها آن‌هم در عصر اطلاعات یکی از هیجان‌انگیزترین و کلیدی‌ترین مفاهیمی است که روزبه‌روز اهمیت بیشتری می‌گیرد.
- داده‌کاوی به بررسی و تجزیه‌وتحلیل مقادیر عظیمی از داده‌ها به‌منظور کشف الگوها و قوانین معنی‌دار گفته می‌شود.
- داده‌کاوی، استخراج اطلاعات مفهومی، ناشناخته و به‌صورت بالقوه مفید از پایگاه داده می‌باشد.
- داده‌کاوی استخراج نیمه اتوماتیک الگوها، تغییرات، وابستگی‌ها، نابهنجاری‌ها و دیگر ساختارهای معنی‌دار آماری از پایگاه‌های بزرگ داده می‌باشد. (Berson, 2004)

بنابراین به‌طورکلی داده‌کاوی که گاهی اوقات آنرا اکتشاف اطلاعات یا دانش می‌نامند؛ عبارت از فرآیندی است که از چشم‌اندازهای مختلف به تحلیل داده‌ها می‌پردازد و جمع‌بندی آن‌ها را در قالب اطلاعات مفیدی ارائه می‌کند. این اطلاعات را می‌توان برای افزایش کارایی سازمان، بهبود بهره‌وری افزایش درآمد، کاهش هزینه‌ها یا همه آنها بکار برد. به لحاظ فنی، داده‌کاوی عبارت از فرآیندی است که در میان حوزه‌های گوناگون بانک‌های اطلاعاتی بزرگ، همبستگی‌ها یا الگوهایی را پیدا می‌کند (ملک محمدی، ۱۳۸۵ : ۳۲).

حوزه‌های مختلف مانند کتابخانه دیجیتال، آرشیو تصاویر، اطلاعات زیستی، تصاویر پزشکی، اطلاعات مالی و سرمایه‌گذاری، ساخت و تولید، کسب‌وکار و بازاریابی، شبکه راه دور و اینترنت وجود دارد که در آن‌ها حجم بسیاری از داده در پایگاه‌های داده متمرکز یا توزیع‌شده ذخیره می‌شود.

داده‌ها اغلب حجیم می‌باشند و به‌تنهایی قابل‌استفاده نیستند، بلکه دانش نهفته در داده‌ها قابل‌استفاده هست؛ بنابراین بهره‌گیری از قدرت فرآیند داده‌کاوی جهت شناسایی الگوها و مدل‌ها و نیز ارتباط عناصر مختلف در پایگاه داده جهت کشف دانش نهفته در داده‌ها و نهایتاً تبدیل داده به اطلاعات، روزبه‌روز ضروری‌تر می‌شود.

دلایل اصلی لزوم استفاده از داده‌کاوی در تصمیم‌گیری فرماندهان نظامی

- حجم داده‌ها با سرعت زیادی در حال افزایش است.
- اطلاعات موجود برای تصمیم‌گیری در مورد این داده‌ها کم است لذا اطلاعات کافی از این داده‌ها استخراج نمی‌شود.
- دانش فرماندهان ما نسبت به این اطلاعات در حد صفر است.

جالب اینجاست که مشکل تمام این‌ها یک کلمه است و آن هم داده‌کاوی است.

توسعه فناوری‌های ذخیره و بازیابی اطلاعات تنها راه‌کار ممکن برای محقق شدن اهداف داده‌کاوی و کشف دانش می‌باشد. همچنین مسائلی از قبیل افزایش روزافزون حجم اطلاعات ذخیره‌شده در بانکهای اطلاعاتی سازمانها، تنوع بسیار زیاد اطلاعات ذخیره شده، افزایش و توسعه بانک‌های اطلاعاتی، افزایش بکارگیری فایل‌های چندرسانه‌ای مثل تصاویر متحرک و فایل‌های صوتی و اطلاعات متنی و فاقد ساختار مواردی هستند که تنها با استفاده از داده‌کاوی می توان تصمیم‌گیری سازمانی را بهینه سازی نمود. (Gyordi, 2004)

آرشیوهای اطلاعاتی در سطح نیروهای نظامی، به دلیل حجم بسیار زیاد، غالباً به مقبره‌های اطلاعات تبدیل می‌شوند و اغلب آنها را در بازه‌های زمانی کوتاه یا برای همیشه امحاء شده و یا بایگانی را که صورت می‌گیرد. علیرغم هزینه‌های سنگین در بخش تکنولوژی اطلاعات، بسیاری از تصمیم‌ها همچنان در فقر اطلاعاتی اتخاذ می‌گردند. از قابلیت‌های بالقوه اطلاعات ذخیره‌شده استفاده نمی‌شود. نیاز به تبدیل اطلاعات به دانش در بسیاری زمینه‌ها آشکار گردیده است. (Ye, 2003:12)

فنون داده‌کاوی

برخی از فنون رایج بکار گرفته‌شده تحت عنوان داده‌کاوی عبارتند از (Berry,2007,35):

- درختان تصمیم‌گیری
- قوانین وابستگی
- شبکه‌های عصبی
- الگوریتم ژنتیکی

- ابزارهای پرس‌وجو
- روش‌های آماری
- مصورسازی
- پردازش تحلیلی پیوسته
- یادگیری مبتنی بر مورد

کاربردهای داده‌کاوی

داده‌کاوی یک‌رشته جدید با کاربردهای وسیع و گوناگون است که به‌عنوان ده علم برتر که منجر به ایجاد تحول در عصر تکنولوژی می‌شود و در تمام زمینه‌ها کاربرد دارد، معرفی می‌شود. اصولاً هرجایی که داده وجود داشته باشد داده‌کاوی نیز معنا می‌یابد، از قبیل: امور تجاری و مالی، امور پزشکی، زیست پزشکی، تجزیه‌وتحلیل‌های مربوط به DNA، کشف ناهنجاری‌ها و اسناد جعلی، ارتباطات از راه دور، ورزش و سرگرمی، کتابداری و اطلاع‌رسانی. امروزه عملیات داده‌کاوی به‌صورت گسترده توسط تمامی شرکت‌هایی که مشتریان در کانون توجه آن‌ها قرار دارند، استفاده می‌شود، ازجمله: فروشگاه‌ها، شرکت‌های مالی، ارتباطاتی، بازاریابی و غیره. استفاده از داده‌کاوی به این شرکت‌ها کمک می‌کند تا ارتباط عوامل داخلی ازجمله: قیمت، محل قرارگیری محصولات و مهارت کارمندان را با عوامل خارجی ازجمله وضعیت اقتصادی، رقابت در بازار و محل جغرافیایی مشتریان کشف نمایند.

