

نگاهی بر مفاهیم و کاربردهای «داده‌کاوی»

انفجار اطلاعات بی معنا است

سعید نوری آزاد

می‌کند. در کل، باید بدانیم وقتی همه چیز برایمان معنا داشته باشد، دچار سردرگمی میان آن‌ها نخواهیم شد و در نتیجه انفجار اطلاعات نخواهیم داشت.

تعریف

داده‌کاوی که برخی آن را کشف دانش در پایگاه داده‌ها^۲ و نیز علم «معنابخشی داده» تفسیر کرده‌اند، خلاف تصور بسیاری، علم تازه و جدیدی نیست چراکه در مکان‌های مختلف، از این علم به‌اشکال گوناگون استفاده شده است که به‌عنوان نمونه می‌توان به داده‌کاوی حدود ۲۵ سال پیش شرکت بنزآلمان اشاره کرد که با جمع‌آوری اطلاعات از نمایندگی‌ها و تعمیرگاه‌های مختلف در نقاط مختلف آب و هوایی سراسر دنیا، بانک اطلاعاتی

اگر اهل مطالعه باشید حتما واژگانی مانند عصر کشاورزی، عصر صنعتی، عصر فراصنعتی و عصر اطلاعات را شنیده‌اید. در نوعی از تقسیم‌بندی تاریخی از نظر نوع زندگی مردمان، اعصار به دسته‌هایی مانند عصر کشاورزی، عصر صنعتی و عصر الکترونیکی تقسیم می‌شود و برخی عقیده دارند عصر حاضر که برآمده از عصر الکترونیک است، عصر اطلاعات نام دارد که در آن اطلاعات حرف اول را می‌زند و نه چیز دیگر.

در مورد این نظریه حرف بسیار است، اما کلیت آن را می‌توان پذیرفت، آن‌هم وقتی فکر کنید چه رخ خواهد داد وقتی یک روز حتی یک ساعت خطوط انتقال اطلاعات از کار بایستند.

برخی می‌گویند به دلیل افزایش سرعت تولید و انتقال اطلاعات و رشد تصاعدی حجم اطلاعات، به‌زودی پدیده‌ای رخ خواهد داد به نام انفجار اطلاعات که در آن انسان‌ها دچار سردرگمی و اختلال در زندگی می‌شوند.



تصویر: assetssearchpros.com

کاملی را در اختیار واحد تحقیق و توسعه خود قرار داد و بر اساس آن توانست تا امروز بهترین و پایدارترین تولیدکننده خودرو باشد. از کارکردها و کاربردهای داده‌کاوی در ادامه بیشتر خواهیم گفت.

در نگاه نخست می‌توان گفت، پیشرفت کشورها

بسیاری در نگاه اول داده‌کاوی را

همان تحلیل آماری می‌دانند

اما باید گفت که داده‌کاوی چیزی

فراتر از آمار است

در زمینه‌های مختلف و نگاه ویژه دولت‌ها به مقوله الکترونیکی شدن و یکپارچگی سیستم‌ها که نتیجه نفوذ استفاده از سیستم‌های رایانه‌ای در صنعت و ایجاد بانک‌های اطلاعاتی بزرگ توسط ادارات مختلف و نیز استفاده روزافزون بانک‌ها و بخش خصوصی است، نیاز به استفاده از این علم را هر روز

سفارش داده شود و این میزها دقیقاً در کنار تلویزیون‌ها چیده شوند تا مشتری به سهولت و با امکان انتخاب بهتر، خرید بیشتری انجام دهد. این روش می‌تواند سود بهتری برای مالک فروشگاه داشته باشد.

با توجه به این نمونه‌ها به این نتیجه می‌رسیم که این نوع بررسی‌ها می‌تواند در تهیه مواد اولیه نیز تاثیر بگذارد و این یعنی یکپارچگی سازمان^۵. کاربرد داده‌کاوی در مدیریت فروش را می‌توان این‌گونه تفسیر کرد که چه چیزی، به کدام مشتری، در چه زمانی و از طریق چه راهی ارایه شود؟

تحلیل آماری یا داده‌کاوی؟

بسیاری در نگاه اول داده‌کاوی را همان تحلیل آماری می‌دانند، اما این درست نیست و داده‌کاوی چیزی فراتر از آمار است و البته باید دانست که داده‌کاوی، هم علمی برآمده از آمار است و هم از آمار در کارهای خود استفاده می‌کند.

هر چند داده‌کاوی شباهت‌هایی به تحلیل آماری دارد، ولی داده‌کاوی از جهات زیادی با آمار متفاوت است و مزیت‌های زیادی نسبت به آمار دارد. یکی از اصلی‌ترین تفاوت‌های داده‌کاوی با تحلیل‌های آماری این است که در آمار، ما فرضیه‌ای طرح می‌کنیم و با استفاده از تحلیل‌های آماری به اثبات یا رد فرضیه می‌پردازیم اما داده‌کاوی به فرضیه احتیاجی ندارد.