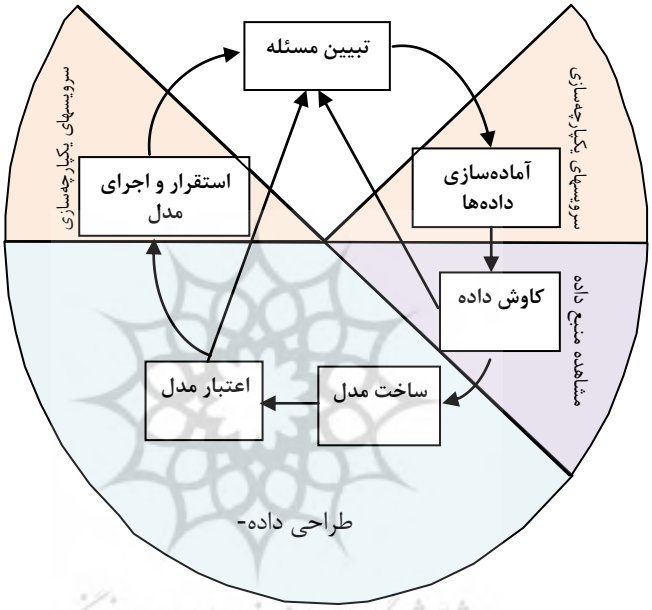
مراحل اصلی داده‌کاوی

استخراج دانش در پایگاه‌داده[1]، به‌عنوان روالی برای شناسایی الگوهای معتبر، جدید، بالقوه مفید، و سرانجام قابل‌فهم در داده‌ها، تعریف‌شده است. روال سراسری شامل تبدیل داده سطح پایین به دانش سطح بالاست. داده‌کاوی یا استخراج دانش داده‌ها دارای مراحل مختلفی می‌باشد که در اینجا به‌صورت خلاصه آن‌ها را بیان می‌کنیم:

1. Knowledge discovery in databases

مرحله درک پروژه و فهم حوزه کاربرد: اولین مرحله پردازش استاندارد می‌باشد که به‌صورت آشکار اهداف و نیازمندی‌ها آن مشخص می‌شود. ترجمه اهداف و محدودیت آن در قاعده سازی، تعریف مسئله داده‌کاوی و مهیاکردن استراتژی اولیه برای نائل شدن به اهداف تعریف می‌شود. (مطابق شکل ۲)

شکل ۲) مراحل داده‌کاوی



۱ -تشکیل انبار داده. استخراج اطلاعات از چندین منبع داده (پایگاه داده)

در این مرحله برای تشکیل محیطی پیوسته و یکپارچه جهت انجام مراحل بعدی و داده‌کاوی در آن، انجام می‌گیرد. در حالت کلی انبار داده مجموعه پیوسته و طبقه‌بندی‌شده است که دائماً در حال تغییر بوده و دینامیک است که برای کاوش آماده می‌شود. انبار داده‌ها، مجموعه‌ای است موضوعی، مجتمع، متغیر در زمان و پایدار از داده‌ها که به‌منظور پشتیبانی از فرآیند مدیریت تصمیم‌گیری مورداستفاده قرار می‌گیرد.

۲ -پاک‌سازی داده‌ها: از بین بردن نویز و ناسازگاری داده‌ها.

۳ -یکپارچه‌سازی اطلاعات و حذف داده‌های زائد: در این مرحله داده‌های غیر معتبر از مجموعه داده‌های آموزشی خارج می‌شوند. داده‌های دارای نویز، اطلاعات ناکامل و ... نمونه‌هایی از داده‌هایی هستند که باید پاک‌سازی در مورد آن‌ها انجام گردد.

انتخاب داده‌ها

داده‌های مرتبط به فرایند داده‌کاوی از سایر داده‌ها جدا می‌شود. این مبحث را می‌توان بخشی از فرایند کاهش اطلاعات نیز دانست. در مواردی نیاز به کاوش در تمام محتویات پایگاه نیست بلکه ممکن است به‌منظور کاهش هزینه عملیات، نمونه‌هایی از عناصر انتخاب و کاوش شوند. این چیزی نیست جز انتخاب یک مجموعه داده یا یک زیرمجموعه از متغیرها، با استفاده از تکنیک‌های رتبه‌بندی و انتخاب است. انتخاب داده‌ها (داده‌های مرتبط با آنالیز از پایگاه داده بازیابی می‌شوند) آماده کردن داده‌های اولیه خام به داده‌های نهایی، این دادها در کلیه مراحل بعدی استفاده می‌شود و از این نظر این مرحله تحلیل و تلاش بیشتری را می‌طلبد. انتخاب عناصر و شناسه‌های تحلیل‌شده را برای کاوش داده‌ها اختصاص می‌دهیم؛ و با تمیز کردن دادهای خام آن را برای ابزارهای مدل‌سازی آماده می کنیم.

درنهایت در این مرحله برای کم کردن هزینه‌های عملیات داده‌کاوی، داده‌هایی از پایگاه داده انتخاب می‌شوند که موردمطالعه هستند و هدف داده‌کاوی دادن نتایجی در مورد آن‌هاست.

تبدیل داده‌ها:

داده‌ها به فرمی مناسب برای داده‌کاوی مانند خلاصه‌سازی و همسان‌سازی، داده‌ها به قالبی قابل‌استفاده برای داده‌کاوی درمی‌آیند. از اعمالی که در این مرحله صورت می‌گیرد، می‌توان به خلاصه‌سازی و یا محاسبه مقادیر تجمعی اشاره کرد. تبدیلاتی ساده همچون تبدیل نوع داده‌ای به نوع دیگر تا تبدیلات پیچیده‌تر همچون تعریف صفات جدید با انجام عملیات‌های ریاضی و منطقی روی صفات موجود. این مرحله برای افزایش کیفیت داده بکار گرفته‌شده برای داده‌کاوی، لازم است. همچنین برای بهبود کارایی کاوش داده لازم است. پیش‌پردازش داده شامل پاک‌سازی داده، انتقال داده، یکپارچه‌سازی داده، کاهش یا فشرده‌سازی داده برای نمایش فشرده، و غیره است.

مشخص است برای انجام عملیات داده‌کاوی لزوماً باید تبدیلات خاصی روی‌داده‌ها انجام گیرد ممکن است این تبدیلات خیلی راحت و مختصر مثل تبدیل بایت به عدد صحیح باشد یا خیلی پیچیده و زمان‌بر و با هزینه‌های بالا مثل تعریف صفات جدید و یا تبدیل و استخراج داده‌ها از مقادیر رشته‌ای و ... باشد.

کاوش در داده‌ها:

انجام عملیات داده‌کاوی توسط نرم‌افزارهای مخصوص.در این مرحله است که داده‌کاوی انجام می‌شود. در این مرحله با استفاده از تکنیک‌های داده‌کاوی داده‌ها مورد کاوش قرارگرفته، دانش نهفته در آن‌ها استخراج‌شده و الگوسازی صورت می‌گیرد. فرایندی که در آن با استفاده از روش‌ها و تکنیک‌های خاص، استخراج الگوهای دانش صورت می‌گیرد. داده‌های تبدیل‌شده با استفاده از تکنیک‌ها و عملیات‌های داده‌کاوی مورد کاوش قرار می‌گیرند تا الگوهای موردنظر کشف شوند. شامل تفسیر الگوهای استخراج‌شده، و تا حد امکان، بصری سازی این الگوهاست. بصری سازی یک کمک‌کننده مهم در قابل‌فهم سازی الگوهاست.

ارزیابی الگوها:

تشخیص الگوهای صحیح موردنظر، از سایر الگوها در این مرحله انجام می‌شود. صحت الگوها بر اساس یک سری معیارهای جذابیت سنجیده می‌شود. در این مرحله خروجی‌های مرحله داده‌کاوی آزمایش می‌شوند تا مشخص شود که آیا این نتایج قابل‌استفاده و جالب هستند یا نه؟ اگر نتایج بهینه نباشد می‌توانیم فرآیند داده‌کاوی را با صفات و نمونه‌های جدید تکرار کنیم. همچنین می‌توانیم به «انبار داده» مراجعه کنیم و فرآیند استخراج دانش را تکرار کنیم.

۴ -تفسیر نتیجه: نمایش نتایج به‌صورت قابل‌فهم مانند گزارش و گراف.

در این مرحله نتایج و الگوهای ارائه‌شده توسط ابزار داده کاو موردبررسی قرارگرفته و نتایج مفید معین می‌شود. در این بخش به‌منظور ارائه دانش استخراج‌شده به کاربر، از یک سری ابزارهای بصری سازی استفاده می‌گردد. این مرحله شامل تلفیق این دانش با کارایی سیستم و گرفتن تصمیمات عملی بر اساس این دانش است.

طرز کار ابزار داده کاو این‌گونه است که ابزار به دنبال اثبات این است که وجود چیزی به معنای وجود چیز دیگری است و سعی می‌کند در درجه اول از توالی ارتباطات برای کشف یک الگو بهره بگیرد و درنهایت اطلاعات به‌دست‌آمده را دسته‌بندی کند تا به الگوی خاصی برسد که بتواند آن را بر اساس فاکتورهای داخلی به مخاطبش ارائه دهد.

همانطوریکه در شکل ۲ نشان داشده شده مراحل داده‌کاوی شامل ۶ گام مهم می باشد که عبارتند از: تبیین و تعریف مسئله[۱]، آماده سازی محیا شدن برای داده‌کاوی[۲]، کاوش داده-ها[۳]، مدلسازی داده‌کاوی[۴]، تعیین اعتبار مدل[۵] و استقرار و اجرای مدل[۶]
خدمات یکپارچه‌سازی[۷]، مشاهده منبع داده[۸] و طراح داده‌کاوی[۹]
البته از یک الگوریتم جامع‌تری نیز برای اجرایی پروژه‌های داده‌کاوی توسط پیشنهاد شده که در بخش تجزیه و تحلیل این مقاله از این مدل استفاده گردیده است.

جایگاه داده‌کاوی در میان علوم مختلف

ریشه‌های داده‌کاوی در میان سه خانواده از علوم، قابل‌پیگیری می‌باشد. مهم‌ترین این خانواده‌ها، آمار کلاسیک[۱۰] می‌باشد. بدون آمار، هیچ داده‌کاوی وجود نخواهد داشت، به‌طوری‌که آمار، اساس اغلب فنّاوری‌هایی می‌باشد که داده‌کاوی بر روی آن‌ها بنا می‌شود. آمار کلاسیک مفاهیمی مانند تحلیل رگرسیون، توزیع استاندارد، انحراف استاندارد، واریانس، تحلیل خوشه و فاصله‌های اطمینان را که همه این موارد برای مطالعه داده و ارتباط بین داده‌ها هست را در برمی‌گیرد. مطمئناً تحلیل آماری کلاسیک نقش اساسی در تکنیک‌های داده‌کاوی ایفا می‌کند.

دومین خانواده‌ای که داده‌کاوی به آن تعلق دارد هوش مصنوعی[۱۱] می‌باشد. هوش مصنوعی که بر پایه روش‌های ابتکاری می‌باشد و با آمار ضدیت دارد، تلاش دارد تا فرایندی مانند فکر انسان را برای حل مسائل آماری بکار بندد. چون این رویکرد نیاز به توان محاسباتی بالایی دارد تا اوایل دهه ۱۹۸۰ عملی نشد. هوش مصنوعی کاربردهای کمی را در حوزه‌های علمی و حکومتی پیدا کرد، اما نیاز به استفاده از کامپیوترهای بزرگ باعث شد همه افراد نتوانند از تکنیک‌های ارائه‌شده استفاده کنند.

---

1. Defining the problem
2. Preparing data
3. Exploring data
4. Building models
5. Validating models
6. Deploying and uploading models
7. Integration services
8. Data source view
9 . Data mining designer
10 . Classic Statistics
11. Artificial Intelligence

سومین خانواده داده‌کاوی، یادگیری ماشین[1] هست که به مفهوم دقیق‌تر، اجتماع آمار و هوش مصنوعی هست. درحالی‌که هوش مصنوعی نتوانست موفقیت تجاری کسب کند، یادگیری ماشین در بسیاری از موارد جایگزین آن گردید. از یادگیری ماشین به‌عنوان تحول هوش مصنوعی یاد شد، چون مخلوطی از روش‌های ابتکاری هوش مصنوعی به همراه تحلیل آماری پیشرفته هست. یادگیری ماشین اجازه می‌دهد تا برنامه‌های کامپیوتری در مورد داده‌ای که آن‌ها مطالعه می‌کنند، مانند برنامه‌هایی که تصمیم‌های متفاوتی بر مبنای کیفیت داده مطالعه شده می‌گیرند، یادگیری داشته باشند و برای مفاهیم پایه‌ای آن از آمار استفاده می‌کنند و الگوریتم‌ها و روش‌های ابتکاری هوش مصنوعی را برای رسیدن به هدف بهره می‌گیرند.

چطور می‌توان از قدرت و اهرم داده‌کاوی برای حل مشکلات روزانه‌کاری و اتخاذ تصمیم‌های مناسب تر استفاده کرد؟ فنّاوری پشت پرده داده‌کاوی چیست؟ چرخه یک پروژه داده‌کاوی چیست؟ در این فصل به تمام این سوا لات پاسخ داده خواهد شد و داده‌کاوی به‌طور کامل معرفی خواهد شد.

معرفی زمینه و نمونه‌های اجراشده پروژه داده‌کاوی

داده‌کاوی پیش‌بینی وضع آینده، گرایش‌ها و شناخت سلیقه‌های عمومی آن‌ها را برای سازمان‌ها ممکن می‌سازد که از جمله مهم‌ترین حوزه‌هایی که اطلاعات داده کاوی می‌تواند در اتخاذ تصمیم‌گیری مناسب کمک نماید عبارتند از :

- بازارهای هدف با توجه به آخرین وضعیت‌های جهتی و جغرافیایی و اقتصادی و ....
- پیدا کردن الگوی فعالیت در حوزه حریف و دشمن (برای نیروهای امنیتی و نظامی)
- برنامه‌ریزی برای معرفی محصول جدید
- دسته‌بندی اهداف بر اساس ماهیت و نوع آن‌ها در عملیات‌های هوایی و دریایی
- آنالیز نیازهای تجهیزاتی نیروهای دشمن و خودی
- تشخیص تجهیزات مناسب برای دسته بندی نیروهای خودی
- تشخیص فاکتورهایی برای جذب نیروهای انسانی کیفی.
- تجزیه‌وتحلیل سبد فعالیت اطلاعاتی، تجهیزاتی و آماد در پشتیبانی دشمن
- پیشگویی میزان و پیش‌بینی الگوهای نفوذ نیروهای دشمن