در حقیقت، ابزار داده‌کاوی فرض می‌کند که خود شما هم نمی‌دانید به دنبال چه می‌گردید و این نکته‌ای است که باعث می‌شود کارآمدی داده‌کاوی در مواقع بروز مشکل، به‌کمک بیاید و دلایل نامعلوم مشکلات را برایمان روشن کند.

برای مثال، در آمار فرض می‌شود که ۲ گروه فاصله‌ای باهم ارتباط دارند سپس با استفاده از ضریب همبستگی پیرسون مشخص می‌کنیم که ارتباط وجود دارد یا خیر؟ ولی داده‌کاوی بدون توجه به اینکه ما این‌گونه فرضی داشته باشیم یا نه، با کاوش میان داده‌ها ارتباط پنهان و معنادار را به اطلاع ما می‌رساند. تفاوت بعدی آمار و داده‌کاوی در این است که آمار فقط می‌تواند از داده‌های عددی استفاده کند، ولی داده‌کاوی از داده‌های غیر عددی هم استفاده می‌کند.

منابع

- [۱] گفتگو نگارنده با پرفسور غلامرضا نخعی زاده (مدیر گروه دانشگاه کالسرروه و مدیر سابق بخش داده‌کاوی شرکت بنزآلمان)
- [2] Robert Orfali, Dan Harkey, Jeri Edwards (1998), Client/Server Survival Guide
- "Review of Data mining", The American Statistician, 52, 112-118.
- [3] Jeffery W. Seifert (2004), "Data Mining: An Overview", Analyst in information science and technology policy.

پی‌نوشت‌ها

- Data mining
- Database
- Artificial Intelligence (AI)
- Machine learning
- Integration of organization

بیش از پیش نشان می‌دهد. داده‌کاوی علمی است که با تلفیق علم آمار، فناوری پایگاه داده، فناوری هوش مصنوعی^۳، یادگیری ماشین^۴ و تکنولوژی‌های دیگر رویای کشف رازهای پنهان را محقق می‌کند.

برای شناخت معنای داده‌کاوی در مقام تعریف، می‌توان گفت داده‌کاوی استخراج دانش و اطلاعات پنهان در یک یا چندین پایگاه داده است. این کار با استفاده از تجزیه و تحلیل ماشینی داده‌های ذخیره شده در یک مجموعه از پایگاه‌های داده برای پیدا کردن الگوهای مفید و تازه و قابل استناد انجام می‌شود.

نباید تصور کرد که این علم تنها کاربردهای سراسری و بزرگ دارد، چون این علم می‌تواند در کارهای کوچک و حتی در طرح‌های خوداشتغالی که در حد یک کارگاه کوچک شکل می‌گیرد استفاده شود. داده‌کاوی در هر جایی که اطلاعات به‌صورت بانک اطلاعاتی ذخیره شوند، می‌تواند به کار رود. در نهایت، کاربرد این علم در یک کلام این است: «پشتیبانی از تصمیم».

دو نمونه از کاربرد داده‌کاوی

نمونه اول: تصور کنید یک شرکت خودروسازی محصولاتش را به چندین منطقه آب و هوایی مختلف می‌فرستد. پس از مدتی شکایتی درباره خرابی قطعه‌ای خاص در یک سری از تولیدات دریافت می‌کند. اگر مدیر مجموعه خودروساز، به‌روش ساده بخواهد به این موضوع رسیدگی کند، باید اول خسارت مشتریان را بپردازد و بعد کسانی را که در این مشکل نقش داشته‌اند توبیخ کند، اما این روش حتما درست نیست و با استفاده از داده‌کاوی این‌گونه می‌توان به موضوع رسیدگی کرد که: اطلاعات مربوط به محصولات معیوب اعم از نوع عیب، فاصله زمانی و مکانی فروش تا محل استفاده و نیز شرایط محیطی را در یک پایگاه داده جمع کند و با استفاده از آن به این نتیجه برسد که آیا شرایط محیطی در این موضوع دخیل بوده است؟

این بررسی می‌تواند نشان دهد که محصولات معیوب در ناحیه‌ای بیابانی که جاده‌های آن دارای سطح بالاتری از نمک هستند تردد داشته‌اند یا نه و آیا این موضوع باعث معیوب شدن خودرو شده است یا خیر؟

نمونه دوم: در یک فروشگاه لوازم خانگی، پس از داده‌کاوی مشخص می‌شود که درصدی از مشتریان پس از خرید تلویزیون، همان روز میز تلویزیون هم خریداری کرده‌اند. پس مالک فروشگاه می‌تواند تصمیم بگیرد که براساس موجودی تلویزیون، میزهای متناسب با آن نیز برای عرضه در فروشگاه