1. Machine Learning

- تعیین میزان استفاده از سامانه‌های مختلف ارتباطی (بی سیم و با سیم) توسط نیروهای دشمن
- شناسایی فاکتورهای اصلی در ریسک اقدامات پدافندی و یا آفندی
- پیش‌بینی میزان خسارت احتمالی برای گروه‌های مختلف در بحران‌های طبیعی توسط نیروهای مدیریت بحران.
- بررسی میزان تأثیر دارو بر بیماری و اثرات جانبی آن
- تشخیص و پیش‌بینی انواع بیماری‌ها مانند تشخیص و یا پیش‌بینی انواع سرطان
- تجزیه‌وتحلیل داده‌های موجود در سیستم‌های اطلاعات سلامت
- تحلیل عکس‌های پزشکی

مثال داده‌کاری اجرا شده در صنعت شرکت فولادسازی پوهانگ کره برای صرفه‌جویی در مصرف انرژی در کوره‌های بلند خود از الگوریتم‌های داده‌کاوی استفاده و در حدود ۱۵٪ از مصرف انرژی خود را کاهش داد، که باعث ۱/۳ میلیون دلار صرفه‌جویی در هزینه‌های شرکت شده و از طرف دیگر باعث کاهش قیمت محصولات آن شرکت و افزایش تعداد مشتریان و سود بیشتر آن شرکت نیز شد. یکی از هتل‌های مشهور در لاس‌وگاس آمریکا، برای بالا بردن رضایت مسافران از الگوریتم‌های داده‌کاوی استفاده کرد، به این صورت که با استفاده از اطلاعات جمع‌آوری‌شده از مسافران به‌وسیله پرسشنامه، و آنالیز آن داده‌ها توانست عواملی که باعث می‌شد مسافران دوباره به این هتل بازگردند را پیداکرده و با طبقه‌بندی مسافران، مسافران وفادار به هتل را پیدا کنند.

- مثال داده کاوی اجرا شده در امور بازرگانی و خدماتی

دریکی از بانک‌های بزرگ کانادا با استفاده از الگوریتم‌های داده‌کاوی، مدلی را برای داده‌ها ارائه داده و به‌وسیله نتایج آنالیز آن، مسئله مهم تقلب در حساب‌ها و چگونگی و میزان برگشت وام‌های داده‌شده توسط بانک را حل نمودند و تصمیمی صحیح را برای مشتری‌های جدید بانک گرفتند.

یکی از نمونه های کاربردی داده‌کاوی در حوزه نظامی استفاده از این ابزار توسط نیروهای ویژه آمریکایی[1] در افغانستان است که توسط پهبادهای آمریکایی بصورت مستمر از سال

1. US Special Operations Command Contracting: Data-Mining the Public Record

۲۰۰۳ صورت می‌گیرد. در این پروژه داده‌کاوی نیروهای آمریکایی نه تنها اطلاعات مربوط به نیروهای طالبان را دائما ذخیره و داده کاوی نمودند بلکه اطلاعات وسیعی از کشورهای منطقه شامل چین، روسیه، ایران و هند و پاکستان و سایر کشورهای منطقه را با استفاده از پهبادهای اسکن‌ایگل جمع‌آوری نمودند و با استفاده از تکنیکهای داده‌کاوی دانش‌های پنهانی را استخراج و در حال و آینده مورد بهره‌برداری قرار می دهند. (شیرزاد، ۱۳۹۱: ۲۰۵)

(http://defenceindustrydaily.com/images/AIR_UAV_ScanEagle)

- مثال در نیروی دریایی سوئد: به عنوان یک مثال کاربردی دیگر در سطح نیروهای مسلح می توان به جمع‌آوری اطلاعات توسط نیروی دریایی سوئد می باشد، در این پروژه داده کاوی نیروی دریایی سوند حرکت زیردریایی‌های کشورهای دیگر در منطقه را رصد و اطلاعات حرکت آنها را در رایانه مربوطه ذخیره‌سازی نمودند، با استفاده از تحلیل الگوریتم ژنتیک توانستند، احتمالات مختلف حرکت زیردریاییها را مدل و با سطح اطمینان بالایی مسیرهای حرکت بعدی و اقدامات آتی آنها را مشخص نمایند. (Bergsten, 1997:130)

داده کاوی اجرا شده در استراتژی‌های داده‌کاوی.

۱)دسته‌بندی[1]

فرایند گروه گروه کردن مجموعه‌ای از اشیاء فیزیکی یا مجرد به‌صورت طبقه‌هایی از اشیاء مشابه هم را دسته‌بندی می‌گویند. دسته‌بندی به تخصیص موارد به گروه‌ها بر اساس یک خصوصیت قابل پیش‌بینی، برمی‌گردد. هر مورد شامل مجموعه‌ای از خصوصیات است که یکی از آن‌ها خصوصیت کلاس (خصوصیت قابل پیش‌بینی) نام دارد. دسته‌بندی شامل یافتن مدلی است که خصوصیت کلاس را بر اساس تابعی از خصوصیات ورودی تشریح کند. برای آموزش مدل دسته‌بندی، شما به مقدار کلاس موارد ورودی در مجموعه داده‌های آموزشی نیاز دارید. الگوریتم‌های داده‌کاوی که به یک هدف برای یادگیری نیاز دارند، الگوریتم‌های نظارت‌شده[2]، نام دارند. در مسائل دسته‌بندی هدف شناسایی ویژگی‌هایی است که گروهی را که هر مورد به آن تعلق دارد نشان دهند. از این الگو می‌توان هم برای فهم داده‌های موجود و هم پیش‌بینی نحوه رفتار موارد جدید استفاده کرد. (Jeffrey, 2014)

1. Classification
2. supervised algorithm

دسته‌بندی یکی از رایج‌ترین وظایف داده‌کاوی است. مسائل تجاری‌ای مانند تحلیل رویگردانی[۱]، مدیریت ریسک و هدف‌گیری موردی[۲]، که قبلاً در موردشان توضیح دادیم، شامل دسته‌بندی می‌شوند. در این وظیفه تعداد دسته‌ها یا کلاس‌های دسته‌بندی از قبل مشخص می‌باشد. تولید یک درخت تصمیم یا مجموعه‌ای از قوانین دسته‌بندی، که برای فهم بهتر داده‌های موجود در پایگاه داده و همچنین دسته‌بندی داده‌هایی که در آینده به پایگاه داده اضافه می‌شوند به کار می‌رود. داده‌کاوی مدل‌های دسته‌بندی را با بررسی داده‌های دسته‌بندی‌شده قبلی ایجاد می‌کند و یک الگوی پیش‌بینی کننده را به‌صورت استقرایی می‌یابد.

برخی از کاربردهای طبقه‌بندی شامل بازاریابی مستقیم، شناخت تقلب، شناسایی و طبقه‌بندی مشتری و ارسال کاتالوگ می‌باشد. الگوریتم‌های رایج دسته‌بندی شامل درخت تصمیم، شبکه عصبی و بیز[۳] است.

۲) خوشه‌بندی[۴]

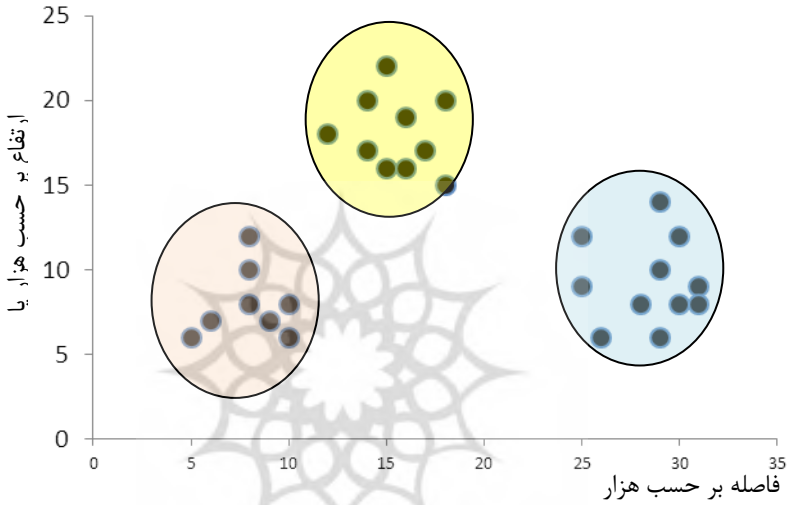
خوشه‌بندی را گاهی قطعه‌بندی نیز می‌خوانند که برای شناسایی گروه‌بندی طبیعی موارد بر اساس یک مجموعه خصوصیت بکار می‌رود. موارد درون یک گروه بیشترین یا کم‌ترین میزان شباهت را در مقادیر خصوصیات دارند. برخلاف دسته‌بندی تعداد کلاس‌ها در ابتدا مشخص نیست. خوشه‌بندی داده‌ها بر اساس اصل مفهومی حداکثرسازی شباهت‌های بین اعضای هر کلاس و حداقل‌سازی شباهت‌ها بین اعضای مربوط به کلاس‌های مختلف صورت می‌گیرد.

شکل ۳ یک مجموعه داده ساده دریافتی از هدف‌های کشف شده راداری که شامل دو خصوصیت ارتفاع و فاصله است را نشان می‌دهد. الگوریتم خوشه‌بندی مجموعه داده را بر اساس این دو خصوصیت به سه گروه تقسیم کرده است. خوشه اول شامل هدف‌های کشف شده با سرعت و ارتفاع پایین، خوشه دوم شامل ارتفاع متوسط سرعت بالا، و خوشه سوم شامل اهداف سرعت بسیار بالا ولی ارتفاع نسبتاً پایین است. متفاوت نگریسته شود خوشه-

1. Churn analysis
2. Ad targeting
3. Naïve Bayes
4. Clustering

بندی این اهداف در تخصیص و واگذاری این اهداف توسط سامانه‌های تصمیم‌سازی و تصمیم‌گیری فرماندهی و کنترل سرعت تصمیم‌گیری را افزایش می‌دهند.

خوشه‌بندی یک وظیفه داده‌کاوی بدون نظارت[1] است. هیچ خصوصیت تکی برای هدایت روال آموزش بکار نمی‌رود و تمام خصوصیات ورودی به‌طور یکسان هدف هستند. بیشتر الگوریتم‌های خوشه‌بندی، مدل را در تعداد تکرارهای معینی می‌سازند و هنگامی‌که مدل همگرا شد، متوقف می‌شوند و آن هنگامی است که مرزهای گروه‌ها تثبیت شد.



شکل ۳- خوشه‌بندی

کاربردهای خوشه‌بندی:

ازآنجاکه خوشه‌بندی یک روش یادگیری بدون نظارت محسوب می‌گردد، در موارد بسیاری می‌تواند کاربرد داشته باشد.

- در بازاریابی: دسته‌بندی مشتری‌ها به دسته‌هایی برحسب رفتارها و نیازهای آن‌ها از طریق مجموعه زیادی از ویژگی‌ها و آخرین خریده‌ای آن‌ها.
- زیست‌شناسی: دسته‌بندی حیوانات و گیاهان از روی ویژگی‌های آن‌ها.
- کتابداری : دسته‌بندی کتاب‌ها.
- نقشه‌برداری شهری: دسته‌بندی خانه‌ها بر اساس نوع و موقعیت جغرافیایی آن‌ها.
- مطالعات زلزله‌نگاری: تشخیص مناطق حادثه‌خیز بر اساس مشاهدات قبلی.
- وب: دسته‌بندی اسناد و یا دسته‌بندی مشتریان به سایت‌ها و...

1.Unsupervised algorithm

- تشخیص گفتار: در ساخت کتاب کد از بردارهای ویژگی، در تقسیم کردن گفتار برحسب گویندگان آن و یا فشرده‌سازی گفتار.
- تقسیم‌بندی تصاویر: تقسیم‌بندی تصاویر پزشکی و یا ماهواره‌ای.
- بیمه: تشخیص افراد متقلب، تشخیص افرادی که بیمه موتور دارند و بیشترین میزان درخواست از بیمه را نیز در سال مشخصی داشته‌اند.

یک نمونه کاربردی خوشه‌بندی در حوزه نظامی در شکل ۳ تشریح گردیده است.

۳) تخمین.

تخمین با نتایجی که با ارقام پیوسته نشان داده‌شده‌اند سروکار دارد. در تخمین دسته‌بندی متغیرهای ناشناس مداوم استفاده می‌شود.

مثال‌هایی دیگر از تخمین:

- تخمین تعداد قطعات بحرانی در مسائل آمادی
- تخمین عمر مفید تجهیزات
- تخمین تعداد نفرات شرکت کننده در یک مانور
- تخمین تعداد تلفات احتمالی در یک نبرد

بطور کلی مدل‌های رگرسیون و شبکه‌های عصبی برای تخمین مناسبند.

۴) وابستگی[1]

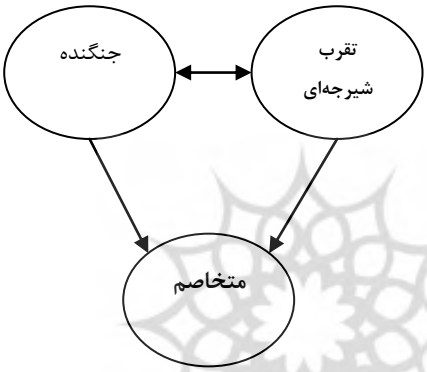
وابستگی یکی دیگر از وظایف معمول داده‌کاوی است که تحلیل سبد بازار[2] نیز خوانده می‌شود. یک مسئله کسب‌وکار معمول وابستگی، تحلیل یک جدول تراکنش فروش و شناسایی آن محصولاتی است که اغلب در یک سبد خرید یکسان قرار می‌گیرند. استفاده رایج از وابستگی شناسایی مجموعه رایج اقلام خرید (مجموعه اقلام متناوب) و قوانینی برای فروش- متقاطع[3] است. البته در این زمینه نیز بسیاری از مسائل نظامی می‌تواند تحلیل و مورد بهره‌برداری قرار گیرد. بطور مثال درک وابستگی بین پدیده‌های مختلف و ارتباط بین آنها در امور نظامی در بسیاری از موارد کاملا شفاف و مشخص نیست ولی ترسیم نمودار وابستگی و تحلیل آن می‌تواند به تجزیه و تحلیل کمک کند. در شکل ۴ یک مثال کاربردی از وابستگی بین فعالیت‌ها و متغیرها نشان داده شده است.

---

1 . Association
2 . Market basket analysis
3 . Cross-selling

بیشتر الگوریتم‌های وابستگی برای یافتن مجموعه اقلام متناوب، مجموعه داده‌ها را چندین بار پیمایش می‌کنند. آستانه تناوب[1] (پشتیبانی)، قبل از شروع پردازش، توسط کاربر تعیین می‌شود. برای مثال پشتیبانی برابر با ۲ درصد به معنی است که مدل فقط آن اقلامی که در ۲ درصد از کارت‌های خرید دیده می‌شوند را تحلیل کند. هر مجموعه قلم یک اندازه دارد که تعداد آیتم‌های آن است. مثلاً اندازه این مجموعه آیتم برابر با ۳ است.

جدای از شناسایی مجموعه اقلام بر مبنای پشتیبانی، بیشتر الگوریتم‌های وابستگی، قوانین را نیز می‌یابند. یک قانون وابستگی شکلی مانند A,B=>C با یک احتمال دارد، که A و B و



C همگی مجموعه اقلام هستند. همچنین احتمال در ادبیات داده‌کاوی اعتماد[2] نیز خوانده می‌شود. احتمال یک مقدار آستانه است که یک کاربر قبل از آموزش یک مدل وابستگی، نیاز دارد که تعیین کند. برای مثال، این یک قانون معمول است: هدف= «جنگنده» و هدف= «حرکت شیرجه‌ای» => محصول= «متخاصم» با احتمال ۸۰٪. تفسیر این قانون بسیار سرراست است. اگر یک هدف جنگنده و بصورت شیرجه‌ای به سمت هدف‌های خاص تقرب نماید با احتمال ۸۰ درصد ممکن است که متخاصم باشد. (Han, 2006)

۵) رگرسیون[3]

وظیفه رگرسیون مشابه با دسته‌بندی است. تفاوت اساسی در خصوصیت پیش‌بینی است که یک عدد پیوسته است. تکنیک رگرسیون سال‌ها در حوزه آمار مطالعه شده است. رگرسیون خطی و منطقی از روش‌های بسیار رایج رگرسیون هستند. سایر تکنیک‌های رگرسیون شامل درخت‌های رگرسیون و شبکه‌های عصبی است. (Schubert, 1996)

---

1 . Frequency threshold
2 . Confidence
3 .Regression

تحلیل رگرسیون می‌تواند بسیاری از مسائل سازمان‌ها را حل کند. برای مثال، آنها می‌توانند در پیش‌بینی دما، رطوبت، فشار هوا و پارامترهای جوی استفاده شود، همین طور در مسئله نظامی زیادی می توان با داشتن اطلاعات آماری زیاد از وضعیت سامانه‌ها و نفرات خودی و دشمن یک برازش آماری استخراج نمود، در صورتیکه تغییرات در طول زمان بصورت خطی باشد می توان یک معادله خطی برای آن لحاظ نمود. که این معادله خطی بر حسب رگرسیون داده‌ها بدست می‌آید.

۶) پیش‌بینی[1]

پیشگویی و پیش‌بینی یک وظیفه مهم دیگر برای داده‌کاوی است. وضعیت انظباطی بعد از اعماب قوانین و محدودیت‌های جدید سازمان چگونه خواهد بود؟ مقدار پایش باران در ماه آینده چگونه خواهد بود؟ پیشگویی می‌تواند به این سؤالات پاسخ دهد.

در عمل پیش‌بینی، تنها روش برای بررسی صحت مدل، دیدن آینده و مقایسه نتیجه مدل و پدیده واقع‌شده می‌باشد. هر یک از تکنیک‌های استفاده‌شده در دسته‌بندی و تخمین را می‌توان برای استفاده در پیش‌بینی تطبیق داد. از داده‌های پیشین برای تهیه یک مدل که بیانگر رفتار مشاهده کنونی است استفاده می‌شود. وقتی این مدل برای ورودی‌های کنونی به کار رفت نتیجه کار پیش‌بینی رفتار آینده خواهد بود.

۷) تحلیل توالی[2]

تحلیل توالی برای یافتن الگوها در رشته‌های گسسته بکار می‌رود. یک توالی از مقادیر (یا نواحی) گسسته ترکیب‌شده است. برای مثال، یک توالی DNA یک‌رشته مرکب از چهار ناحیه مختلف A، G، C و T است. تفاوت داده‌های توالی و رشته‌های زمانی در این است که رشته‌های زمانی شامل اعداد پیوسته و توالی شامل مقادیر گسسته است.

۸) تحلیل انحراف[3]

تحلیل انحراف برای یافتن موارد نادری است که با دیگران بسیار متفاوت هستند. این‌که همچنین کشف خطا[4] نیز نامیده می‌شود که تشخیص تغییرات معنی‌دار از رفتار

---

1. Forecasting
2. Sequence analysis
3. Deviation Analysis
4. Outlier detection

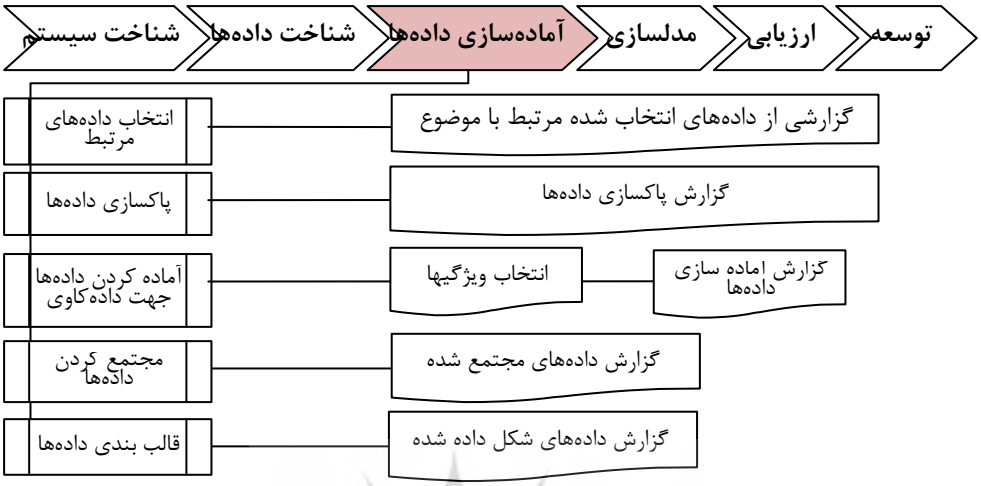
مشاهده‌شده قبلی است. تحلیل انحراف می‌تواند در کاربردهای بسیاری استفاده شود. رایج‌ترین استفاده از آن در کشف تقلبات کارت‌های اعتباری است. شناسایی موارد غیرعادی از بین میلیون‌ها تراکنش یک چالش بسیار مشکل است. سایر کاربردها می‌تواند شامل کشف ورود سرزده به شبکه، تحلیل خطای تولید و غیره باشد.

هیچ استانداردی برای تحلیل انحراف وجود ندارد و این هنوز یک موضوع باز برای تحقیقات است. اما تحلیلگران معمولاً نسخه‌های اصلاح‌شده الگوریتم‌های درخت‌های تصمیم، خوشه‌بندی، یا شبکه‌های عصبی را برای این کار بکار می‌گیرند. به‌منظور ساختن قوانین معنی‌دار، تحلیل‌گران نیاز به داشتن نمونه‌های بسیار از این موارد نادر در مجموعه‌های آموزشی خود دارند.

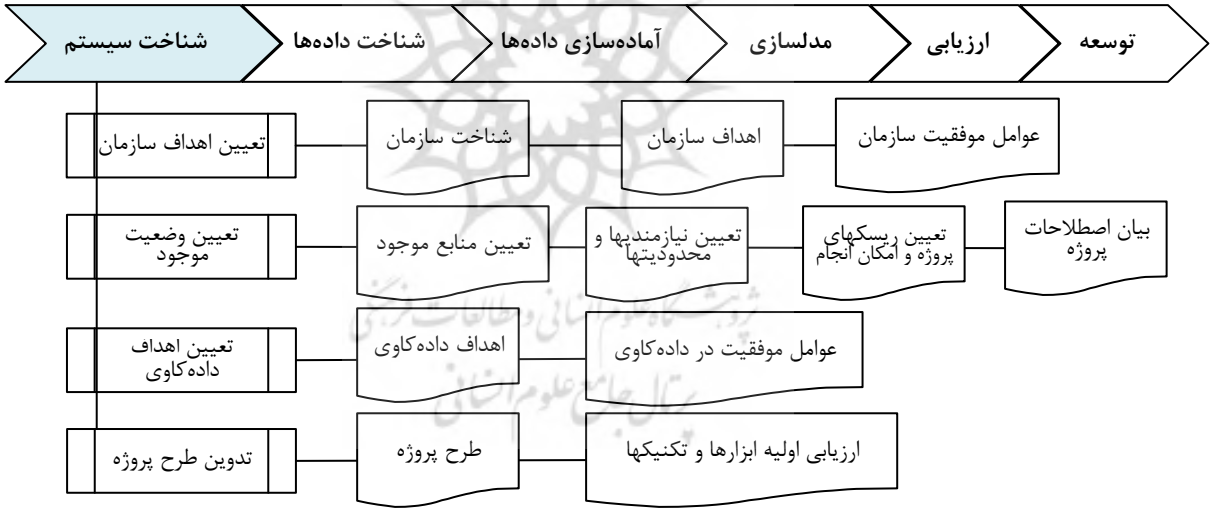
جدول ۱) کاربردهای داده‌کاوی

وظیفه / فعالیت	الگوریتم
پیشگویی یک خاصیت گسسته برای مثال پیش‌بینی اینکه آیا دریافت‌کنندگان یک نامه پیام خاص در جهت اجرای اهداف آن پیام اقدام خواهند نمود. مثلا ارائه تبلیغات یا پیام بازرگانی در رسانه‌ها یا انتشار پیام به منظور عملیات روانی بر روی نیروهای دشمن و ...	الگوریتم درخت تصمیم الگوریتم بیز الگوریتم خوشه‌بندی الگوریتم شبکه عصبی
پیشگویی یک خاصیت پیوسته برای مثال پیشگویی فروش سال آینده	الگوریتم درخت تصمیم الگوریتم سری زمانی
پیشگویی یک ترتیب برای مثال تحلیل نحوه ترتیب انتخاب صفحه در وب‌سایت یک شرکت	الگوریتم خوشه‌بندی ترتیبی
پیدا کردن گروهی از آیتم‌های معمولی در یک تراکنش برای مثال استفاده از تحلیل سبد خرید برای ارائه پیشنهاد خرید یک کالای اضافه به مشتری	الگوریتم وابستگی الگوریتم درخت تصمیم
پیدا کردن گروهی از گزینه‌های مشابه برای مثال تقسیم جمعیتی داده‌ها به گروه‌ها برای درک بهتر روابط بین خاصیت‌ها	الگوریتم خوشه‌بندی الگوریتم خوشه‌بندی ترتیبی

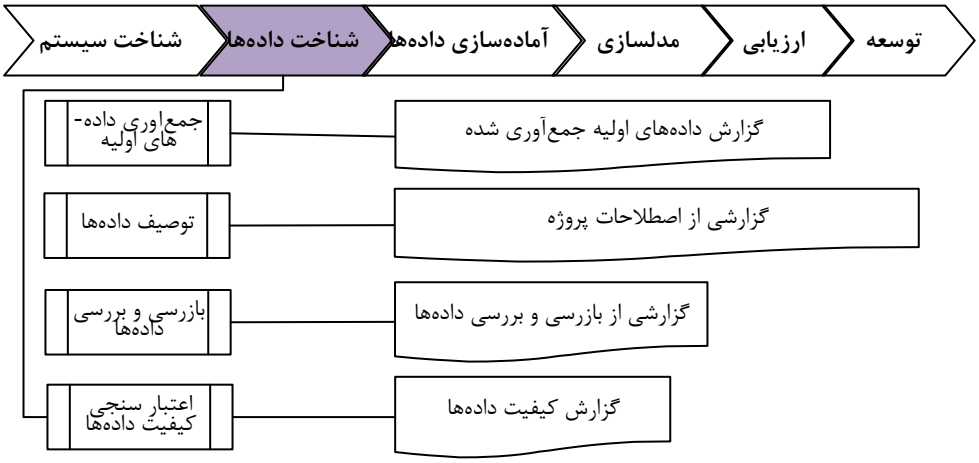
الگوی پیشنهادی مراحل پیاده سازی و اجرای داده کاوی در سطح نیروهای مسلح:



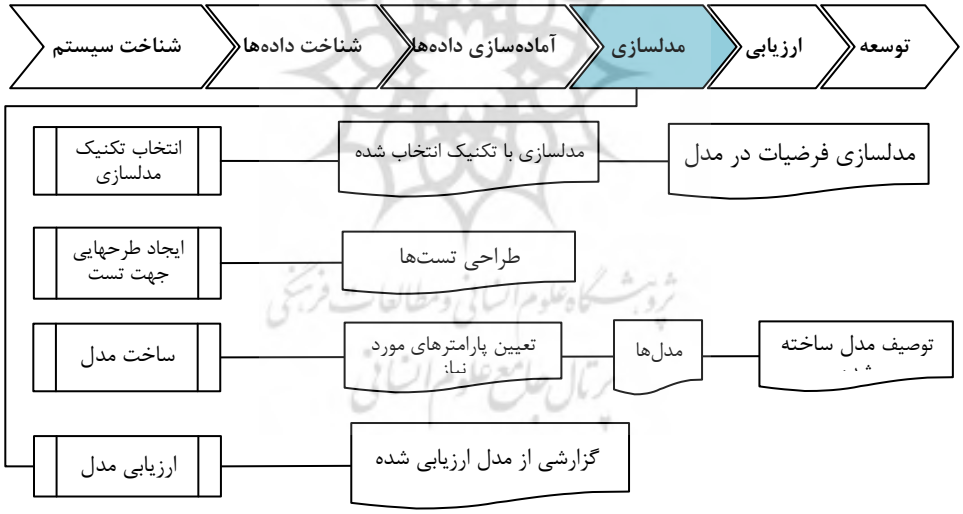
شکل ۵- گام اول) شناخت سیستم:



شکل ۶- گام دوم: شناخت داده‌ها

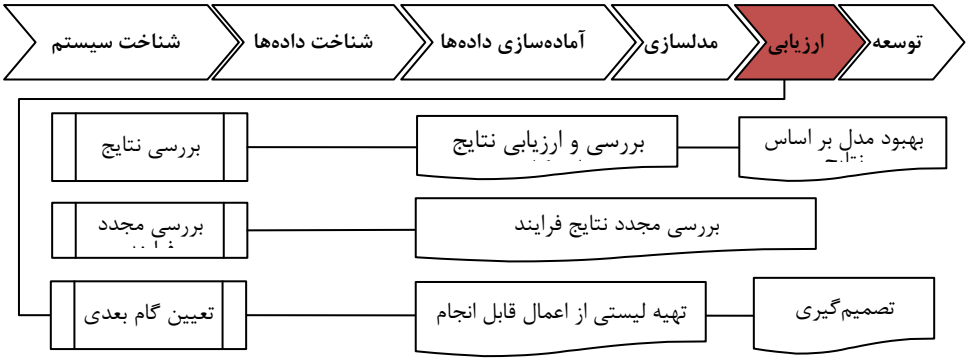


شکل ۷- گام سوم: آماده‌سازی داده‌ها



شکل ۸- گام چهارم: مدلسازی

شکل ۹- گام پنجم: ارزیابی



شکل ۱۰- گام ششم: توسعه

نتیجه‌گیری:

همانطوریکه در بخش چکیده، مقدمه و متن مقاله اشاره گردید با افزایش حجم استفاده از رایانه‌ها در سازمان‌ها و بخصوص بخش‌های نظامی و دفاعی کشورها، حجم زیادی از داده‌ها و اطلاعات بصورت مستمر رصد، جمع‌آوری و پایش و ذخیره می‌گردند، این گونه اطلاعات شامل اطلاعات کارکنان، تجهیزات، سامانه‌های مختلف و حتی عملکرد و رفتار کشورهای حریف و دشمن می‌باشد، اینگونه اطلاعات بطور برنامه‌ریزی شده عمدتا بعد از یک بازه زمانی مشخص به بایگانی را ارسال و یا امحا می‌گردند، در حالیکه تحقیقات نشان داده است ارزش اینگونه اطلاعات برای تصمیم‌سازی و تصمیم‌گیری غیر قابل محاسبه است. در این مقاله با ذکر متال‌های کاربردی علم داده‌کاوی و کشف دانش در حوزه‌ای مختلف بویژه نظامی مشخص گردید که در چه زمینه‌هایی می‌توان از داده‌کاوی بهره گرفت و در صورت استفاده از داده‌کاوی با استفاده از تکنیک‌های دسته‌بندی، خوشه‌بندی، تخمین، وابستگی، رگرسیون، پیش‌بینی، تحلیل توالی و تحلیل انحراف تصمیم‌هایی اتخاذ شده بشدت کاراتر، موثرتر و مفید تر خواهد بود، حتی در بعضی از مواقع داده‌کاوی از داده‌های ذخیره شده به مدت طولانی منجر به استخراج دانش جدیدی می‌گردد که می‌تواند برای سازمان‌های نظامی که در صحنه‌های اجرایی و عملیاتی و نبرد با کمبود نوآوری و خلاقیت مواجه می-باشند کمک نماید.

در انتهای این تحقیق شکل‌های ۵ الی ۱۰ مراحل و گام‌های پیاده سازی پروژه داده‌کاوی در بخش‌هایی از نیروی نظامی، دفاعی، امنیتی که دارای پایگاه داده‌های چند و چندین ساله می‌باشند پیشنهاد شده است.

منابع:

- شیرزاد مهران، صلواتی زینب، هاشمی سید محمد (۱۳۹۱)، داده‌کاوی؛ ابزاری نیرومند در سیستم مدیریت دانش نیروهای مسلح، پایگاه علمی- پژوهشی پارس مدیر، صص ۲۱۱-۱۹۹
- صادقی ایرج، (۱۳۷۷) تجزیه و تحلیل و طاحی سیستم‌ها، نشر آتیه، تهران
- علیزاده، سمیه و ملک‌محمدی سمیرا (۱۳۹۰)، داده کاوی و کشف دانش گام به گام با نرم افزار Clementine، انتشارات دانشگاه خواجه نصیرالدین طوسی، چاپ اول، تهران
- غضنفری مهدی، سمیه علیزاده، بابک تیمورپور، (۱۳۸۷)، داده‌کاوی و کشف دانش، انتشارات دانشگاه علم و صنعت، تهران
- غضنفری، مهدی، رضایی، (۱۳۸۵)، مجموعه‌های فازی، انتشارات علم و صنعت، تهران
- ملک محمدی، سمیرا (۱۳۸۵)، کاربرد داده‌کاوی در مدیریت ارتباط با مشتری، سمینار کارشناسی ارشد دانشگاه علم و صنعت، تهران
- ____، بی‌تا، کاربرد داده‌کاوی در شناسایی و کشف الگوهای پنهان جرم سرقت
- _____، بی‌تا، بررسی کاربردهای داده کاوی در علوم نظامی و مباحث مرزبانی،
- داده‌کاوی ابزاری نیرومند در سیستم مدیریت دانش نیروهای مسلح

retrieved in http://textmine.persianblog.ir/post/108(2015/12/16)

- قراردادهای فرماندهی عملیات ویژه ایالات متحده: داده‌کاوی از اسناد عمومی

retrieved in http://www.asipress.ir/vdcbw5bz.rhb90piuur.html (2015/12/16

- Bergsten Ulla, Schubert Johan, Svensson Per (1997), Applying Data Mining and Machine Learning Techniques to submarine Intelligence analysis, 3rd conference of knowledge discovery and data mining , Menlo park, pp. 127-130
- Berry, M., and Linoff, G.(2007), Data Mining Techniques: For Marketing, Sales, and Customer Support" New York: John Wiley and Sons.
- Berson A., Smith S., and Thearling K.(2004), "Building Data Mining Applications for CRM," Tata McGraw-Hill, New York.
- Gyordi Cornelia, Gyordi Robert, Holban Stefan (2004), A Comparative Study of Association Rules Mining Algorithms", SACI 2004, 1st Romanian-Hungarian Joint Symposium on Applied Computational Intelligence, Timisoara, Romania, May 25-26, pp. 213-222, 2004.

- Fayyad U., Piatetsky-Shapiro G., and Smyth P., "From Data Mining to Knowledge Discovery in Databases," American Association for Artificial Intelligence, 1996.
- Gertisio Christine and Alan Dussauchoy(2004) , Knowledge Discovery from Industrial Data base, Journal of Intelligent Manufacturing, 15, pp. 29-37.
- Han, J, Kamber. M. (2006), Chapter 2: Data Processing, data mining concepts and techniques, 2nd edition, Morgan Kaufmann Publishers.
- Han, J, Kamber. M. (2006), Chapter 5: Mining Frequent Patterns, Associations and Correlations, data mining concepts and techniques, 2nd edition, Morgan Kaufmann Publishers.
- Han, J, Kamber. M. (2006), Chapter 7: Cluster Analysis, data mining concepts and techniques, 2nd edition, Morgan Kaufmann Publishers.
- Jeffery W. Seifert (2004), Analyst in information science and Technology Policy, Data Mining: An Overview, December 2004.
- Jeffrey W. Seifert (2014) ,Data Mining: An Overview, CRS Report for Congress, December 16.
- Schubert Johan (1996), specifying nonspecific evidence, International Journal of Intelligent system, pp. 525-563.
- Ye N. (2003), The handbook of Data mining, Lawrence Erlbaum Associations, Publishers Mahwah, New Jersey London.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی