

Advanced Analytics

فهرست مطالب

فصل اول : مقدمه

- ۱. طرح موضوع
- ۲. لغت نامه

فصل دوم: تعاریف و مفهوم های ترکیبی

- اصطلاحات
- مفاهیم ترکیبی

فصل سوم: تهدیدها و فرصتها

- فرصتهای پیش رو در زمینه های مختلف
- تهدیدهای پیش رو در زمینه های مختلف

فصل چهارم: معرفی زیست بوم تجزیه و تحلیل پیشرفته اطلاعات

- ویژگیهای اصلی زیست بوم
- عناصر اصلی تشکیل دهنده زیست بوم
- ساختار کلی زیست بوم
- ذینفعان زیست بوم
- نحوه ایجاد ارزش در این زیست بوم
- روش انجام تجزیه و تحلیل اطلاعات
- حوزه های کاربرد در صنایع مختلف
- مروری بر وضعیت بازار تجزیه و تحلیل پیشرفته
- نگاهی به آینده

فصل پنجم: نتیجه گیری

- نتیجه گیری

فصل ششم: پیوستها

فصل هفتم: منابع

فصل اول

مقدمه

۱. طرح موضوع

آیا تا به حال فکر کرده اید که در جهانی مملو از داده ها و اطلاعات خام و پردازش شده قدم میزنیم؟ هر جا که هستیم، هر کاری که میکنیم و هر تفکر که داشته باشیم فرقی نمیکند! ما در حال احاطه شدن توسط سرویسها، خدمات و تجهیزاتی هستیم که بطور مستمر به تولید داده از زوایای مختلف زندگی شخصی و جهان اطراف ما می پردازند. آیا میتوان از این اطلاعات برای پاسخ به سوالات زیر استفاده کرد؟

- اگر بتوانیم از این اطلاعات به تحلیل اتفاقات گذشته بپردازیم؟
- آیا میتوان با تجزیه و تحلیل اطلاعات به رمز موفقیت رقبا پی برد ؟
- آیا میتوان به علل شکست خود پی ببریم ؟
- آیا میتوان به تصمیم های قبلی و آثار و نتایج آنها نگاهی سیستمی بیندازیم؟
- آیا میتوانیم از روی داده های گذشته، وقایع آینده را پیش بینی کنیم؟
- آیا میتوانیم از روی داده های گذشته، وقایع آینده را شبیه سازی کنیم؟

پاسخ سوالات فوق و دهها سوال دیگر، در داده ها و اطلاعاتی هستند که به صورت مستمر در حال تولید شدن هستند. داده هایی که با ظهور موتورهای جستجوگر، شبکه های اجتماعی و شرکت های خرید اینترنتی حجم بسیار عظیمی از آنها به صورت لحظه یی، ساعتی و روزانه تولید می شوند. این داده ها به صورت غیر قابل باوری شامل اطلاعات پیدا و پنهان بسیار زیادی از کاربران است که برخی از این اطلاعات شامل دسته بندیهای ذیل است:

- اطلاعات مربوط به عادات و رفتارهای کاربران شامل طبقه اجتماعی، ساعات فعالیت و استراحت، دوستان و همکاران، زمینه های فعالیت، عادات غذایی ، عادات سلامتی، الگوهای خرید و فروش.
- اطلاعات مربوط به علاقه مندیهای کاربران شامل گرایشهای سیاسی و اجتماعی، موضوعات مورد پسند و عدم پسند، مسیرهای رفت و آمد، عادات رانندگی، هزینه های زندگی، برندهای مورد علاقه و

ابزارهای و سیستمهای اطلاعاتی به طور خواسته یا ناخواسته داده های بسیاری از الگوهای پنهان زندگی و رفتاری کاربران را به صورت مستمر جمع آوری میکنند. این اطلاعات توسط ابزارهای سنتی تجزیه و تحلیل قادر به پردازش شدن نیستند.

فناوریهای تجزیه و تحلیل پیشرفته به رمزگشایی از این اطلاعات پنهان و همچنین عظیم کمک می کنند، تنها پس از آن است سازمان ها میتوانند از این داده استفاده کنند تا دیدگاه های ارزشمندی را در مورد تفکر مشتریان کشف کنند. در نهایت آگاهی از نیاز مشتریان، به سازمان ها کمک میکنند تا نیازهای مشتریان را در اسرع وقت شناسایی و پوشش دهند.

تجزیه و تحلیل پیشرفته کمک می کند تا در تصمیم گیری های کسب و کار آگاهانه عمل شده و با درک کامل از بازار، تحلیل عمیق از بازار نیز به مدیران سازمانها ارائه گردد.

تقاضای روزافزون برای فناوریهای تجزیه و تحلیل داده های بزرگ باعث ایجاد یک بازار رو به رشد برای این صنعت شده است. لازم به ذکر است که اصلی ترین عامل در توسعه این فناوریها، نرخ رشد بسیار سریع داده های در سازمانها و به طور کلی در جهان میباشد.

منطقه آمریکای شمالی دارای بزرگترین سهم بازار در سراسر جهان است و پس از آن اروپا و آسیا و اقیانوسیه قرار دارند. ایالات متحده و کانادا به دلیل پیشرفت های تکنولوژیکی و افزایش استفاده از هوش مصنوعی در بازار آمریکای شمالی جلوتر از بقیه کشورها میباشدند. این منطقه همچنین دارای زیرساخت مناسبی است که امکان اجرای سریعتر فناوریهای پیشرفته را فراهم می نماید. عامل مهم دیگر در پیشرو بودن این منطقه، افزایش روز افزون رسانه های اجتماعی در این کشورها است.

هم اینک، تجزیه و تحلیل داده به یک ابزار استراتژیک برای اکثر سازمان ها تبدیل شده است . بدیهی است که جمع آوری و یکپارچه سازی داده ها به تنهایی ارزشی را تولید نمی کند و نیازمند مدل های پیشرفته جهت تجزیه و تحلیل داده می باشیم ، مدلهایی که می بایست همسو با استراتژی تجزیه و تحلیل و با هدف تولید ارزش تعریف شوند. به باور بسیاری از کارشناسان ، تجزیه و تحلیل داده صرفا یک فناوری نیست ، بلکه روشی است برای انجام

کسب و کار. به عنوان نمونه ، تنها با تجزیه و تحلیل مبتنی بر داده صحیح و به روز است که می توان در مسیر درست جذب ، رشد و حفظ مشتریان حرکت کرد.

تصمیم گیری بر اساس تحلیل داده های صحیح و به روز شده و انجام اقدامات مناسب و به موقع یکی از آرزوهای هر فرد و یا بنگاه تجاری جهت نیل به موفقیت مستمر است . اهمیت تحلیل داده در عصر حاضر با توجه به واقعیت هایی نظیر داده های عظیم و فضای شدیداً رقابتی ، بیش از گذشته شده است . هم اینک ، تعداد بسیار زیادی از سازمان ها با یک رویکرد کاملاً استراتژیک با هوش کسب و کار و تجزیه و تحلیل داده برخورد می کنند و برای آن حساب ویژه ای باز کرده اند . طعم تلخ تصمیم گیری مبتنی بر داده جزیره ای و پیامدهای منفی آن را بسیاری از سازمان ها چشیده اند و یا در حال مقابله با آن هستند .

در گذشته و همزمان با بکارگیری سیستم های اطلاعاتی در سازمان ها که اغلب در قالب یک سری از پروژه های جداگانه اجراء می شد، شاهد تولد سیلوهای (بخشهای جداگانه) اطلاعاتی یکی پس از دیگری بوده ایم که در عمل نمی توانستند به مدیران ارشد و میانی یک سازمان بینش عمیق و وسیعی جهت تصمیم گیریهای بهتر را ارائه کنند.

تبعیت از یک رویکرد استراتژیک در حوزه هوش کسب و کار و تجزیه و تحلیل داده ، یک فرآیند تک مرحله ای و یک شبه نیست. ایجاد زیرساخت لازم به همراه پلتفرم مورد نیاز و همچنین مهارت های لازم جهت تجزیه و تحلیل به زمان کافی و منابع مختلفی نیاز دارد . روش ها و مفاهیم جدیدی نظیر پیاده سازی چالاک و نوآوری های فنی نظیر رایانش ابری ، موبایل ، کشف الگوها داده و داده های عظیم باعث شده است که نیاز سازمان ها به تجزیه و تحلیل داده افزایش پیدا کند . بسیاری از مراکز و موسسات تحقیقاتی و پژوهشی معتبر مدلهای مختلفی از سطوح بلوغ در حوزه تجزیه و تحلیل داده را ارائه کرده اند با این امید که سازمان ها بتوانند به کمک آنها ، ضمن شناسایی دقیق سطح بلوغ فعلی خود، گام های اساسی جهت گذر از یک مرحله به مرحله بعد را مشخص نمایند. سازمانها با استفاده از این مدلهای بلوغ میتوانند به صورت هدفمند و با برنامه زیری - هم نقاط ضعف خود را شناسایی کنند و هم به سمت نقاط مقصد بعدی - حرکت کنند.

۲. لغت نامه

با توجه به گستردگی فناوریهای مرتبط با تجزیه و تحلیل داده ها در این بخش به برخی مفاهیم پایه پرداخته شده است. این بخش صرفاً شامل عناوین و لغت‌های مورد استفاده مکرر در این صنعت است. در بخش بعدی تمامی مطالب مرتبط با این مفاهیم آمده است.

Title	عنوان
Hyper-Connected Collaborative Networks	شبکه های تعامل پذیری که کاملاً به هم متصل هستند.
Big Data	کلان داده
Data Warehouse	انبار داده
Data Analytics	تجزیه و تحلیل داده
Business Intelligence	هوش تجاری
Advanced Analytics	تجزیه و تحلیل پیشرفته
Predictive analytics	تجزیه و تحلیل پیش بینی
Data Mining	داده کاوی
Machine Learning	یادگیری ماشین
Predictive Analytics	تجزیه و تحلیل پیش گوینه
Dark Data	داده های تیره
Prescriptive Analytics	تجزیه و تحلیل تجویزی
Descriptive Analytics	تجزیه و تحلیل توصیفی
Diagnostic Analytics	تجزیه و تحلیل تشخیصی

فصل دوم

مفاهیم و تعاریف

۱. تعاریف و اصطلاحات

- شبکه های تعامل پذیری کاملاً به هم متصل

شبکه های همکاری متقابل (HCNs) شامل صدها نفر و میلیون ها آیتم میباشد که مشغول پردازش و تجزیه و تحلیل داده های عظیم تولید شده از طریق شبکه به هم متصل (CN) میباشند. این شبکه نیازمند به فرآیندهای هوش تجاری در همه جای شهر دارد تا در سطوح مختلف (از سطح افراد تا سطح سازمان) اجرا گردد.

- کلان داده (مطابق تعریف گارتنر)

کلان داده (Big Data) به معنای دارایی های اطلاعاتی [یک مجموعه یا سازمان] است که دارای ویژگیهای زیر باشد:

- حجم بالا داشته باشند.

- با سرعت زیاد تولید می شوند و / یا تنوع گسترده دارند

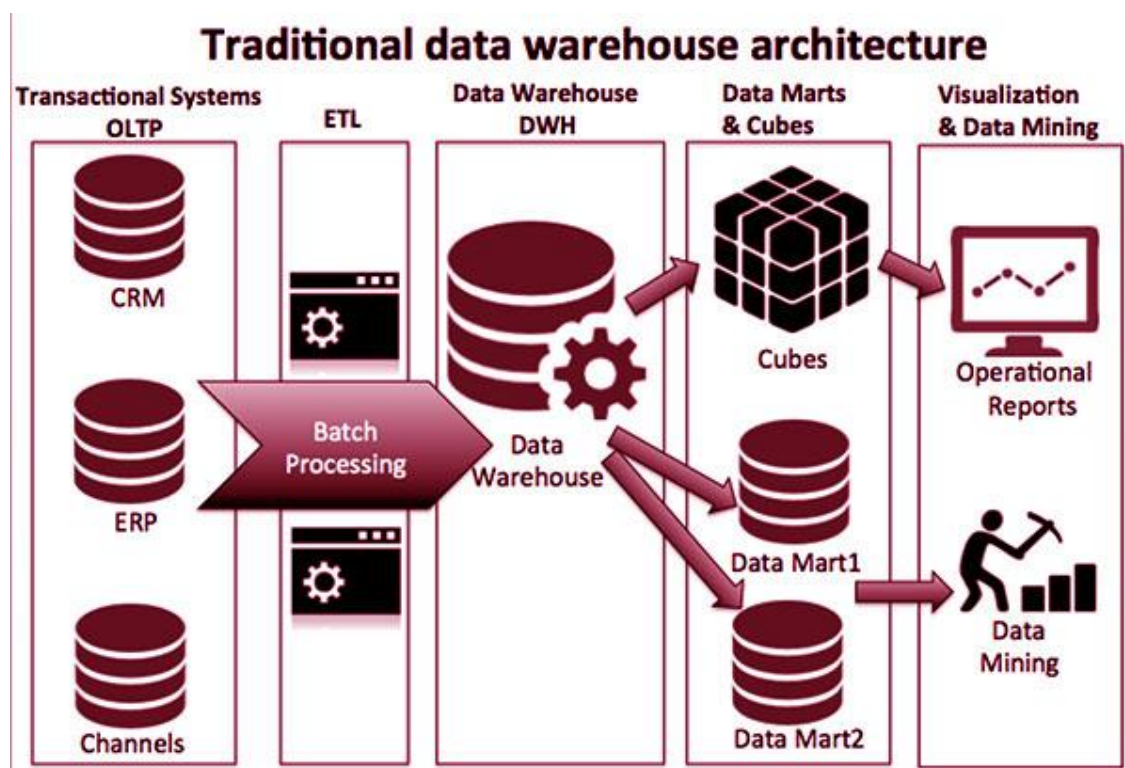
و نیازمند شیوه های پردازش نوآورانه با هزینه ی مناسب هستند تا بتوان از آن برای اتوماسیون فرایندها، تصمیم گیری و بهبود شهود و بینش [در سازمان] بهره گرفت.

- انبار داده

انبار داده به مجموعه ای از داده ها گفته می شود که از منابع مختلف اطلاعاتی سازمان جمع آوری، دسته بندی و ذخیره می شود. انبار داده یا Data Warehouse پایگاه داده ای است که برای گزارش گیری و تحلیل داده به کار می رود و بعنوان هسته اصلی یک سیستم BI به شمار می آید. به عبارت دیگر انبار داده یک مخزن داده مرکزی از داده های تجمیع شده است که از سیستم ها و منابع مختلف سازمان جمع آوری شده است.

انبار داده یک بانک اطلاعاتی رابطه ای غیرنرمال است که داده های حال و گذشته را در یک مکان واحد تجمیع کرده است و هدف اصلی آن پوشش گزارش گیری و نیازهای تحلیلی یک سازمان به کار گرفته می شود.

از اواسط سالهای 1980 نیاز به انبار داده‌ها به وجود آمد و در کنار سیستم‌های اطلاعاتی (OLTP) سیستم‌های تحلیلی (OLAP, DW/BI) ایجاد شد. در شکل بعدی به یک معماری کلی از یک انبار داده اشاره شده است.



• تجزیه و تحلیل داده

برای تجزیه و تحلیل تاکنون تعاریف مختلفی ارایه شده است . به دست آوردن بینش از داده به منظور تصمیم گیری بهتر ، یکی از تعاریفی است که بر روی آن اتفاق نظر فراوانی وجود دارد . به عبارت دیگر ، تجزیه و تحلیل فرآیندی علمی جهت تبدیل داده به بینش و تصمیم گیری بهتر و متعاقب آن اقدامات مناسب است.

• هوش تجاری

به طور معمول تمرکز بر استفاده از یک مجموعه سازگار از معیارهای اندازه گیری عملکرد گذشته و راهنمای برنامه ریزی کسب و کار است. اطلاعات کسب و کار شامل پرس و جو، گزارش دهی، OLAP (پردازش تحلیلی آنلاین) می باشد و می تواند به سؤالاتی گذشته نگرا نه زیر جواب بدهد :

✓ چه اتفاقی افتاده است؟

✓ چگونه این اتفاق رخ داده است؟

به طور کلی هوش تجاری رویکرد بررسی گذشته اتفاقات و رخدادها میباشد.

• تجزیه و تحلیل پیشرفته

اما فراتر از هوش کسب و کار، استفاده از تکنیک های مدل سازی پیچیده برای پیش بینی وقایع

آینده یا کشف الگوهایی که نمی توانند در حالت عادی شناسایی شوند، لذا موضوع تجزیه و تحلیل

پیشرفته می تواند به مجموعه ای از سوالات آینده نگرانه زیر پاسخ دهد:

✓ چرا این اتفاق می افتد؟

✓ اگر این روند ادامه یابد چه اتفاقی خواهد افتاد (پیش بینی)

✓ بهترین کاری که ممکن است در پاسخ به این رخداد انجام شود چیست؟

• اطلاعات تیره

داده های تیره در فناوری شامل اطلاعات دیجیتالی است که در حال حاضر برای تجزیه و تحلیل

کسب و کار استفاده نمی شود. این داده ها از طریق عملیات های شبکه کامپیوتری مختلف به دست می

آیند اما به شیوه ای مرسوم برای به دست آوردن بینش یا تصمیم گیریها مورد استفاده قرار نمیگیرند.

باید توجه داشت که تجزیه و تحلیل داده ها و اطلاعات عملکرد روزانه سازمان ها را بهبود داده و نیاز به

درک بیشتر تمامی اطلاعات وجود دارد زیرا هر گونه اطلاعات ناشناخته، فرصتی است که ممکن است

به یک خطر امنیتی بالقوه منجر شود و یا اینکه میتوان از اطلاعات به عنوان یک فرصت بهبود استفاده

کرد که با عدم پردازش صحیح آنها صبیحاً از دست خواهد رفت.

• محاسبات کوانتومی

غول های فناوری مانند آی بی ام، مایکروسافت، گوگل و اینتل به سختی با یکدیگر رقابت می کنند

تا اولین کامپیوتر کوانتومی را بسازند. محاسبات کوانتومی، رمزنگاری داده های بدون درز، پیش بینی

آب و هوا، حل مشکلات پیچیده پزشکی، مکالمات واقعی و مدل سازی مالی بهتر را برای ایجاد سازمان

ها، اجزای محاسبات کوانتومی، الگوریتم ها، برنامه های کاربردی و ابزارهای نرم افزاری بر روی خدمات ابری را قادر می سازد.

• تجزیه و تحلیل تشخیصی

در تجزیه و تحلیل تشخیصی (Diagnostic Analytics) ، هدف اصلی پاسخ به این نوع سوالات است:

- چرا این اتفاق افتاده است ؟
- چرا این اتفاق دارد می افتد ؟
- روند به چه صورت است ؟
- چه الگویی وجود دارد ؟

با بررسی و تحلیل سوابق داده (ارتباط با داده گذشته) ، بینش ایجاد می گردد .

• تجزیه و تحلیل توصیفی

در تجزیه و تحلیل توصیفی (Descriptive Analytics) ، هدف اصلی پاسخ به این نوع سوالات است :

چه اتفاقی افتاده است ؟ چه چیزی دارد اتفاق می افتد؟ تا در نهایت برای تصمیم گیری بینش تولید شده را پیش روی تصمیم گیرندگان انسانی قرار دهد .

• تجزیه و تحلیل تجویزی

در تجزیه و تحلیل تجویزی (Prescriptive Analytics) ، هدف اصلی پاسخ به این نوع سوالات است:

چه کاری می بایست انجام داد تا این اتفاق بیافتد؟ پاسخ به این سوال شامل مجموعه اقداماتی است که می بایست هم اینک انجام داد که می توان آنها را در زمره تصمیمات عملیاتی در نظر گرفت و یا اقداماتی که می بایست در آینده انجام داد که می توان آنها را به عنوان تصمیمات تاکتیکی برای کوتاه مدت و تصمیمات استراتژیک برای بلند مدت در نظر گرفت .

۲. مفاهیم ترکیبی

• تفاوت بین هوش تجاری و تجزیه و تحلیل پیشرفته

باید در نظر داشت که مفاهیم هوش تجاری بر روی گزارشگیری و پرس و جو متمرکز است، اما مفاهیم تجزیه و تحلیل پیشرفته در مورد بهینه سازی، همبستگی و پیش بینی برای انجام بهترین اقدام بعدی و یا صرفاً اقدام بعدی است.

در حالی که ابزارهای تجزیه و تحلیل سنتی که شامل اطلاعات کسب و کار پایه (BI) داده های گذشته را بررسی می کنند، ابزارهای تجزیه و تحلیل پیشرفته بر پیش بینی رویدادها و رفتارهای آینده تمرکز می کنند و شرکت ها را قادر می سازد تا تجزیه و تحلیل هایی را انجام دهند تا اثرات تغییرات بالقوه در استراتژی های تجاری را پیش بینی کنند.

• ارتباط بین کلان داده (Big Data) و تجزیه و تحلیل پیشرفته

کلان داده ها خوراک ورودی برای سیستمهای تجزیه و تحلیل داده میباشند. عامل و پیشرانه اصلی برای سیستمهای تجزیه و تحلیل داده ها دسترسی کلان داده ها به منظور خودکار سازی تحلیل رفتار گذشته و یا روابط بین مشتریان و نیز مبادلات مبتنی بر بازار میباشد. در ادامه به برخی از منابع تولید کلان داده پرداخته میشود:

○ داده های مربوط به شبکه های اجتماعی : حجم بسیار زیادی از اطلاعات ساخت

یافته، متن، عکس، صوت و تصویر در سایتهای مربوط به شبکه های اجتماعی یافت

میشود. این داده ها شامل اطلاعات خامی است که میتواند توسط هر شخص یا شرکتی

مورد بررسی قرار گرفته و دانش و بینشهای ارزشمندی تولید گردد.

○ داده های مربوط به موبایل : در حال حاضر حدود 5 میلیارد گوشی موبایل در

سرتاسر جهان مورد استفاده قرار میگیرد. این گوشیها میتوانند داده های ارزشمندی

مانند موقعیت کاربران، نحوه استفاده از تجهیزات موبایل و همچنین امکاناتی که ممکن است مورد استفاده قرار گیرد.

○ **تراکنشها:** بسیاری از تجهیزات مانند کارتهای اعتباری و بانکی، کیفهای پول الکترونیکی امکان خرید و فروش – بدون وجه نقد- کالا و خدمات را فراهم آورده است. سیستمهای مرتبط با این تجهیزات اطلاعات مربوط به معاملات را برای استفاده بعدی ذخیره میکنند.

○ **مرور صفحات وب و جستجوی اطلاعات:** بسیاری از وبسایتهای عادی مربوط به جستجو و مرور اطلاعات را ذخیره میکنند. البته فقط برخی از این وبسایتهای اطلاعات مربوطه را تجزیه و تحلیل کرده و از نتیجه این تحلیل به منظور بهبود محتوای سایت استفاده میکنند.

○ **مستندات:** مستنداتی مانند بیانیه ها، قراردادهای بیمه، رکوردهای پزشکی و مکاتبات با مشتری میتوانند در سامانه ها ذخیره شوند و سپس اطلاعات پایه ارزشمند از درون این اسناد استخراج و ذخیره خواهند شد.

○ **اطلاعات مربوط به کانالهای تلویزیونی و رادیویی دیجیتال:** تعامل کاربران با این تجهیزات، اطلاعاتی را در مورد نحوه استفاده کاربران از آنها فراهم نموده و می تواند تجربه کاربر را در تعامل با محتوای چند رسانه ای بهبود ببخشد.

○ **اینترنت اشیاء:** اشیاء متصل به اینترنت حجم بسیار زیادی از اطلاعات را تولید میکنند که دارای انواع گوناگونی بوده و از منابع اطلاعاتی مختلف وارد شبکه میشوند. برخی از این منابع تولید اطلاعات شامل کتابهای الکترونیکی، دوچرخه، ماشین، بازیهای ویدئویی، عینک و ساعت و کنتورهای برق و آب و گاز و هزاران وسیله دیگر.

○ **رخدادهای شبکه ارتباطی:** شبکه بیشتر به عنوان تعدادی وسایل به هم متصل شناخته میشود. نتایج اتصال نیازمند به مانیتور کردن حجم بسیار زیادی از اطلاعات است که باید سریعاً بررسی و تغییرات را اعمال نمایند.

○ **ثبت رکوردهای تماس :** بررسی تماسهای تلفنی کمک میکند تا عاداتهای گوناگون

مشتریان کشف شود با شناسایی شبکه ارتباطی افراد میتوان خدمات بهتری به مشتریان

ارائه نمود.

○ **اطلاعات مربوط به RFID :**

RFID ها اطلاعات فراگیر و ارزشمند هستند که اغلب مورد بررسی قرار می گیرند و تحلیل

نمی شوند.

○ **داده های ترافیکی و حمل و نقل:**

سنسورهای جاده یی، دوربین های ویدئویی و داده های شناور ماشین، تکنیک های اخیر

برای مطالعه الگوهای ترافیکی هستند. برای جلوگیری از ازدحامهای ترافیکی، این داده ها

نیاز به تجزیه و تحلیل سریع دارند.

○ **داده های آب و هوا:** در حال حاضر، داده های آب و هوایی میتوانند به منابع مختلف

دیگری از اطلاعات مانند اطلاعات فروش، بازاریابی و اطلاعات مرتبط با محصول متصل

شوند، به طوری که شرکتها میتوانند محصولات خود را به طور موثر تر به فروش رسانده

و هزینه ها را کاهش دهند.

• **موارد استفاده تجزیه و تحلیل پیشرفته داده**

فناوری تجزیه و تحلیل پیشرفته داده ها در تمامی صنایع و برای پیش بینی وقایع استفاده می شود.

برخی از این صنایع و کاربردها عبارتند از :

✓ تیم های بازاریابی اینترنتی آن را برای پیش بینی احتمال کلیک کردن برخی از کاربران

وب روی یک لینک مشخص مورد استفاده قرار میدهند.

✓ ارائه دهندگان مراقبت های بهداشتی از تجزیه و تحلیل توصیفی برای شناسایی بیمارانی -

که ممکن است از یک درمان خاص بهره مند شوند- متمرکز میشوند.

✓ ارائه دهندگان شبکه تلفن همراه از تجزیه و تحلیل تشخیصی برای پیش بینی شکست

بالمقه شبکه های مخابراتی استفاده نموده و آنها را قادر به نگهداری پیشگیرانه از شبکه

ارتباطی می کنند.

فن آوری های تجزیه و تحلیل پیشرفته مانند داده کاوی، استخراج متن و جستجوی وب بیشتر توسط شرکت

هایی مورد استفاده قرار میگیرد که بروی مصرف کنندگان و مشتریان خود تمرکز بسیار ویژه ای دارند.

تجزیه و تحلیل پیشرفته داده ابزاری هستند که در مدیریت تصمیم گیری نقش مهمی داشته و به سرعت در حال

رشد است. این ابزارهای تجزیه و تحلیل در تمامی سطوح مفهومی سازمان - سطح استراتژیک سازمانی، سطح

تاکتیکی و سطح عملیاتی - راه گشا بوده و امکان فراهم آوردن خدمات زیر برای سطوح مختلف سازمانی فراهم

می آورد.

✓ مدیران استراتژیک برای تصمیم گیری و همچنین در زمینه هوش رقابتی ، شناسایی فرصت های بازار،

تصمیم گیری در مورد راه اندازی خط جدید محصولات استفاده میکنند.

✓ مدیران تاکتیکی از آنها برای تصمیم گیری در زمینه پیش بینی فروش، بازاریابی مستقیم، خرید، نگهداری

و اهداف توسعه یی، و تجزیه و تحلیل کمپینهای بازاریابی استفاده می کنند.

✓ مدیران عملیاتی نیز از آنها برای تصمیم گیری در مورد استفاده بهتر از امکانات و یا مدیریت زنجیره تامین

استفاده میکنند.

• فرآیند تجزیه و تحلیل پیشرفته داده

اصطلاح تجزیه و تحلیل پیشرفته، در واقع معرف یک فرایند کلی است از تبدیل داده های ریز و جزئی به داده ها

و دانشهای سطح بالا که معمولاً با استفاده تکنیکهای طبقه بندی و گروه بندی داده ها، پایگاه داده، متون و وب

انجام میشود.

استخراج داده ها از فرمتهای سه گانه (داده، متن و وب) ، الگوهای مخفی شده در آن اطلاعات را با استفاده از

تکنیک های پیش بینی آینده آشکار میکنند. این الگوها نقش مهمی در تصمیم گیریها را ایفا می کنند زیرا آنها

زمینه هایی بهبود روندها را مشخص میکنند. با استفاده از داده کاوی ، سازمان ها می توانند سودآوری مشتریان خود را افزایش دهند، قلب را تشخیص داده و مدیریت ریسک را نیز بهبود بخشند. الگوهای موجود در فرایند داده کاوی به سازمانها کمک می کنند تا تصمیماتی بهتر و سریع تر اخذ کنند.

نمایه سازی (مصورسازی) اطلاعات مربوط به یک مشتری خاص و همچنین طبقه بندی و تقسیم بندی مشتریان از نیازهای روزمره همه سازمانها و شرکتهای میباشد. وظیفه اصلی نمایه سازی مشتریان در واقع شامل شناسایی گروه های همگن از مشتریان است که الگوهای رفتاری مشابهی را از خود نشان می دهند.

با استفاده از این اطلاعات، در واقع شرکت ها یاد می گیرند تا بجای انتخاب مشتریان به صورت تصادفی، مشتریان خاصی را هدف تبلیغات و برنامه های خود قرار دهند کاملاً روشن است که انتخاب مشتریان به صورت تصادفی باعث افزایش هزینه های بازاریابی و عدم دریافت نتیجه مناسب میگردد.

تقسیم بندی مشتریان می تواند به نوآوری های شرکت در طراحی و قیمت گذاری محصولات جدید - که برای مصرف کنندگان جذاب تر هستند - نیز کمک کند.

یک مثال از موارد فوق مربوط به ارائه طرحهای بیمه برای صنعت سلامت میباشد، شرکت های بیمه درمانی می توانند برنامه های سلامتی را - که مبتنی بر مشخصات ریسک مشتریان میباشد - ارائه دهند،

در ادامه گردش کاری نشان داده شده است که شامل مراحل مختلف لازم برای استفاده از تجزیه و تحلیل پیشرفته اطلاعات میباشد. هر یک از این مراحل به صورت تکرار پذیر و تعاملی انجام میشود. تعیین اهداف کسب و کاری میتوانند برای فرایند تجزیه و تحلیل پیشرفته بسیار موثر باشند زیرا این تحلیلها بدنبال پیش بینی و یا به دنبال توضیح رفتارها بوده و در واقع به دنبال یافتن الگوهای رفتاری در اطلاعات میباشد.

برای مثال، ممکن است یک هدف کسب و کاری شامل تهیه بهترین نمایه از مشتریان و یا پیش بینی لیست مشتریانی باشد که بیشترین ریسک را به سازمان تحمیل میکنند.

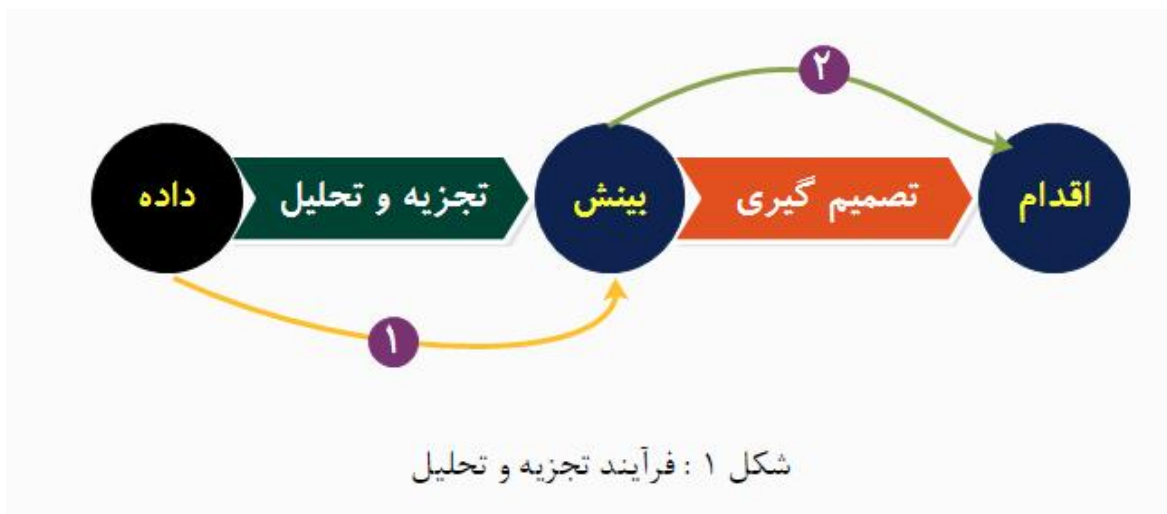
هنگامی که اهداف کسب و کاری تعریف شوند در گام بعدی نیاز است که مشخص کنیم چه اطلاعاتی برای رسیدن به اهداف مورد نیاز بوده و نحوه دستیابی به آن اطلاعات چگونه است؟

در ادامه مراحل 9 گانه مربوط به استفاده از تجزیه و تحلیل پیشرفته اطلاعات نشان داده شده است. این مراحل به صورت کلی نمایش داده شده و هر یک از آنها توضیحات تفصیلی بسیار زیادی را در بر میگیرد.



• چگونگی عملکرد تجزیه و تحلیل داده

تولید بینش از داده ها و تصمیم گیری و اقدام بر اساس آن بینش در واقع هدف کلی این فناوریها میباشد. در شکل 1، فرآیند تجزیه و تحلیل را نشان داده شده است.

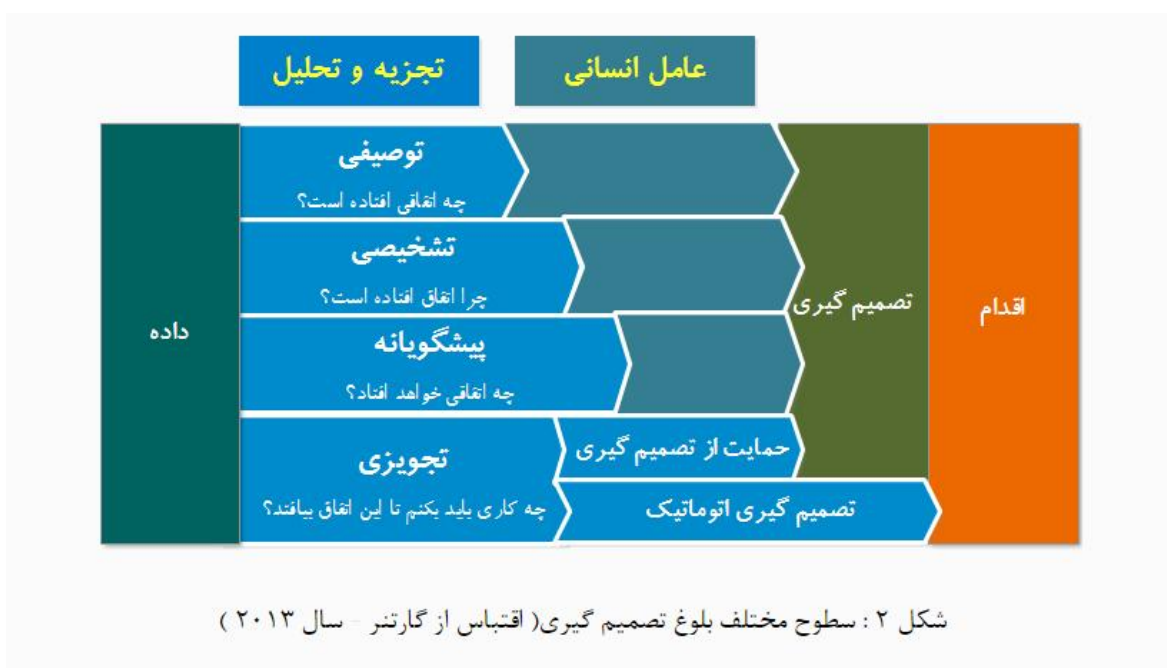


مشاهده شکل فوق نشان دهنده اطلاعات کلی زیر میباشد:

- فاصله بین داده تا بینش با عملیات تجزیه و تحلیل تکمیل می گردد و به نوعی می توان گفت که این موتور تجزیه و تحلیل است که با دریافت داده، بینش لازم را تولید می میکند.
- فاصله بین زمانی که به بینش لازم می رسیم تا زمانی که اقدام متناسب با آن را انجام دهیم به تصمیم گیری بر میگردد. سرعت، دقت و کیفیت در هر دو مرحله فوق بسیار تعیین کننده و مهم می باشد چراکه در یک فضای شدیداً رقابتی زندگی می کنیم و رقبا، منتظر سوار شدن و تاختن ما بر اسب سرنوشت نخواهند شد. چه بسیار سازمان هایی که علی رغم تولید بینش لازم از داده، در مرحله تصمیم گیری دارای اشکالات فراوان هستند و به دلیل نبود یک ساختار منسجم تصمیم گیری، تصمیمات با تاخیر و احتمالاً با ابهام اتخاذ می گردد. همین موضوع باعث می شود که آنها نتوانند در زمان لازم اقدامات مناسب و بایسته ای را انجام دهند و در عمل به اسطوره ای از فرصت سوزی تبدیل شوند. فرصت هایی که شاید شرایط تحقق آنها در هیچ بازه ای از حیات یک بنگاه کسب و کار دیگر تکرار نگردد.

- شاید اگر می توانستیم نقش انسان را در فرآیند تصمیم گیری کمتر کنیم ، فاصله زمانی بین داده تا اقدام کمتر می گردید و می توانستیم مناسب ترین واکنش ها را متناسب با بینش حاصل از تحلیل داده در سریع ترین زمان ممکن انجام دهیم .

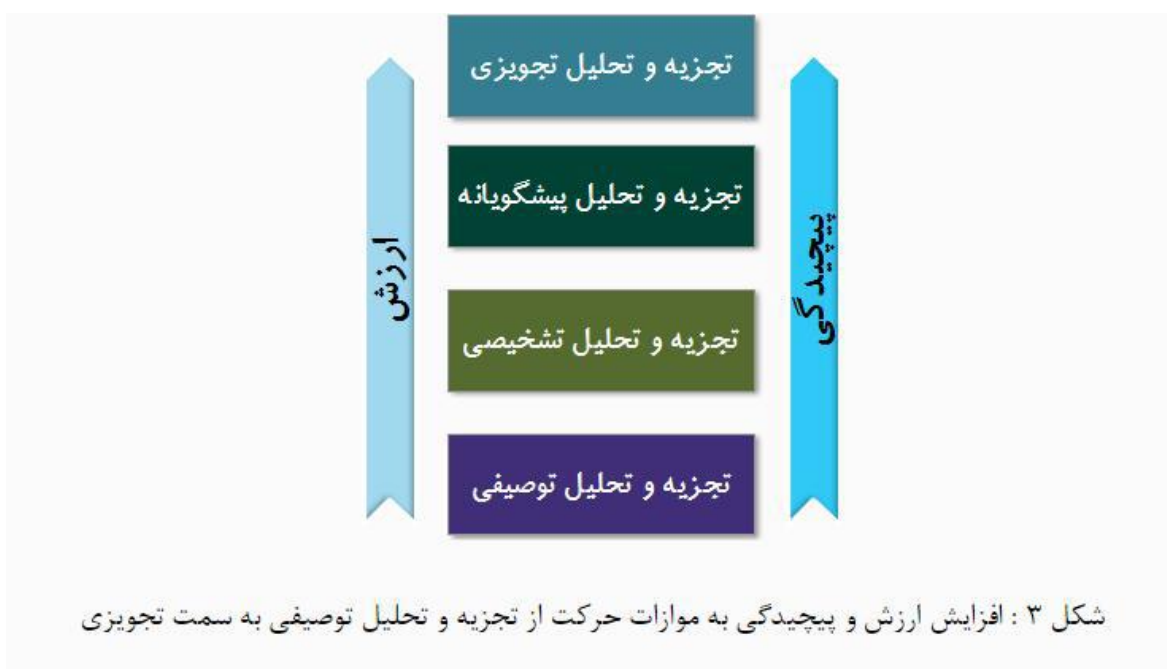
چه میزان از فرآیند تصمیم گیری خودکار است و چه میزان از آن با دخالت عوامل انسانی انجام می شود؟ پاسخ به این پرسش می تواند علل ایجاد سطوح مختلف تجزیه و تحلیل را تشریح میکند. در واقع ، می توان سطوح مختلفی از بلوغ تجزیه و تحلیل را با توجه به میزان خودکار بودن تصمیم گیری تعریف کرد . شکل 2 ، سطوح مختلف بلوغ تصمیم گیری را نشان می دهد .



در شکل فوق ، سوالاتی نظیر چه اتفاقی افتاده است ؟ چرا این اتفاق افتاده است ؟ چه اتفاقی خواهد افتاد و چه کاری را می بایست انجام داد به خوبی تعریف شده اند . تفکیک آن بخشی از تجزیه و تحلیل که به صورت اتوماتیک انجام می شود و آن بخشی که توسط انسان انجام خواهد شد ، یکی از نکات مهم مدل نشان داده شده در شکل 2 می باشد. به عنوان نمونه در تجزیه و تحلیل توصیفی (Descriptive Analytics) که عصاره آن در بکارگیری ابزارهایی نظیر داشبورد و یا query/drill down تبلور پیدا می کند ، هوشمندی کار به طور کامل در اختیار انسان باقی خواهد ماند . هر اندازه که به سمت جلو برویم ، درصد هوشمندی تجزیه و تحلیل افزایش می یابد و

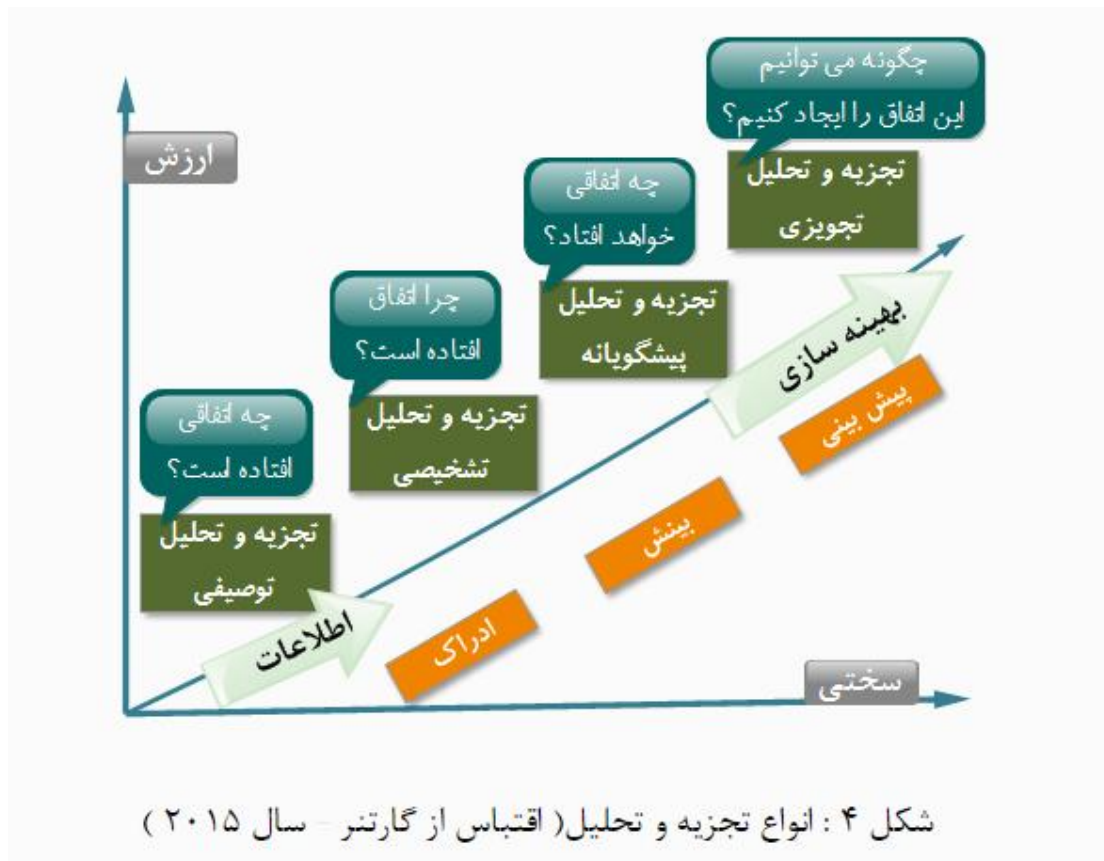
وارد دنیای تجزیه و تحلیل پیشرفته می شویم که دخالت انسان کمتر می گردد. کار انسان از جایی شروع می شود که تجزیه و تحلیل متوقف می شود. به موازات حرکت از تجزیه و تحلیل توصیفی به سمت تجزیه و تحلیل تجویزی به نقاط عطف مهم تصمیم گیری و اقدام می رسیم .

هر نوع و یا مدل تجزیه و تحلیل نشان داده شده در شکل فوق ، از روش ها و الگوریتم های مختلفی استفاده می کنند . این بدان معنی است که با توجه به نوع نیاز و انتظار یک سازمان از تجزیه و تحلیل، شاهد تنوع در بکارگیری نوع های مختلف داده ، امکانات ذخیره سازی و پردازش خواهیم بود تا نتایج دلخواه را ارایه نماید . شکل 3 ، بیانگر این واقعیت است که همزمان با تولید ارزش بالا از نتایج تجزیه و تحلیل ، شاهد افزایش پیچیدگی و هزینه محیط تجزیه و تحلیل خواهیم بود .



• انواع سطوح تجزیه و تحلیل داده

با توجه به میزان اتوماتیک بودن فرآیند تجزیه و تحلیل می توان آن را به چهار گروه توصیفی ، تشخیصی ، پیشگویانه و تجویزی تقسیم کرد . شکل 4 ، انواع مدل های تجزیه و تحلیل را نشان می دهد .



در تجزیه و تحلیل توصیفی (Descriptive Analytics) ، هدف اصلی پاسخ به این نوع سوالات است : چه اتفاقی افتاده است ؟ چه چیزی دارد اتفاق می افتد؟ تا در نهایت برای تصمیم گیری بینش تولید شده را پیش روی تصمیم گیرندگان انسانی قرار دهد . در این نوع تجزیه و تحلیل با دستیابی به سوابق که ریشه در گذشته و حال دارد ، امکان تحلیل داده مبتنی بر رویدادهای گذشته و حال و تولید بینش میسر می گردد . امکاناتی نظیر OLAP, BI, داشبوردها، گزارشات و بصری سازی را می توان در این گروه قرار داد . Google Analytics یک نمونه شناخته شده در این رابطه است. با تجزیه و تحلیل توصیفی ، می توان دنیا را بگونه ای که هست توصیف کرد.

در تجزیه و تحلیل تشخیصی (Diagnostic Analytics) ، هدف اصلی پاسخ به این نوع سوالات است : چرا این اتفاق افتاده است ؟ چرا این اتفاق دارد می افتد ؟ روند به چه صورت است ؟ چه الگویی وجود دارد ؟ با بررسی و تحلیل سوابق داده (ارتباط با داده گذشته) ، بینش ایجاد می گردد . شناخت و تحلیل بخش های مختلف بازار ، تجزیه و تحلیل احساسات ، ریشه یابی بروز مشکلات ، تجزیه و تحلیل لاگ فناوری اطلاعات ، الگوهای تقلب و کشف قیمت و ... نمونه هایی در این زمینه می باشند . تحقیقات علمی نمونه مثال خوبی است که هدف آن ایجاد دانش از تجزیه و تحلیل داده است .

در تجزیه و تحلیل پیشگویانه (Predictive Analytics) ، هدف اصلی پاسخ به این نوع سوالات است : چه چیزی اتفاق خواهد افتاد ؟ چگونه می توان واکنش نشان داد ؟ آیا این یک تقلب است ؟ می توان هر واژه ای را جایگزین واژه تقلب در آخرین پرسش کرد. با تحلیل داده و تولید بینش ، امکان پیش بینی داده جدید در رابطه با حال و آینده که به آن foresight گفته می شود ، فراهم می گردد . پیش بینی رفتار مشتری (گرایش خرید) و میزان ریزش مشتری (ترک یک مشتری از یک شرکت و جذب شرکت رقیب شدن) ، نمونه هایی در این زمینه می باشند . استخراج معنی از متن آزاد نیز یک نوع از تجزیه و تحلیل پیش گویانه است .

در تجزیه و تحلیل تجویزی (Prescriptive Analytics) ، هدف اصلی پاسخ به این نوع سوالات است : چه کاری می بایست انجام داد تا این اتفاق بیافتد؟ پاسخ به این سوال شامل مجموعه اقداماتی است که می بایست هم اینک انجام داد که می توان آنها را در زمره تصمیمات عملیاتی در نظر گرفت و یا اقداماتی که می بایست در آینده انجام داد که می توان آنها را به عنوان تصمیمات تاکتیکی برای کوتاه مدت و تصمیمات استراتژیک برای بلند مدت در نظر گرفت . در حالت کلی ، قواعد تصمیمات عملیاتی از قبل تعریف شده است و می تواند از قالب ساختار معروف جملات شرطی در زبان های برنامه نویسی تبعیت کند (If condition then) . شرط و یا شرایط می تواند بر اساس داده پیش بینی شده باشد. به عنوان نمونه در صورتی که گرایش به خرید یک محصول خاص بیش از میزان آستانه باشد ، چه واکنش و یا اقداماتی می بایست انجام داد . در خصوص تصمیمات تاکتیکی و استراتژیک ، هدف اصلی بهینه سازی و شبیه سازی ابزارهای لازم است . مرتب سازی مجدد برنامه ها ، سیاست ها و زمانبندی نمونه هایی در این رابطه می باشند که می تواند بر روی داده پیشگویانه تاثیرگذار باشد . مثلا زمانی

که میزان فروش در آینده پیش بینی می شود (پیش بینی تقاضا) ، می توان برنامه ریزی لازم جهت راه اندازی یک فروشگاه و بازرنگری در فرآیندهای تدارکات یک فروشگاه و به نوعی مدیریت انبار را انجام داد.

مثال

در شکل 5 با توجه به ماهیت و ماموریت هر یک از انواع سطوح تجزیه و تحلیل ، مثال هایی آورده شده است .

نوع تجزیه و تحلیل	پاسخ به سوالات	مثال
توصیفی	چه اتفاقی افتاده است ؟	سال گذشته چه تعداد اتومبیل فروختیم ؟ در شش ماه گذشته چه تعداد مشتری جذب کرده ایم ؟ در سه ماه گذشته چه تعداد مشتری را از دست داده ایم ؟
تشخیصی	چرا این اتفاق افتاده است ؟	چرا در شش ماه گذشته فقط ۱۰۰۰ اتومبیل فروخته ایم ؟ چرا در شش ماه گذشته فقط ۱۰ مشتری جدید جذب کرده ایم ؟ چرا میزان فروش فصل بهار بیش تر از زمستان بوده است ؟
پیشگويانه	چه اتفاقی خواهد افتاد ؟	اگر ۵ کمپین جدید را اجراء کنیم ، چه تعداد اتومبیل خواهیم فروخت ؟ اگر یک مشتری محصولات الف و ب را خریده باشد ، آیا این احتمال است که محصول ج را هم بخرد ؟ اگر کیفیت محصولات و خدمات خود را ارتقاء دهیم ، چه تعداد به مشتریان خود اضافه می کنیم ؟
تجویزی	چگونه می توانیم کاری کنیم که این اتفاق بیافتد ؟	چه کار باید بکنیم تا بتوانیم ۱۰۰۰۰ اتومبیل را بفروشیم ؟ از این سه دارو ، کدامیک بهترین نتایج را در بهبود بیمار می تواند داشته باشد ؟ بهترین موقع خرید سهام شرکت الف چه زمانی است ؟

شکل ۵ : نمونه مثال هایی از انواع سطوح تجزیه و تحلیل داده

• مزایای تجزیه و تحلیل پیشبینانه چیست؟

در طول تاریخ، پیش بینی نمودن آینده با هر روش و شگردی، همواره سه هدف اساسی را دنبال میکرد است: پول، شهرت و قدرت. تجزیه و تحلیل پیشگیرانه نیز مستثنی نبوده نیست چرا که مردم همواره میخواهند بدانند که ظرف هفته ، ماه آینده و یا سال آینده، چه اتفاقی خواهد افتاد اما باید بدانیم که فناوری جدید دقیقتر از متدها و روشهای دیگر است. در ادامه به سه انگیزه اصلی جهت پیش بینی آینده نگاهی دقیقتر خواهیم انداخت:

پول: تجزیه و تحلیل پیش گیرانه میتواند به افراد کمک کند تا راهی برای ذخیره پول و کسب درآمد پیدا کنند. خرده فروشان اغلب از مدل های پیش بینی برای تخمین نیازهای موجودی انبارها، مدیریت برنامه های حمل و نقل و پیکربندی، طرح های فروشگاه ها برای به حداکثر رساندن فروش استفاده می کنند.

اغلب آژانسها و خطوط هوایی از تجزیه و تحلیل پیشگیرانه استفاده می کنند تا بتوانند قیمت بلیط را در بازه های زمانی آینده تخمین بزنند. همچنین رستورانها و هتلها و دیگر ذینفعان صنعت توریسم بسیار علاقه مند هستند تا از تکنولوژی برای پیش بینی تعداد مهمان ها در هر شبانه روز استفاده کنند تا امکان ایجاد بیشترین میزان اشغال و درآمد را داشته باشند.

شهرت: استفاده از تجزیه و تحلیل پیگیرانه به منظور بهینه سازی کمپین های بازاریابی، سازمانها می توانند پاسخ های مناسب به مشتریان جدید یا ارائه پیشنهادهای مناسب خرید به آنها و همچنین ارتقاء فرصت های فروش متقابل را ایجاد کنند. مدل های پیشگیرانه می توانند به کسب و کارها برای جذب، نگهداری و حمایت از مشتریان با ارزش خود کمک کنند.

قدرت: تجزیه و تحلیل پیشبینی کننده قابل استفاده برای تشخیص و متوقف کردن انواع مختلف رفتارهای جنایی و امنیتی خطرناک قبل از ایجاد هر گونه آسیب جدی است. با استفاده از تجزیه و تحلیل پیشگیرانه برای مطالعه رفتار و عملکردهای کاربر، یک سازمان می تواند فعالیت هایی را که از عادت معمولی نشات میگیرند را - نسبت به عادات متقلبانه که به منظور سوء استفاده از کارتهای اعتباری و جاسوسی از شرکت ها تا حملات سایبری میباشد - تشخیص دهند.

• تفاوت بین تجزیه و تحلیل پیشبینانه و تجزیه و تحلیل سنتی

تفاوت بین تجزیه و تحلیل های معمول و تجزیه و تحلیل پیشبینانه ساده است. تجزیه و تحلیل سنتی بر روی داده های فعلی تمرکز دارد، در حالیکه تحلیل پیشرفته به کاربران امکان می دهد تا آینده نزدیک و درازمدت را با دقت به روندها و رفتارهای آینده تشخیص میدهند.

• مثالهای از تجزیه و تحلیل پیشبینانه

امروزه سازمانها از روشهای تجزیه و تحلیل پیش بینانه در تعداد زیادی از کاربردها استفاده می کنند. این فناوری به افرادی - که در زمینه امور مالی، مراقبت های بهداشتی، خرده فروشی، توریسم، دارو، خودرو، هوا فضا و تولید فعالیت میکنند - کمک می کند.

✓ **صنعت هوا و فضا:** پیش بینی اثرات عملیات تعمیر و نگهداری خاص در قابلیت اطمینان هواپیما،

استفاده از سوخت، دسترسی و زمانهای پرواز.

✓ **خودرو:** ثبت اسناد و مدارک مربوط به استحکام و شکست بخشهای مختلف خودرو، برنامه های

تولید خودرو بعدی. مطالعه رفتار رانندگان ، کمک به رانندگی بهتر با استفاده از اتومبیلهای خود

ران.

✓ **انرژی:** پیش بینی قیمت و تقاضای درازمدت انرژی. تأثیر حوادث زیست محیطی بر انرژی ، خرابی

تجهیزات، مقررات و سایر متغیرها بر هزینه های خدمات انرژی.

✓ **خدمات مالی:** توسعه مدل های ریسک اعتباری، پیش بینی روند بازار مالی. پیش بینی تاثیر سیاست

ها، قوانین و مقررات جدید در مورد کسب و کار و بازار.

✓ **ساخت:** پیش بینی مکان و سرعت شکست دستگاه. بهینه سازی تحویل مواد خام بر اساس خواسته

های پیش بینی شده در آینده.

✓ **امنیت و پلیس:** از داده های روند جرم استفاده کنید تا محدوده هایی را تعیین کنید که ممکن

است در زمان های خاصی از حفاظت اضافی مورد نیاز باشد.

✓ **خرده فروشی:** مشتری آنلاین را به صورت بلادرنگ دنبال کنید تا تعیین کنید که آیا ارائه اطلاعات

مربوط به محصول یا انگیزه های اضافی احتمال انجام معامله را افزایش می دهد.

فصل سوم

تهدیدها و فرصتها

فرصت‌های پیش روی این فناوری

در این بخش گزارش به بررسی برخی از فرصت‌های موجود برای فناوری تجزیه تحلیل پیشرفته پرداخته خواهد شد این فرصت‌ها در واقع مزایای رقابتی برای بسیاری از سازمانها بوجود می آورد؛ در جدول بعدی به برخی از این فرصت‌ها اشاره شده است:

عنوان فرصت	توضیحات
وجود داده های فراوان - داده کاوی (جستجوی داده)	پیوست شماره یک
وجود متون فراوان - متن کاوی (جستجوی متن)	
وجود اینترنت - وب کاوی (جستجوی وب)	
امکان پیش بینی آینده	
اخذ تصمیم با دقت و سرعت بیشتر	
ارائه خدمات به افراد با شخصی سازی حداکثری	
پیش بینی عرضه و تقاضا در بازار	
تعمیر و نگهداری پیش بینانه تجهیزات	
پیش بینی رفتارهای خریداران و فروشندگان در بازارهای مالی	
کاهش هزینه ها	
امکان شناسایی تقلب و یا کشف اطلاعات نا منطبق	میتوان با استفاده از اطلاعات مجتمع شده توسط منابع مختلف و نیز در طول زمان، امکان کشفهای رفتارهای متقلبانه وجود دارد.
امکان حاکمیت بهتر برای دولتها و حکومتها	اجزای اصلی هر نوع حاکمیتی و در هر سطحی (حاکمیت در سطح شرکت و جامعه) به شرح ذیل است:

<ul style="list-style-type: none"> • راهبردها و استراتژیها – بخش مهمی از ابزارهای حاکمیتی هستند که میتوانند توسط پیش بینی نیازها و نیز محورهای توسعه تدوین شوند. • نظارت – بخش دوم از حاکمیت نظارت بر روند اجرای استراتژیها میباشد. که با تجزیه و تحلیل داده ها محقق خواهد شد. • کنترل - بخش سوم از حاکمیت قدرت کنترل میباشد. باید کنترل شود که چرا راهبردها و استراتژیها درست اجرا نشده و فاصله وضعیت فعلی با وضعیت مطلوب اندازه گیری شود. برای فهمین این که چرا اقدامات درست اجرا شده اند و یا اینکه چرا درست اجرا نشده با تجزیه و تحلیل و مقایسه تصمیمات حاصل میشود. برای تجویز اقدامات اصلاحی باید تصمیم گیری نمود. 	
<p>با ایجاد امکان سنجش نیازهای مشتری و کسب رضایت وی از طریق تجزیه و تحلیل پیشرفته اطلاعات، امکان شناخت نیازمندیهای آتی مشتریان بسیار راحتتر خواهد بود لذا امکان تعریف سرویس و خدمات جدید برای مشتریان فراهم میگردد.</p>	<p>امکان ارائه خدمات و سرویسهای جدید به مشتریان</p>
<p>امکان پیش بینی و کنترل شیوع بیماریها امکان پیش بینی و کنترل وضعیت بیماران امکان پیش بینی</p>	<p>کمک به صنعت بهداشت و درمان</p>

تهدیدات پیش روی این فناوری

در حالی که تجزیه و تحلیل اطلاعات پیشرفته، هم برای کاهش هزینه و افزایش درآمد مناسب می باشد اما نمیتوان به اهداف فوق رسید مگر اینکه یک ساز و کار مناسب جهت استفاده از این فناوری در سازمان وجود داشته باشد. در ادامه به برخی از مشکلات و چالشهای این فناوری اشاره شده است:

عنوان تهدید / نقطه ضعف	توضیحات
یک نقطه ضعف مرکزی	تجمیع اطلاعات در انبارهای بسیار بزرگ ریسکهای بزرگی نیز به دنبال دارد. هک کردن سیستم های تصمیم گیری و هوشمند شدن، سرقت اطلاعات، یادگیری های غلط برای سیستمهای یادگیر تصمیم گیر میتوانند مشکلات بسیار عظیمی را به وجود بیاورند.
مدیریت داده های	مدیریت داده های کلان شامل ذخیره سازی، آرشیو و نحوه استفاده از آنها خود یک چالش بزرگ برای سازمانها است.
صحت داده های ورودی که معمولاً حجم بسیار زیادی دارد.	معمولاً کیفیت داده های ورودی در خروجیها تاثیر بسزایی دارد لذا اطمینان از اینکه داده های ورودی با کیفیت باشند از نقاط ضعف این فناوری می باشد.
اطمینان از نتایج تجزیه و تحلیل پیشرفته	هر سازمانی به منظور توسعه بینشهای خود از فناوری تحلیل داده استفاده میکند. در صورتیکه نتایج تجزیه و تحلیل داده قابل اتکا نباشد و یا بدتر از آن برخی مواقع درست و

<p>برخی مواقع نادرست باشد. به عبارت دیگر نتایج تحلیلی ناقص باشد میتوان چالشهای بزرگی برای استفاده کنندگان فراهم آورد.</p>	
<p>پیوست شماره دوم</p>	<p>نداشتن مشتری</p>
	<p>پیاده سازی تجزیه و تحلیل پیشرفته</p>
	<p>مقررات و حفظ حریم خصوصی – این بزرگترین چالش خواهد بود.</p>
	<p>مالکیت داده در صنعت کلان داده ها و تجزیه و تحلیل</p>
	<p>مشکلات ناشی از فناوری</p>
	<p>فراهم سازی و به اشتراک گذاری داده ها بین سازمانها</p>
	<p>خارج از کنترل شدن هوش مصنوعی</p>
<p>با توجه به اینکه منابع تولید داده مختلف بوده و معمولاً در سازمانها، ارگانها و با روشهای مختلفی تولید میشوند لذا تجمیع این اطلاعات گوناگون و یکدست کردن آنها از مشکلات و تهدیدات بزرگ است. البته مساله دیگر این است انتقال اطلاعات از سازمانها به یکدیگر از مشکلات بزرگ محسوب میشود.</p>	<p>تجمیع اطلاعات از منابع مختلف</p>

فصل چهارم

معرفی زیست بوم

تجزیه و تحلیل پیشرفته اطلاعات

۱. مقدمه

زیست بوم به مجموعه ای از اکوسیستم ها اطلاق میشود که اجزای مختلفی تشکیل یافته است. این اجزاء بر یکدیگر اثر متقابل دارند. برخی از عناصر در یک زیست بوم مهمتر از بقیه عناصر بوده اما تمامی عناصر مفید بوده و نبود یکی از آنها باعث گسست در زنجیره تامین و ارتباطی اجزای زیست بوم در کوتاه مدت یا دراز مدت خواهد شد. یکی از ویژگیهای زیست بوم تعدد اکوسیستمها و نیز تعداد بسیار زیاد اجزای آن زیست بوم میباشد.

با توجه به مفهوم کلی که در این گزارش پی گرفته ایم. اغراق نیست اگر از واژه زیست بوم برای فناوری تحلیل داده ها استفاده کنیم. مفهوم داده و اطلاعات آنقدر کلی و فراگیر است که براحتی میتوان برای آن واژه زیست بوم را بکار برد.

داده ها در همه جای زندگی بشر حضور دارد از ساعت و گوشی و عینک گرفته تا تلویزیون لپ تاپ و دوربینهای راهنمایی رانندگی، دوربینهای امنیتی، خودروها، بانکها، بیمه ها، پزشک و داروخانه، آزمایشگاه و بیمارستان، هواشناسی و زنجیره تامین، تبلیغات و خرید و فروش و تجارت، سیاست و شبکه های اجتماعی و تولید دانش، همه و همه با داده ها و اطلاعات کار میکنند.

جمع آوری اطلاعات، ذخیره سازی، پایش و پالایش اطلاعات، و تجزیه و تحلیل آنها همگی نیاز به قوانین و مقررات، ابزارها و نرم افزارهای خاص و نیز الگوریتمهای و فناوریهای متعددی دارند. مجموعه این ابزارها و اطلاعات و تکنیک و روشها جملگی یک زیست بوم داده را تشکیل میدهند.

۲. ویژگیهای اصلی زیست بوم تجزیه و تحلیل داده

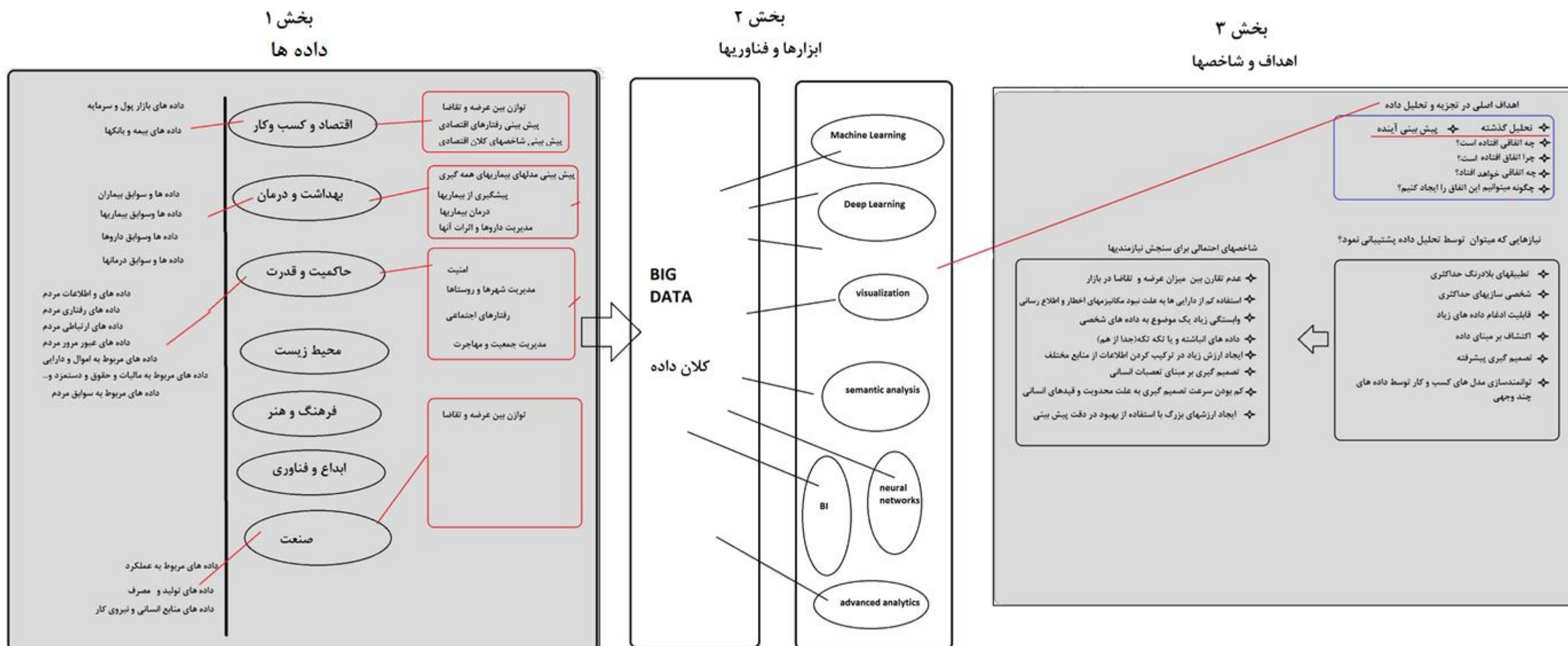
- زیست بوم داده در همه جای زندگی بشر حضور دارد.
- این زیست بوم چرخه عمر داده ها را به طور کامل پشتیبانی میکند. تولید داده، ذخیره سازی داده ها، پالایش و جلا دادن داده، تجزیه و تحلیل داده ها و تولید بینش و آینده پژوهی از داده ها و در نهایت آرشیو و حذف داده های تاریخ مصرف گذشته
- داده های موجود در این زیست بوم تنوع بسیار زیادی دارد.
- حجم اطلاعات در این زیست بوم میتواند به طور شگفت آوری زیاد باشد.
- نیاز به انواع تخصصها دارد، به منظور استفاده از داده ها به طور موثر و با هدف آینده نگری نیاز به تخصصها و رشته های مختلفی از علم وجود دارد. تنها مهندسان فناوری اطلاعات نمیتوانند از داده ها استفاده کنند.
- پیچیدگی فنی بسیار زیادی در تکنولوژیهای مرتبط با تحلیل داده وجود دارد اعم فناوریهای هوش مصنوعی، یادگیری ماشین، یادگیری عمیق و ...

۳. عناصر اصلی تشکیل دهنده زیست بوم

- عنصر اول - که مهمترین عنصر در این زیست بوم میباشد، داده ها و اطلاعات میباشند.
- عنصر دوم - الگوریتمها و فناوریهای تجزیه و تحلیل داده ها میباشند.
- عنصر سوم - کاربردها و نیز ابزارهای نمایش اطلاعات و دانشها میباشند.

۴. ساختار کلی زیست بوم

در شکل بعدی به ساختار کلی زیست بوم مربوط به تجزیه و تحلیل و اطلاعات پرداخته شده است. در این زیست بوم سه بخش اصلی که شامل فضای تولید داده ها در حوزه های مختلف، فضای ابزارها و فناوریها و فضای مربوط به اهداف و کاربردها وجود دارد.



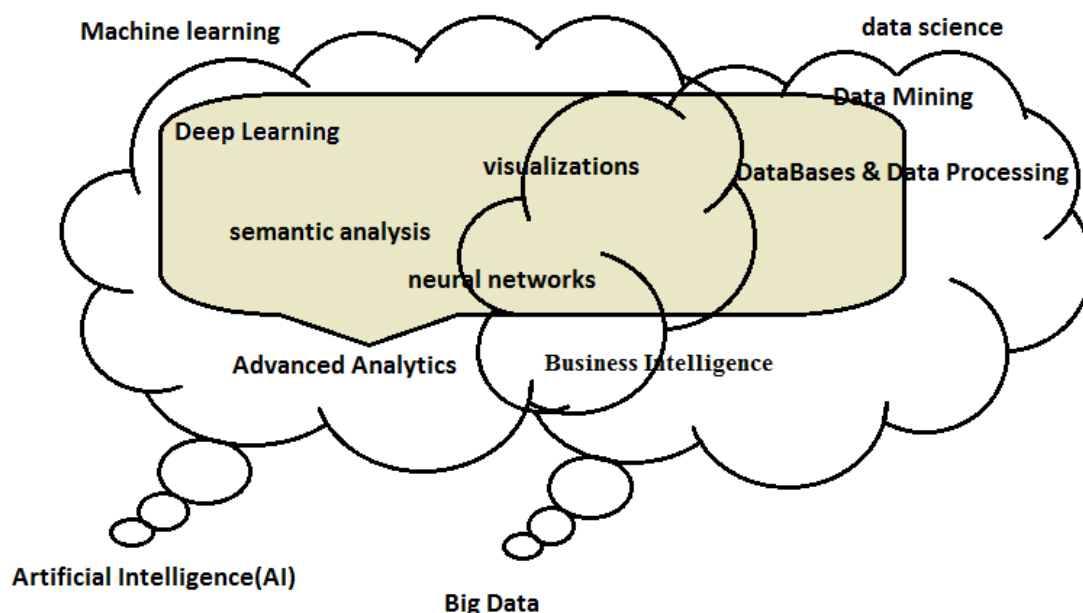
شکل فوق در واقع نشان دهنده یک اکوسیستم کلان از نحوه کارکرد و نیز تحقق اهداف و شاخصهای مربوط به فناوری تحلیل داده پیشرفته را نشان میدهد. در ادامه به توضیح هر یک از بخشهای تصویر فوق پرداخته میشود:

بخش 1: حوزه های مختلف تولید داده : هدف از این بخش موارد ذیل است :

- انواع کلان داده هایی که در بخشهای مختلف از زندگی بشر نقش دارند توضیح داده شده است. به عنوان مثال در بخش بهداشت و درمان داده های مربوط بیماران، بیماریها، داروها و نیز درمان و تمامی تاریخچه ها نشان داده شده است.
- معرفی جنبه های مختلف زندگی بشر که به صورت مستمر داده های زیاد تولید میکنند نمایش داده شده است. مثلاً داده های مربوط به اقتصاد، بهداشت و درمان، داده های مربوط به مردم و ...

بخش 2: روشها و فناوریهای مختلف تجزیه و تحلیل : هدف از این بخش موارد ذیل است :

همانگونه در شکل فوق نشان داده شده است ابزارها، فناوری و دانشهای متعددی در خصوص تجزیه، تحلیل، غنی سازی، تبدیل اطلاعات از یک فرمت به فرمتهای دیگر وجود دارد. شکل زیر در واقع نشان دهنده یک نمای دیگر از روابط کلی بین فناوریها و دانشهای عمده در حوزه کلان داده میباشد.



در بخش سمت راست تصویر آنچنان که نشان داده شده است **Big Data** (کلان داده) است که باعث بوجود آمدن علم داده (**data science**) شده است. دور از ذهن نیست که در زیر مجموعه علم داده میتوان تکنیکها و مفاهیمی چون داده کاوی، پایگاه داده و هوش تجاری و تحلیل پیشرفته اطلاعات را مشاهده نمود. در بخش چپ

تصویر مفهوم هوش مصنوعی نشان داده شده است که در واقع یک چتر است که مفاهیم دیگری چون یادگیری ماشین، یادگیری عمیق، شبکه های عصبی و تحلیلهای معنایی را در بر میگیرد. لازم به ذکر است که مفاهیم کلان داده و هوش مصنوعی در واقع در ترکیب با یکدیگر امکان تولید بینشهای عمیق از اطلاعات را بوجود می آورند. که در این گزارش از آن به عنوان تجزیه و تحلیل پیشرفته یاد میکنیم.

این گزارش استفاده از یادگیری ماشین و هوش مصنوعی در تحلیل اطلاعات را از دو زاویه مورد بررسی قرار داده است. اولاً، تحقیق در مورد اینکه آیا میتوان از این روش در کدام صنایع میتوان استفاده نمود؟ دوم، بررسی نمودن اینکه که فعالیت های کاری فعلی که توسط افراد انجام می شود، به طور بالقوه می توانند از طریق یادگیری ماشین ها خودکار شوند و چگونه می توانند در دیگر مشاغل نیز همین اتفاق تکرار شود. شکل زیر نشان میدهد که با استفاده از تحلیلهای یادگیری ماشین میتوان کدام موارد کاربرد را انجام داد:



داده ها و الگوریتم های تحلیل پیشرفته میتوانند بروی کیفیت و کمیت تصمیم گیری انسان اثرگذار باشند.

انسانها معمولاً بر اساس حالات روحی و روانی (شرایط)، تعصبات و همچنین عدم توانایی در تحلیل اطلاعات زیاد تصمیم گیری میکنند.

استفاده از داده های زیاد و همچنین الگوریتمهای موثر جهت تحلیل آنها میتواند حصارها و محدودیتهای انسانی را کنار زده و یک تصمیم سازی شفاف، سریع، دقیق و البته سازگار با شرایط را فراهم آورد.

نمونه ها و مثالهای زیادی از نحوه بهینه سازی تصمیمات با استفاده از تحلیل داده در کاربردها و صنایع مختلف وجود دارد از جمله این کاربردها میتوان به مباحث شهر هوشمند اشاره نمود.

یکی از زمینه های دلایل اصلی استفاده از توانایی ماشین ها و الگوریتم ها، امکان پردازش مقادیر عظیمی از اطلاعات در کسری ثانیه است. کاری که هیچ انسانی قادر به انجام آن در این زمان کوتاه نیست.

استفاده از سنسورها برای بهبود جریان های ترافیکی سطح شهر و و نیز استفاده از اینترنت اشیاء برای مدیریت و بهینه سازی مصرف آب و برق در شهرها که به نگهداری و توسعه سیستم های زیربنایی شهرها کمک میکنند. تنها دو مورد از کاربردهای تحلیل اطلاعات در شهرهای هوشمند است. یکی از کاربردهای اصلی در تحلیل داده ها، پیشگیری از رخداد خطاهای پزشکی میباشد.

ابزارهای تحلیل پیشرفته می توانند حساسیت بالقوه و یا تعاملات خطرناک دارویی را برای پزشکان و داروسازان به طور کاملاً نمایان نموده و لذا اطمینان حاصل شود که تصمیمات آنها سازگار و قابل اطمینان است.

و سرانجام، شاید هیچ یک از حوزه های تصمیم گیری انسانی مانند موضوع استخدام کارمند جدید شامل اطلاعات مبهم و پیچیده نباشد. لذا تجزیه و تحلیل پیشرفته داده ها می تواند به ایجاد یک بازار کار شفاف تر با تحلیل اطلاعات (کارفرمایان، مشاغل، جویندگان کار مهارتهای موجود و مورد نیاز، اطلاعات مربوط به عرضه و تقاضا برای مهارت های خاص، دستمزدها و ...) کمک نماید.

بخش 3:

بخش سوم از این اکوسیستم در واقع یکی از مهمترین بخشها میباشد که مرتبط با اهداف، شاخصها و نیازمندیها تجزیه و تحلیل داده میباشد. در این بخش از اکوسیستم شاخصهای احتمالی نیازمندیها تعریف میشوند. نوع نمایش اطلاعات و نوع تصمیم گیریها و تنظیمات مرتبط با مدلهای تصمیم گیری و پیش بینی در این بخش انجام میشود.

۵. ذینفعان تجزیه و تحلیل پیشرفته

در ادامه به لیست ذینفعان مربوط به فناوری تجزیه و تحلیل داده پرداخته میشود. این فناوری بطور کلی دارای سه نوع ذینفع به شرح ذیل است:

- نوع اول: ذینفعان فنی و تخصصی: افرادی که زیرساختهای فنی و مفهومی مربوط به تجزیه تحلیل داده را فراهم می آورند.
- نوع دوم: ذینفعان سطح بالا و استفاده کننده که غالباً مدیران ارشد و استراتژیک سازمان هستند و از گزارشات و پیش بینیهایی این فناوری استفاده میکنند.
- نوع سوم: ذینفعانی نهایی سیستم هستند که پیش بینی های سیستم بروی آنها اثر گذار هستند.

گروه بندی ذینفعان	نام ذینفع	StakeHolder	توضیحات
ذینفعان فنی و تخصصی	آمارگیر	Actuary	
	مدیر تحلیل	Analytics Manager/Director	
	متخصص تجزیه و تحلیل	Analytics Specialist	
	توسعه دهنده هوش تجاری	BI developer	
	برنامه نویس کلان داده	Big Data Programmer	
	دانشمندان بیوانفورماتیک	Bioinformatics Scientist	
	متخصص آمار	Biostatistician	
	تحلیلگران کسب و کار	Business analysts	
	مشاوران هوش تجاری	Business Intelligence Consultant	
	ریاضیدان	Mathematician	
	مهندس یادگیری ماشین	Machine Learning Engineer	
	تحلیلگر سیستم	System Analyst	
	تحلیلگر وب	Web Analyst	
	معماران راه حل کسب و کار	Business Solution Architects	

	Database Administrator	مدیر پایگاه داده	
	Data scientist	دانشمند داده	
	Statistician	آمارشناس	
	Chief Technology Office (CTO)	مدیر ارشد فناوری	
	Client project team	تیم پروژه مشتری	
	Customer Service	خدمات مشتریان	
	Customers	مشتریان	
	Data architect	معمار داده	
	Systems Integrators	مهندسان یکپارچه سازی سیستمها	
	Data Analyst	تحلیلگر داده	
			دینفعان سطح بالا
	Chief Executive (CEO)	مدیر عامل (مدیر اجرایی)	
	Data integration developers	برنامه نویسان مجتمع سازی داده ها	
	Data Warehouse	انبار داده	
	Campaign Experts	متخصصان کمپین	
	Marketing Manager	مدیران بازاریابی و توسعه کسب و کار	
	Financial Analyst	تحلیلگر مالی	
	Chief Information Officer (CIO)	مدیر ارشد اطلاعاتی - مدیر IT	
	Planning Analyst	تحلیلگر برنامه ریزی	
	Risk Analytics	تحلیلگران ریسک	
	Market Research Analyst	تحلیلگر بازار	
	Marketing	بازاریابی	

	Suppliers	تامین کنندگان	ذینفعان نهایی
	People	مردم	
	Goverment	دولت	
	Business SMEs	شرکتهای کوچک و متوسط بازاریابی	
	Competitors	رقبای شرکتهای	

۶. نحوه ایجاد ارزش در این زیست بوم

اخذ تصمیمات با سرعت و دقت بیشتر از مهمترین ارزشهایی است که تجزیه و تحلیل پیشرفته اطلاعات برای سازمانها و مدیران ایجاد میکند. به منظور درک نحوه تولید ارزش به این سوال بپردازیم که یک مدیر در سازمان چگونه تصمیم میگیرد؟ مراحل مختلف تصمیم گیری به شرح ذیل است:

۱. شناسایی تصمیم
۲. جمع آوری داده های مرتبط با تصمیم
۳. مشخص کردن گزینه های تصمیم گیری
۴. تجزیه و تحلیل شواهد مربوط به هر یک از گزینه ها
۵. انتخاب یکی از گزینه ها
۶. انجام کار

فناوری تجزیه و تحلیل پیشرفته میتواند برخی از مراحل فوق را خودکار سازی نماید. خودکار سازی برخی از مراحل فوق میتواند به کاهش هزینه های تصمیم گیری و افزایش دقت تصمیمگیری منجر شود. بگذارید یک تصمیم را تعریف کنیم: "تصمیم گیری در خصوص پیگیری تسهیلات معوق شده مشتریان در بانکها" برای باید مراحل زیر انجام شود:

۱. تعریف تصمیم: پیگیری کردن یا نکردن مشتریان جهت وصول مطالبات
۲. جمع آوری داده های مربوط به مشتریان و نیز داده های مربوط به پرداخت تسهیلات و نیز وثایق آنها
۳. سپس مشتریان به دسته بندیهای بسیار سبز، زرد، قرمز تقسیم میشوند.
۴. طبقه بندی مشتریان و امتیاز دهی به آنها بر اساس داده های موجود. از سبزها میتوان مطالبات را گرفت، از قرمزها نمیتوان مطالبات را اخذ کرد. زردها مشکوک و الوصول هستند
۵. انتخاب یکی از گزینه های فوق جهت پیگیری یا عدم جهت اخذ مطالبات.
۶. انجام کار

در فرآیند فوق میتوان بخشهای 2 تا 4 را کاملاً با استفاده فناوری تجزیه و تحلیل داده خودکار سازی نمود.

ابزارهای هوش مصنوعی و یادگیری ماشین با کمک تحلیل و تجزیه اطلاعات می تواند به حل مسائل در سه سطح زیر کمک کند :

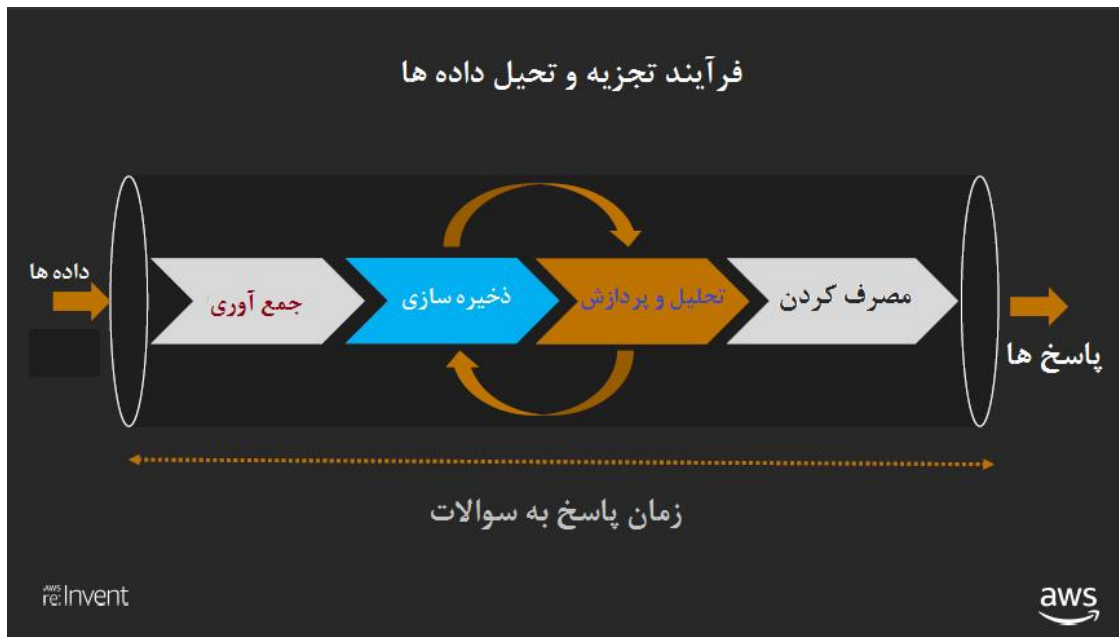
- **سطح اول :** طبقه بندی داده ها و اطلاعات
- **سطح دوم :** پیش بینی نتایج و رفتارها بر اساس اطلاعات
- **سطح سوم :** تولید رفتارها و نتایج جدید بر اساس الگوهای رفتاری قبلی

در شکل بعدی به برخی از کاربردهای مختلف – که میتواند با یادگیری ماشین انجام شود – پرداخته شده است:

طبقه بندی	طبقه بندی / برچسب محتواها بصری	شناسایی اشیاء، چهره در تصاویر و ویدئو
	طبقه بندی / برچسب نوشتار و متن	شناسایی حروف، نمادها، کلمات در نوشتار و متون
	طبقه بندی / برچسب صوتی	طبقه بندی و برچسب گذاری اصوات در فایل‌های صوتی
	خوشه ، گروه بندی سایر داده ها	بخش بندی اطلاعات (مثلا مشتریان، ویژگیهای محصول) به دسته ها، خوشه ها
پیش بینی	کشف ارتباطات	مشخص کردن اینکه افرادی که یک شبکه تلویزیونی خاص تماشا میکنند یک کتاب خاص را نیز مطالعه میکنند
	پیش بینی احتمالی نتایج	پیش بینی احتمال اینکه مشتری یک ارائه دهنده خدمات دیگر را انتخاب کند
	پیش بینی تخمین تابع ارزش	یادگیری روی داده های گذشته، پیش بینی تقاضا برای یک محصول
تولید	تولید محتواهای بصری	یادگیری از مجموعه ای از نقاشی های هنرمندان و طراحی یک نقاشی جدید در همان سبک
	تولید نوشتار و متن	یادگیری در خصوص طراحی نوشتار و تکمیل قسمتهای گم شده یک متن
	تولید صوت	طراحی یک صوت جدید در همان سبک / ژانر
	تولید سایر داده ها	یادگیری با استفاده داده های مربوط به اب و هوا در کشورهای مختلف و پر کردن اطلاعات مربوط به کشورهای با داده های کم و ناقص

۷. روش و معماری تجزیه و تحلیل اطلاعات

در شکل بعدی یک فرآیند کلی به منظور تجزیه و تحلیل اطلاعات آورده شده است:



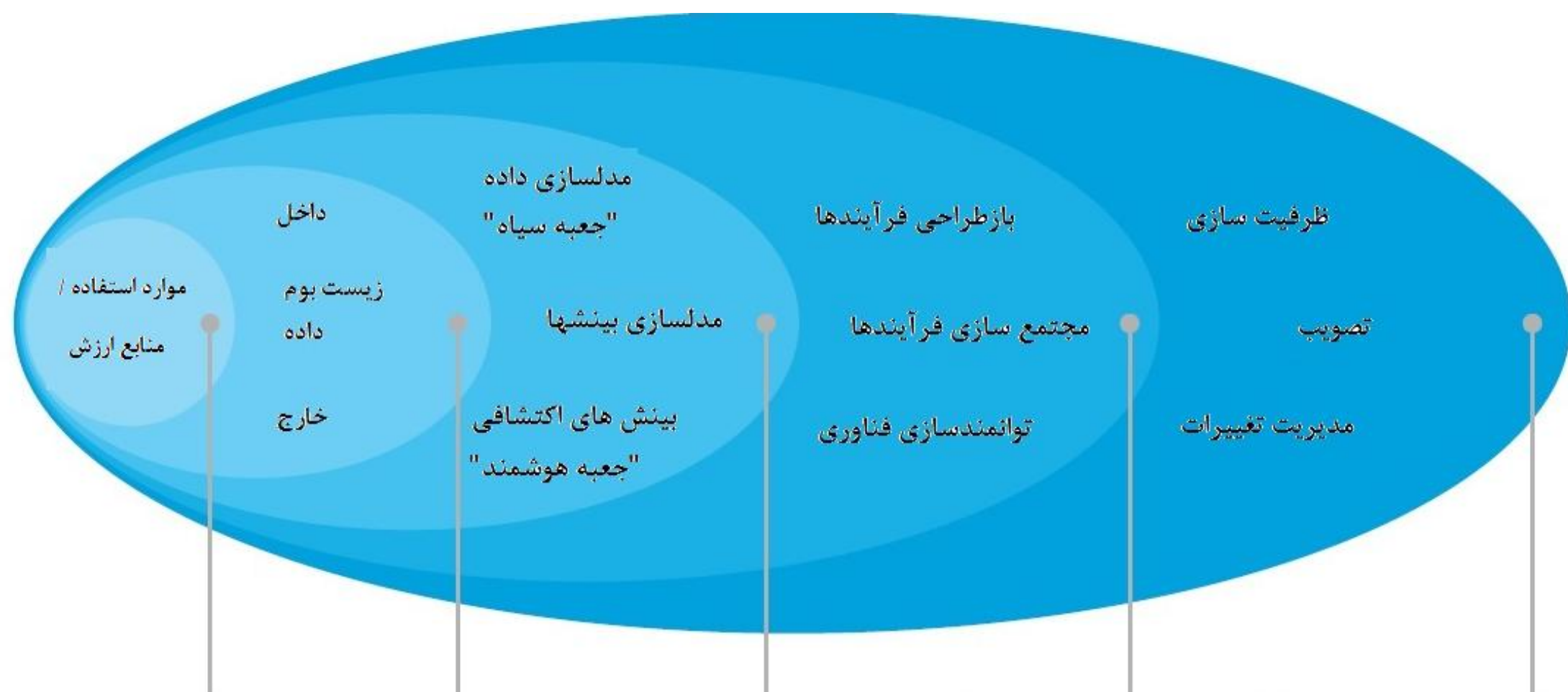
برای آگاهی از معماری جزئی برای این زیست بوم به همراه توضیحات به پیوست شماره سه مراجعه نمایید.

۸. حوزه های کلان کاربرد در صنایع مختلف

در ادامه به ذکر برخی موارد کلان از کاربردهای مربوط به تجزیه و تحلیل پیشرفته اطلاعات پرداخته خواهد شد. بطور کلی علم تجزیه و تحلیل داده با اهداف کلانی به شرح ذیل پیوند یافته است:

- انبار بزرگ از داده ها و اطلاعات بسیار زیاد که میتوان از آن استفاده کرد. این انبار همواره از منابع اطلاعاتی مختلف تغذیه شده و اطلاعات ناهمگون را دریافت ، پالایش و ذخیره میکند.
- از این کلان داده ها میتوان برای استخراج بینش و اکتشاف الگوها و روابط پنهان بین اطلاعات که ناشی از رفتارهای طبیعی ، غیر طبیعی، ارادی و غیر ارادی هر منبع تولید کننده داده استفاده برد.
- از این اطلاعات میتوان برای پیش بینی و بازطراحی فرآیندها و رفتارهای منابع تولید کننده داده استفاده نمود.

- از این اطلاعات میتوان برای مدیریت و کنترل پیشگیرانه و به طور کلی مدیریت تغییرات در اشیاء و یا منابع تولید داده استفاده نمود.
- از این بینشها میتوان جهت خودکار سازی و بهبود فرآیند تصمیم گیری استفاده نمود.
- از این بینشها میتوان جهت پیش بینی رخدادهای آینده استفاده نمود.



- ایجاد قابلیت های مدیریتی و ارتباط با مشتری
- بازطراحی فرآیندها
- عمل مدل سازی خطی و غیر خطی
- جمع آوری اطلاعات از سیستمهای داخلی و منابع خارجی
- توضیحاتی درباره چگونگی استفاده کسب و کار از راه حل را بیان می کند.
- ایجاد یک رابط کاربری گرافیکی که با گردش کارهای روزانه یکپارچه شده است
- گردش کارهای خودکار
- برای به دست آوردن بینش جدید
- کدگذاری و آزمایشات اکتشاف در سراسر فزودن داده های کلیدی خارجی
- نیازهای کسب و کار را شفاف نموده و اثرات پروژه را پیش بینی می کند
- سازمان (متغیرهای پیش بینی کننده) ایجاد محیط آزمایشی تحلیل داده
- افزایش داده ها (ایجاد متغیرهای پیش بینی جدید)
- مدیریت تغییرات پیشگیرانه و پیگیری شاخصهای عملکردی مصوب

در ادامه موارد کاربردی کلی که در بخش قبل عنوان شده به شش الگوی کاربرد کلان اشاره خواهد شد که هر یک از آنها میتوانند در تعدادی از صنایع مختلف مفیدتر واقع شود پرداخته خواهد شد این موارد در جدول بعدی به شرح ذیل توضیح داده خواهد شد :

ستون اول - شاخصهای احتمالی برای سنجش الگوها : در واقع این ستون نشان میدهد که از تجزیه و تحلیل داده در کدام حوزه کلان زندگی میتوان استفاده نمود.

ستون دوم - الگوهای کلان کاربرد : نشان دهنده موضوعات کلانی است که ممکن است در هر صنعت کاربرد و مصداق پیدا کنند.

ستون سوم - صنعت مرتبط : نشان دهنده صناعی است که میتواند از الگوی کلان مد نظر برای بهبود کسب و کار، بهبود عملکرد و یا حل کردن مشکلات موجود در آن صنعت استفاده نمایند.

شاخصهای احتمالی برای سنجش الگوهای کلان کاربرد	الگوهای کلان کاربرد	صنعت مرتبط
<ul style="list-style-type: none"> • استفاده کم از دارایی ها به علت نبود مکانیزمهای اخطار و اطلاع رسانی • عدم تقارن بین میزان عرضه و تقاضا • وابستگی زیاد به داده های شخصی • داده های انباشته و یا تکه تکه (جدا از هم) • ایجاد ارزش زیاد در ترکیب کردن اطلاعات از منابع مختلف • تحقیق و توسعه هسته اصلی کسب و کار است. • تصمیم گیری بر مبنای پیش تعصبات انسانی • کم بودن سرعت تصمیم گیری به علت محدودیت و قیدهای انسانی • وابستگی بودن ایجاد ارزشهای بزرگ با استفاده از بهبود در دقت پیش بینی 	توانمندسازی مدل های کسب و کار توسط داده های چند وجهی	بیمه مراقبتهای پزشکی و بهداشتی سرمایه انسانی و استعداد
	تطبیقهای بلادرنج حداکثری	حمل و نقل و تدارکات
		خوکارسازی
		شهرهای هوشمند و زیرساختها
	شخصی سازیهای حداکثری	مراقبتهای پزشکی و بهداشتی
		خرده فروشی
		رسانه
		آموزش
	قابلیت ادغام داده های زیاد	بانکداری
		بیمه
		بخشهای خدمات عمومی
		سرمایه انسانی و استعداد
	اکتشاف بر مبنای داده	علوم زیستی و داروسازی
		علوم مواد
		فن آوری
	تصمیم گیری پیشرفته	شهرهای هوشمند
		مراقبت های بهداشتی

بیمه		
سرمایه انسانی و استعداد		

۹. مروری بر وضعیت بازار تجزیه و تحلیل پیشرفته اطلاعات

تجزیه و تحلیل پیشرفته اطلاعات آخرین انقلاب تکنولوژیک است که دنیای کسب و کار را به سمت درکی عمیق تر و واقعی تر از کسب و کار و رفتار مشتریان هدایت کرده است. اگر ما از تمام پتانسیل این فناوری استفاده کنیم میتوانیم در بازار رقابتی و جهانی تغییر عمده ایجاد کند.

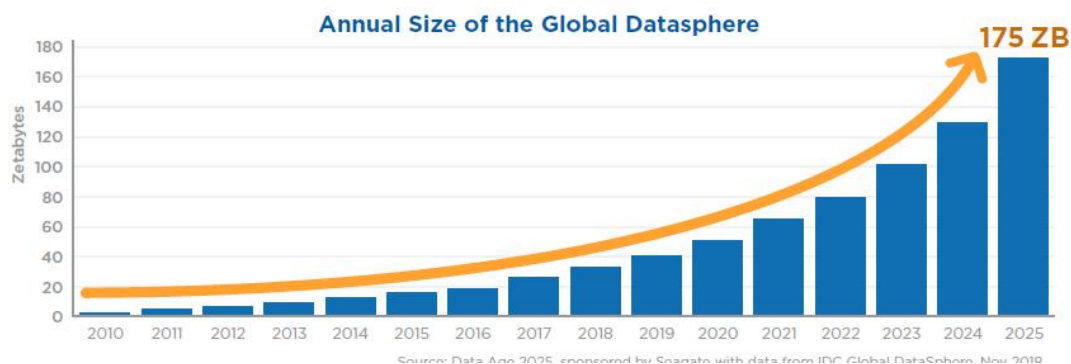
برخی از شرایطی که در آینده ممکن است به واسطه تجزیه و تحلیل پیشرفته در جهان ایجاد گردد:

- مراکز بینش ممکن است عنصر جدید در سازمان ها باشد.
- مجتمع های دانشگاهی به همراه صنعت میتوانند وارد بازار تجزیه و تحلیل اطلاعات شوند.
- ترکیب شدن دو رویکرد بلادرنگ و تحلیل فله ایی اطلاعات با یکدیگر
- در سال 2018، شبکه های عصبی عمیق (DNN) یک بخش استاندارد از مجموعه ابزار Data Science در 80 درصد موارد خواهد بود.
- در سال 2019، دانشمندان داده توجه به تجزیه و تحلیل اطلاعات مربوط به مسائل شهروندی بسیار بیشتر از دیگر انواع داده ها خواهد شد.
- در سال 2019، نسل زبانی طبیعی (NLG) در 90 درصد از راه حل های تجزیه و تحلیل و تجزیه و تحلیل BI به یک ویژگی استاندارد تبدیل خواهد شد. همچنین، 50 درصد از نمایشگاه های تجزیه و تحلیل از جستجو، پرس و جو یا صدای NL یا از منابع خودکار می آیند.
- در سال 2019، امکان تولید زبان طبیعی (NLG) تبدیل به یک ویژگی استاندارد در 90 درصد از تجزیه و تحلیل و راه حل های BI خواهد شد. همچنین، 50 درصد از پرس و جوهای تجزیه و تحلیل اطلاعات از طریق پرس و جو صوتی NL و یا از طریق خودکار انجام میشود.
- آینده تجزیه و تحلیل کسب و کار با مباحثی مانند یادگیری ماشین و یادگیری عمیق کاملاً گره خورده است. به عبارت دیگر تجزیه و تحلیل پیشرفته اطلاعات با استفاده از AI, DL وجه دیگری خواهد یافت.

- ارائه دهندگان خدمات تجزیه و تحلیل پیشرفته به طور فزاینده ای به سمت خدمات مبتنی ابر تبدیل خواهند رفت تا خدمات مذکور را به شیوه ای مقرون به صرفه تر ارائه دهند، و آنها برای شرکتهای کوچک و متوسط نیز توجیه پذیر نموده و هزینه های مربوط را برای شرکتهای بزرگ نیز کاهش دهند.
- شرکت ها به سمت استفاده از زبان نشانه گذاری مدل پیش بینی (PMML) حرکت خواهند کرد: باید توجه کرد که شرکت ها از تجزیه و تحلیل دسته ای اطلاعات به سمت استفاده از جریان داده های بلادرنگ خواهند رفت و به طور مداوم به بهبود دقت مدل های خود، با استفاده از زبان پیش فرض مدل نشانه گذاری (PMML) خواهند پرداخت. PMML یک استاندارد برای مدل آماری و داده کاوی است که توسط گروه داده (Data Mining (DMG تهیه شده است.
- یک کنسرسیوم مستقل به رهبری فروشندگان IBM و SAS عضو کامل DMG هستند، در حالی که SAP، StatSoft، RapidMiner و Angoss در توسعه PMML کمک کردند.
- زبان برنامه نویسی python یک راه ورود خواهد بود برای محقق ساختن نیازهای عمومی برنامه نویسان تجزیه و تحلیل پیشرفته اطلاعات
- توسعه رابطهای گرافیکی کاربر پسند، امکان تجزیه و تحلیل اطلاعات توسط افراد بیشتری را با تخصصهای کمتر فراهم می آورد.
- جریان های داده بلادرنگ و اینترنت اشیاء بسیار پرتعداد خواهند شد.
- نمایه سازی (مصور سازی) داده به یک نیاز عمومی تبدیل خواهد شد.
- نمایه سازی داده ها نقش بسیار مهمی در سازمان ها ایفا می کند، زیرا آنها با استفاده از جریان داده ها، داده های رسانه های اجتماعی، داده های ماشین و سایر حجم های زیاد داده های ساختار یافته، نیمه ساختار یافته و غیر ساختاریافته تغذیه می شوند. تصویرها به تحلیلگران کمک میکنند تا بینش هایی را کشف کنند که ممکن است درک آنها، هیچگاه در یک قالب جداول داده، صفحات گسترده اکسل و نمودارها ممکن نباشد.

- سازمانها از تجزیه و تحلیل پیشرفته اطلاعات به منظور اخذ تصمیم گیری در تمامی فعالیت استفاده گسترده خواهند نمود.
- نرخ رشد کلان داده ها بر اساس پیش بینی موسسه IDC(International Data Corporation) در کل جهان در سال 2018 معادل 33 Zettabytes میباشد. حجم اطلاعات در سال 2025 به عدد 175 Zettabytes خواهد رسید. یعنی حجم کل داده های جهان در حدود 5.3 برابر خواهد شد. به منظور

Figure 1 - Annual Size of the Global Datasphere



درک بزرگی این تبدیلهای زیر توجه فرمایید :

1 ZettaByte= 1,000,000,000 TeraByte

1 ZettaByte = 1,000,000,000,000 Giga Byte

✓ اگر شما قادر به ذخیره کل داده جهان در سال 2025 در دی وی دی باشید، میتوانید با روی

هم گذاشت DVD ها 23 بار به کره ماه برسید و یا 222 بار کل دور زمین بچرخید!

✓ اگر یک نفر بخواهد 175 زتا بایت را با سرعت طور متوسط 25 مگابایت بر ثانیه - سرعت

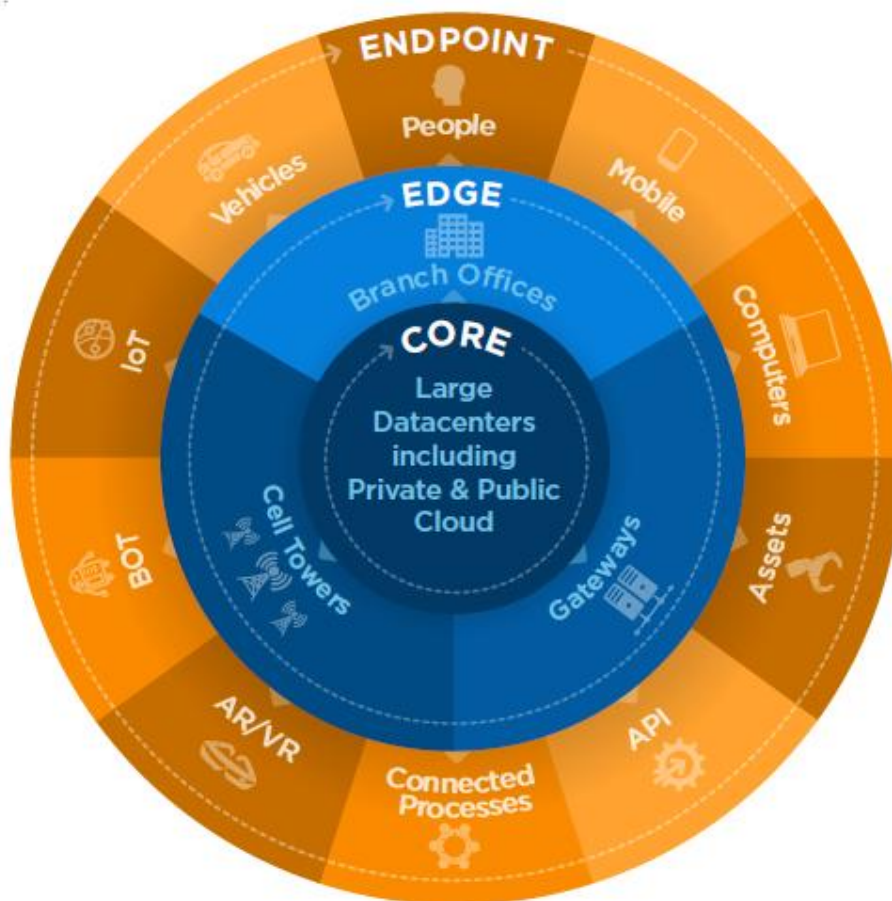
متوسط اینترنت امروز در سراسر ایالات متحده - دانلود کند، نیاز به 1.8 میلیارد سال زمان

خواهد داشت

✓ اگر تمام افراد جهان بخواهند همزمان این مقدار را با سرعت 25 مگابایت در ثانیه دانلود کنند

81 روز طول خواهد کشید.

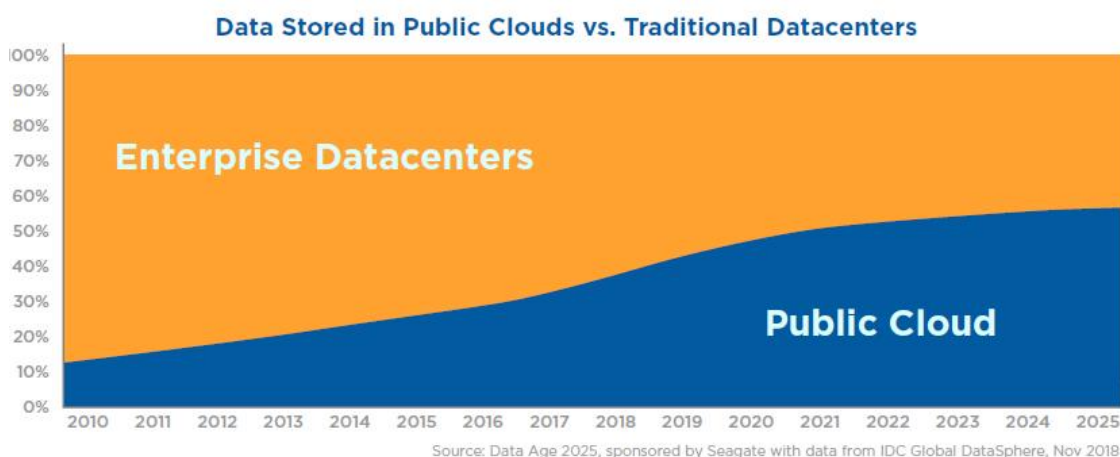
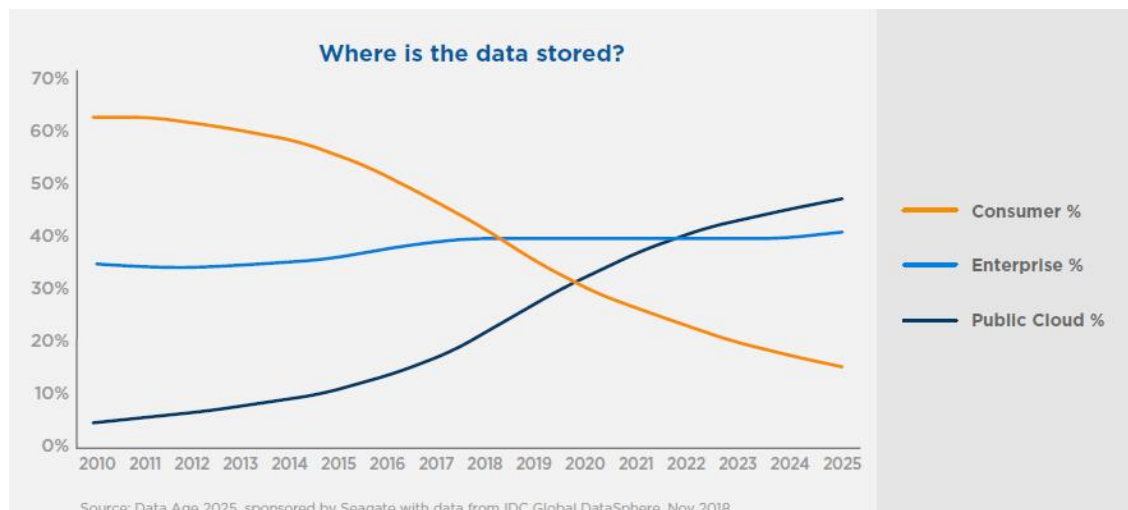
- ساختار کل داده های جهان در سه سطح به شکل زیر خواهد بود:



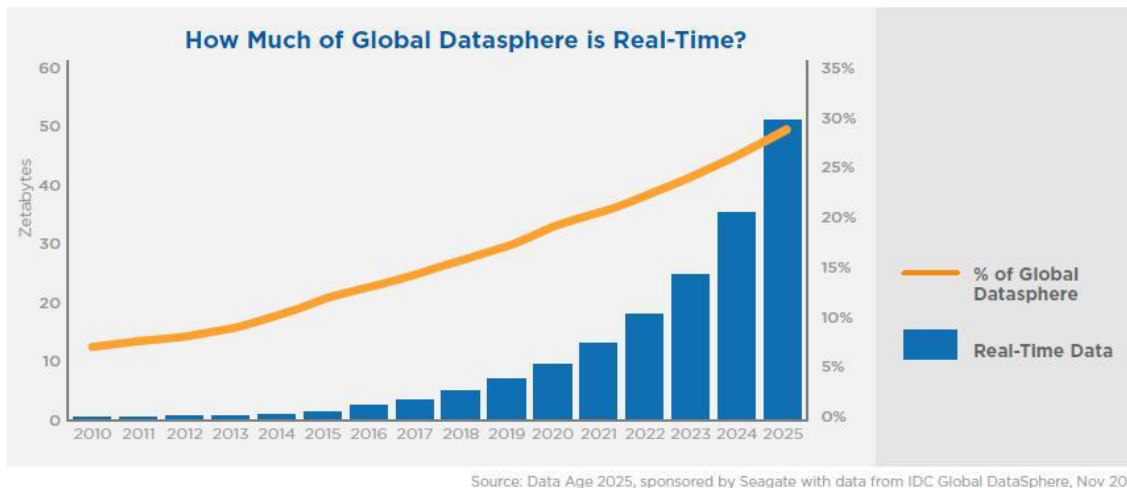
Source: IDC's Data Age 2025 study, sponsored by Seagate

- دنیای مبتنی بر داده آینده دنیایی است که همه چیز در آن ردیابی و نظارت میشود به همه چیز گوش داده میشود و همه چیز توسط دوربینها مورد توجه قرار میگیرد. زیرا این دنیا همیشه در حال تغییر تحول بوده و یادگیری در آن اتفاق می افتد.

- مطابق با پیش بینی IDC در سال 2025 ، حدود 49% از کل داده های دنیا بروی فضاهای ابری ذخیره خواهند شد.

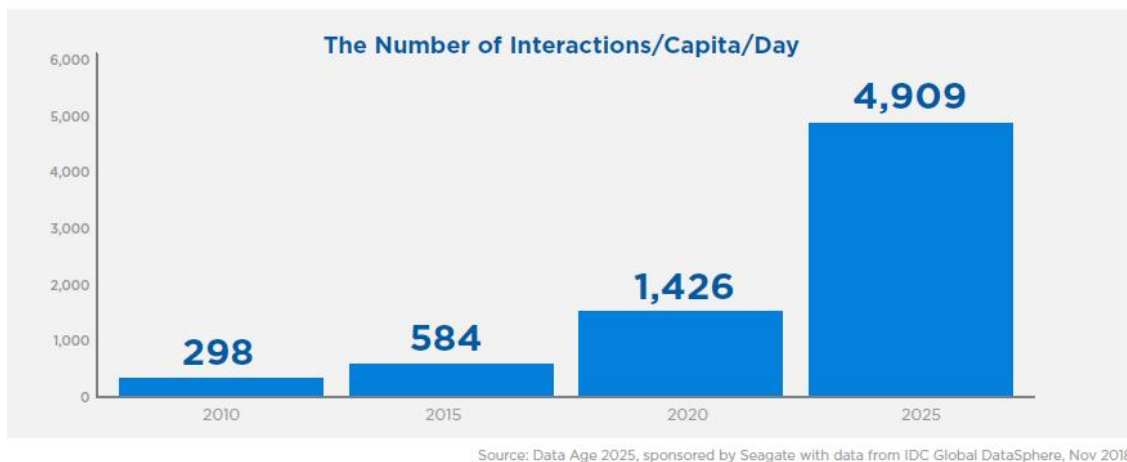


- بر اساس پیش بینیهای IDC ، در سال 2025 حدود 30% از کل داده های جهان دادههایی هستند که نیاز به پردازش بلادرنگ خواهند داشت. در سال 2017 این عدد حدود 15% از کل داده های این سال بوده است.



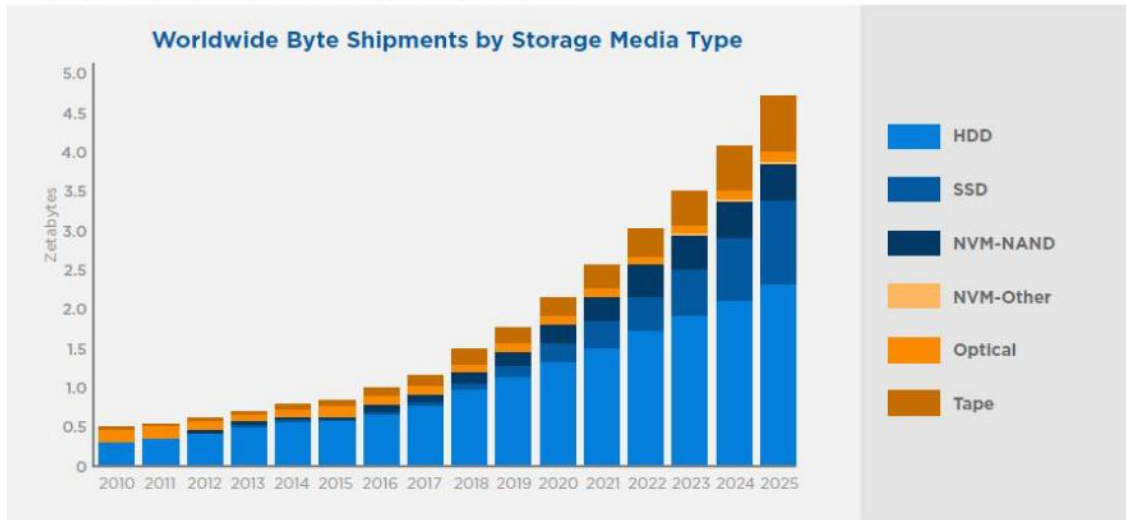
- بر اساس پیش بینیهای IDC میزان تراکنشهای هر فرد در ارتباط با تجهیزات داده محور در سال 2010 حدود 300 تراکنش در روز بوده است که در سال 2025 این عدد به حدود 5000 تراکنش در روز خواهد رسید.

Figure 9 - Data Interactions per Connected Person Per Day



- نمودار تقاضا برای فضاهای ذخیره سازی تا سال 2025 به صورت زیر خواهد بود.

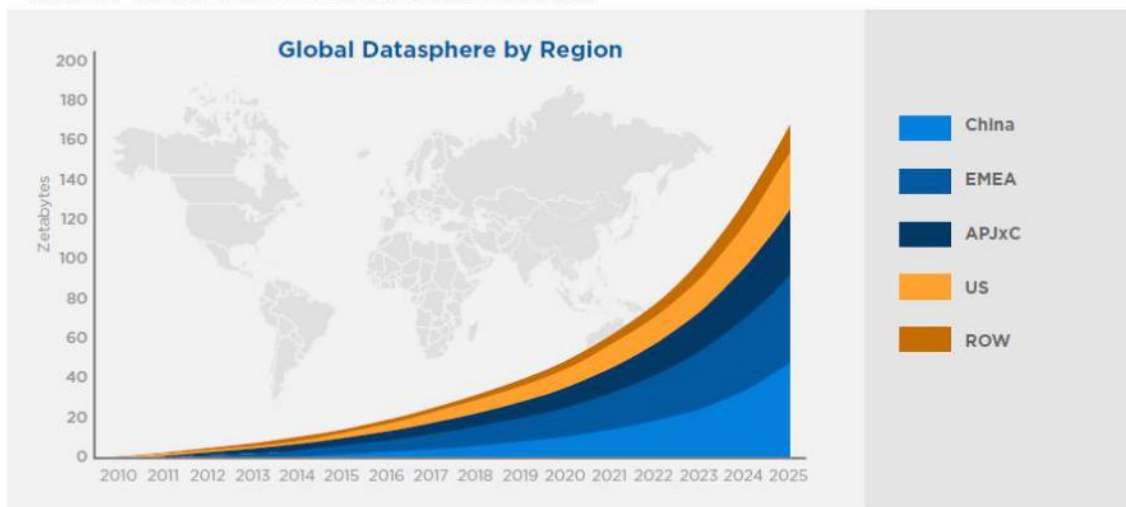
Figure 10 - Worldwide Byte Shipments by Storage Media Type



Source: Data Age 2025, sponsored by Seagate with data from IDC Global DataSphere, Nov 2018

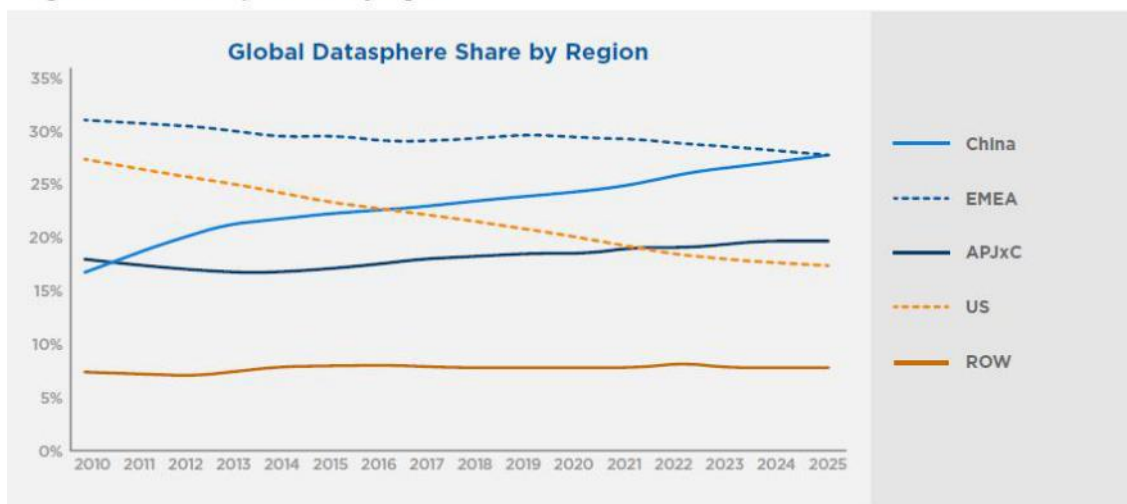
- حوزه های جغرافیایی رشد داده ها بر اساس پیش بینی IDC به شرح ذیل است:

Figure 12 - Size and Growth of the Global Datasphere by Region



Source: IDC's Data Age 2025 study, sponsored by Seagate

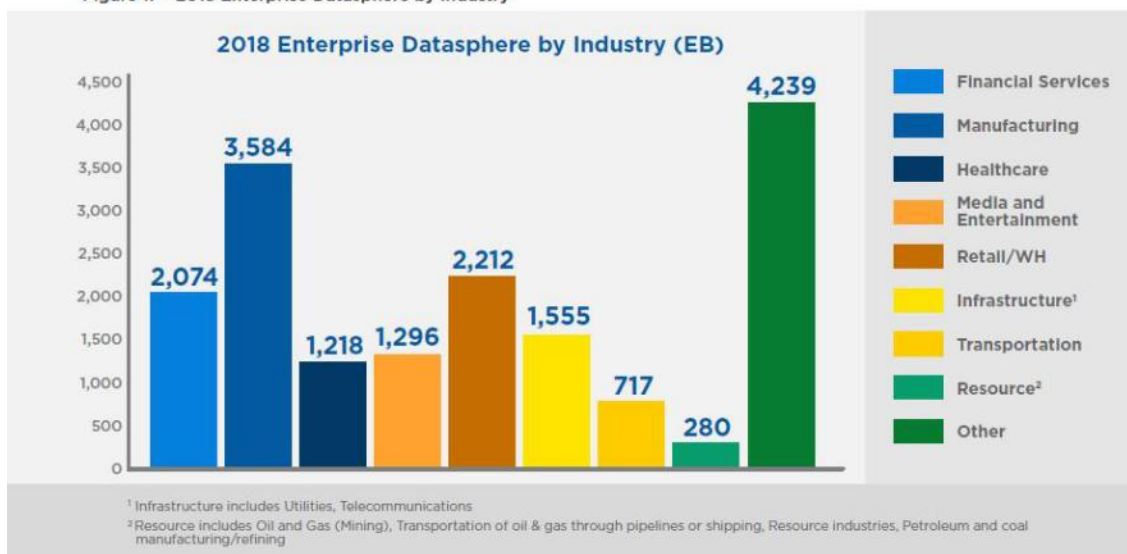
Figure 13 – Global Datasphere Share by Region



Source: IDC's Data Age 2025 study, sponsored by Seagate

- حجم داده های جهانی به تفکیک صنایع مختلف در سال 2018 به شرح زیر است. واحد اندازه گیری داده در شکل زیر **اگزا بایت** است که معادل **1 هزارم یک زتا بایت** میباشد. به عبارت دیگر حجم داده در بخش سرویسهای مالی در شکل زیر حدود 2 زتا بایت و بخش تولید حدود 3.58 زتا بایت میباشد.

Figure 17 - 2018 Enterprise Datasphere by Industry



- انتظار می رود اندازه بازار تجزیه و تحلیل پیشرفته اطلاعات جهانی طی سالهای 2019 تا 2016 تقریباً 106 میلیارد دلار افزایش خواهد یافت. نرخ رشد پیش بینی شده تقریباً 15 درصد به صورت سالیانه افزایش خواهد یافت. رشد در اندازه این بازار به عوامل متعددی از جمله نیاز به بهبود بهره وری کسب و کار و استفاده روزافزون از تجزیه و تحلیل داده های بزرگ، تجزیه و تحلیل IoT و تجزیه و تحلیل مبتنی بر SaaS در برخی از کاربردهایی نهایی از جمله BFSI، خرده فروشی، ورزش و مراقبت های بهداشتی وابسته است.

- با توجه به بررسیهای به عمل آمده بنظر میرسد صنعت تجزیه و تحلیل اطلاعات در بخش بهداشت و درمان از سالهای 2020 تا 2022 یک افزایش قابل توجه را داشته باشد.
- صنعت تجزیه و تحلیل پیشرفته داده ها در واقع یکی از پیشراانه های اصلی مربوط به تحول دیجیتال میباشد.
- انتظار می رود بازار تجزیه و تحلیل پیشرفته تکمیل شده از 8/4 میلیارد دلار در سال 2018 به 18/4 میلیارد دلار تا سال 2023 افزایش پیدا کند.

۱۰. ابزارها و فناوریهای موجود در این زیست بوم

ابزارهای پیشرفته تجزیه و تحلیل برای انتخاب از انواع مختلف وجود دارد که مزایای مختلفی دارند. آنها معمولاً به دو دسته تقسیم می شوند: منبع باز و اختصاصی.

- ابزارهای متن باز برای بسیاری از دانشمندان که داده های مربوط به ماشین را مطالعه می کنند و در تجزیه و تحلیل های مجزا از آنها استفاده می شود، تبدیل به یک گزینه می شود.

آنها عبارتند از زبان برنامه نویسی، و همچنین محیط محاسباتی، از جمله Hadoop و Spark. به طور معمول کاربران می گویند که ابزار تجزیه و تحلیل پیشرفته منبع باز هستند، زیرا عموماً ارزان هستند و

کارایی بالایی دارند و قابلیت پشتیبانی قوی دارند و توسط یک جامعه کاربری پشتیبانی می شوند که به طور مداوم ابزارها را ارتقا می دهد.

- در بخش اختصاصی، فروشندگان از جمله مایکروسافت، آی بی ام و موسسه SAS همه ابزارهای تجزیه و تحلیل پیشرفته ارائه می دهند. اغلب نیاز به یک پس زمینه عمیق فنی و درک تکنیک های ریاضی است.

اما در سالهای اخیر، ابزار تجزیه و تحلیل ابزار خودآموزی بالغ شده است تا عملکرد را برای کاربران کسب و کار آسان تر کند. به طور خاص، تابلو، تبدیل به یک ابزار رایج شده است. در حالی که قابلیت آن محدودتر از ابزارهای فنی عمیقتر است، کاربران را قادر می سازد تجزیه و تحلیل خوشه ای و سایر تحلیل های پیشرفته انجام دهند.

فصل پنجم

نتیجه گیری

نتیجه گیری

بر اساس پیش بینی موسسه IDC حجم کل داده های جهان با سرعتی باور نکردنی در حال رشد است. حجم کل داده های جهان که در سال 2018 معادل 33 زتابایت بوده است در سال 2025 به حجم 175 زتابایت خواهد رسید. این حجم بسیار زیاد از اطلاعات از منابع مختلف و متعددی تولید میشود که برخی از آنها عبارتند از :

- داده های مربوط به تماسهای تلفنی
- شبکه های اجتماعی و اینترنت
- داده های مربوط به عبور و مرور در کل سطح جهان
- تراکنشهای مالی
- داده های مربوط به صنایع مختلف
- نیز داده های مربوط به اینترنت اشیاء و ...

داشتن اطلاعات در هر زمینه ایی همواره مفید بوده است. تجزیه و تحلیل داده ها و اطلاعات با اهداف ذیل میتواند کیفیت زندگی را بسیار بالاتر ببرد:

- سهولت و دقت در تصمیم گیری
- کشف علت وقوع برخی رخدادها
- پیش بینی رخدادها و حوادث آینده با استفاده از تجزیه و تحلیل اطلاعات

افزایش داده ها از یک سو و نیز پیچیده شدن ابزارها و فرآیندهای کار نیازمندی به تجزیه و تحلیل اطلاعات را دو چندان کرده است.

جمع آوری و یکپارچگی داده به تنهایی ارزشی را تولید نمی کند و نیازمند مدل های پیشرفته جهت تجزیه و تحلیل داده می باشیم ، مدل هایی که می بایست همسو با استراتژی تجزیه و تحلیل و با هدف تولید ارزش تعریف گردند. هدف اصلی ، شناسایی موارد و موضوعاتی است که دارای ظرفیت لازم جهت خلق ارزش اضافه برای کسب و کار می باشند . علاوه براین ، می بایست خود را آماده برخورد با چالش هایی نظیر ایجاد مجدد

فرآیندها، تغییرات سازمانی، تغییر رفتار سازمانی و مجموعه ای از تغییرات کرد که به دنبال تصمیم گیری مبتنی بر تجزیه و تحلیل داده اتخاذ می گردد.

تجزیه و تحلیل اطلاعات و تبدیل داده های خام به اطلاعات ارزشمند و دارای دانش و بینش نیازمند به یک ساز و کار نسبتاً پیچیده فنی و عملیاتی دارد. جمع آوری اطلاعات، پالایش و یکسان سازی اطلاعات، کشف الگوهای تکراری و نیز تهیه مدل های تصمیم گیری و پیش بینی آینده نیازمند به استفاده از علوم آماری، شبکه های عصبی و هوش مصنوعی دارد. لذا علم تجزیه و تحلیل اطلاعات در واقع یک اکوسیستم است که دارای زنجیره نسبتاً طولانی از فناوریها و ابزارهای مختلف میباشد. ابزارها و استانداردهای جمع آوری، ذخیره سازی، شناسایی الگوها و تهیه مدل های یادگیری و در نهایت، تبدیل اطلاعات به فرمتهای مختلف ارزشمند شامل این مجموعه می باشند. اکثر سازمان ها به این موضوع اذعان دارند که با تجزیه و تحلیل داده بهره وری آنها افزایش خواهد یافت، فرآیند تصمیم گیری آنها بهبود خواهد یافت و دارای فرصت های بیشتری جهت ایجاد مزایای رقابتی خواهند شد. آینده بسیاری از بنگاه های کسب و کار به تجزیه و تحلیل داده بستگی دارد. توان یک بنگاه کسب و کار جهت تجزیه و تحلیل داده به یکی از تفاوت های اصلی و مهم بین برندگان و بازندگان دنیای کسب و کار تبدیل شده است.

مهم نیست هم اینک در چه سطحی از سطوح تجزیه و تحلیل داده باشیم، مهم شناسایی وضعیت فعلی، ایجاد زیرساخت و پلت فرم مورد نیاز، انطباق خواسته ها و انتظارات با وضعیت فعلی و حرکت هدفمند به سمت سطوح بعدی تجزیه و تحلیل داده است. این کار محقق نمی شود مگر با ایجاد یک استراتژی دقیق تجزیه و تحلیل داده و به دنبال آن یک نقشه راه که جزئیات کار را در طول مسیر مشخص می کند.

فصل ششم

پاسخ به سوالات کلیدی

سوالات کلیدی

در ادامه تعدادی از سوالات کلیدی که میتواند در روشن شدن موضوع کمک شایانی نماید آورده شده است. ممکن است پاسخ برخی از سوالات ذیل در بخشهای قبلی گزارش ذکر شده باشد اما در این قسمت نیز تکرار شده است تا خوانندگان گزارش در صورت نیاز نیاز به مرور بخشهای غیرضروری نداشته باشند.

سوال	اهداف اصلی تجزیه و تحلیل پیشرفته اطلاعات چیست؟
جواب	<p>✓ استفاده مناسب از اطلاعات در بهبود کیفیت زندگی</p> <p>✓ امکان تصمیم گیری سریع و دقیق بر مبنای تصمیمات و نتایج قبلی بر اساس داده ها</p> <p>✓ کشف علل وقوع برخی حوادث در ابزارها و طبیعت و رفتارها در انسانها</p> <p>✓ پیش بینی وقوع برخی حوادث در ابزارها و طبیعت و رفتارها در انسانها به منظور جلوگیری از افزایش هزینه های جانی و مالی</p>

سوال	برای رسیدن به اهداف اصلی مطرح شده در بخش قبل چه باید کرد؟
جواب	<p>✓ باید موارد کاربرد این فناوری را شناسایی نمود.</p> <p>✓ باید داده های گذشته را به صورت موثر و قابل دسترس نگهداری کنیم.</p> <p>✓ باید داده های گذشته را مورد تجزیه و تحلیل قرار داد و بدنبال الگوهای تکراری باشیم.</p> <p>✓ باید نتایج را به صورتهای قابل فهم و قابل تصمیم گیری تبدیل نماییم.</p> <p>✓ باید زیرساختهای سخت افزاری مناسب جهت پردازش سریع و کامل اطلاعات فراهم کرد.</p> <p>✓ باید از فناوریهای مرتبط که برای پیش بینی و تصمیم گیری به ما کمک میکنند در تجزیه و تحلیل اطلاعات استفاده نمود.</p>

سوال	در کدام حوزه های زندگی بشر میتوان از این فناوری استفاده نمود؟
جواب	<ul style="list-style-type: none"> ✓ اقتصاد و بازرگانی ✓ مالی و بانکی ✓ بهداشت و درمان ✓ رفت و آمد ✓ ارتباطات ✓ آموزش ✓ مشتری مداری ✓ امنیت و جرائم ✓ بیمه و ریسک ✓ حوادث رانندگی ✓ تامین انرژی

سوال	فناوریها و ابزارهای مرتبط با تحلیل داده پیشرفته چیست؟
جواب	<ul style="list-style-type: none"> ✓ فناوریهای کلان داده ✓ فناوریهای هوش مصنوعی ✓ فناوریهای یادگیری ماشین و یادگیری عمیق ✓ فناوریهای هوش تجاری ✓ شبکه های عصبی ✓ اینترنت اشیاء

سوال	در کدام حوزه صنعت یا زندگی ، فرصتهای بیشتری برای استفاده از تحلیل داده وجود دارد؟
جواب	<p>✓ در حوزه شهرهای هوشمند</p> <p>✓ در حوزه سلامت و بهداشت</p>

سوال	اهداف اصلی در استفاده از تجزیه و تحلیل پیشرفته در صنعت بانکی و مالی چیست؟
جواب	<p>بر اساس حوزه خدمات بانکداری و سرویسهای مالی میتوان اهداف زیر برای تجزیه و تحلیل پیشرفته متصور بود:</p> <p>✓ اصلاحات نظارتی</p> <p>✓ سودآوری مشتری</p> <p>✓ بهره وری عملیاتی</p> <p>✓ مدیریت ریسک</p> <p>✓ تجزیه و تحلیل ریسک اعتباری</p> <p>✓ تشخیص تقلب و سوء مدیریت</p> <p>✓ بودجه بندی و برنامه ریزی</p> <p>✓ بهینه سازی فرآیند</p> <p>✓ بهبود کمپین های تبلیغاتی</p> <p>✓ ارائه پاسخ سریع به ترندهای بازار و ایجاد مزیت رقابتی</p> <p>✓ ارائه خدمات بهتر به مشتری</p>

سوال	اهداف اصلی در استفاده از تجزیه و تحلیل پیشرفته در صنعت فناوری اطلاعات و ارتباطات چیست؟
جواب	<p>در حوزه خدمات مخابرات و فناوری اطلاعات میتوان اهداف زیر برای تجزیه و تحلیل پیشرفته متصور بود:</p> <p>✓ پیشنهاد هدف گذاری و مدیریت کمپینهای تبلیغاتی</p> <p>✓ Cell Site Optimization</p>

✓ تضمین در آمد	
✓ تجزیه و تحلیل سودآوری مشتری	
✓ دینامیک شبکه	
✓ کنترل جابجایی و تحلیل شبکه اجتماعی	

سوال	اهداف اصلی در استفاده از تجزیه و تحلیل پیشرفته در صنعت بهداشت و سلامت چیست؟
جواب	✓ مدل سازی پیش بینی شده ✓ نظارت بر عملکرد مالی ✓ تشخیص تقلب و سوء مدیریت

سوال	اهداف اصلی در استفاده از تجزیه و تحلیل پیشرفته در اداره مملکت چیست؟
جواب	بهبود روندهای حاکمیتی شامل : ✓ تدوین راهبردها ✓ نظارت برنامه های راهبردی ✓ کنترل و ایجاد برنامه های بهبود با استفاده از تصمیم گیریها ✓ بررسی کیفیت تصمیم گیریهای گذشته ✓ پیش بینی اقدامات آتی جهت بهبود اجرای برنامه های کلان و راهبردی مملکت

سوال	اهداف اصلی در استفاده از تجزیه و تحلیل پیشرفته در صنعت دفاعی و امنیت چیست؟
جواب	✓ تشخیص تقلب و سوء مدیریت ✓ سلامت دفاعی

✓ برنامه ریزی مبتنی بر سناریو	
-------------------------------	--

سوال	اهداف اصلی در استفاده از تجزیه و تحلیل پیشرفته در صنعت حمل و نقل و تدارکات چیست؟
جواب	<p>✓ بهینه سازی موجودی انبارها</p> <p>✓ برنامه ریزی زنجیره تامین</p> <p>✓ فروش و برنامه ریزی عملیاتی</p> <p>✓ مدیریت چرخه عمر و کیفیت محصول</p>

سوال	اهداف اصلی در استفاده از تجزیه و تحلیل پیشرفته در خصوص کیفیت زندگی انسان چیست؟
جواب	<p>در طول تاریخ، ایده‌های نوآورانه از ابتکار و خلاقیت انسانها به وجود آمده است، و این ابتکار و خلاقیت از مولفه های اصلی رشد بهره وری، نوآوری در فرآیندها و محصولات محسوب شود. اما در عصر حاضر داده‌ها و الگوریتمها هم میتوانند برخی این موارد را پشتیبانی، ارتقاء یا حتی جایگزینی قابل انعطاف برای خلاقیت بشر باشند.</p> <p>علم تجزیه و تحلیل داده ها در حوزه نوآوری فرآیندها، به سازمان ها کمک می کند تا چگونگی ساخت تیم ها، منابع و جریان های کاری را مشخص نمایند. داشتن اطلاعات در خصوص تیم های با کارایی بالا و همچنین تیم های با کارآمدی کم میتواند به درک این اختلاف بین فعالیتهای آنها و نیز چگونگی ایجاد همکاری مؤثر بین افراد و تیمها، فرصت بزرگی برای سازمان ها بوجود می آورد.</p> <p>این تحلیل میتواند شامل موارد زیر باشد:</p> <ul style="list-style-type: none"> • بهبود کیفیت تصمیم گیریها • مسائلی مانند مکمل مهارت ها • اندازه تیم های مطلوب • اینکه آیا تیم ها باید با هم کار کنند یا خیر؟ • میزان اهمیت تجربیات گذشته افراد در عملکرد فعلی آنها • آموزش مهم است یا خیر • بررسی ویژگیهای شخصیتی افراد. <p>تجزیه و تحلیل داده ها می توانند فرضیه ها را آزمایش نموده و الگوهای جدیدی را پیدا کنند بطوریکه حتی ممکن است مدیران سازمان به این الگوها واقف نباشند.</p>

<p>تحلیل مقادیر ایمیلها، تقویم کاری، موقعیت مکانی و سایر اطلاعات در دسترس میتواند کمک کند تا بدانند که افراد چگونه با هم کار می کنند و ارتباط برقرار می کنند، لذا بررسی این اطلاعات می تواند بینشهای جدیدی را در مورد بهبود عملکرد افراد و تیمها به وجود آورد.</p> <p>در بخش نوآوری محصول، تحلیل داده ها می توانند تحقیق و توسعه را در زمینه هایی مانند علوم مواد، زیست شناسی مصنوعی و علوم زیستی تغییر دهند. شرکت های پیشرو در زمینه دارویی با استفاده از تجزیه و تحلیل داده ها ه کشف داروهای جدید کمک میکنند. داده های جمع آوری شده از منابع مختلف می توانند ترکیبات شیمیایی را که به عنوان درمان موثر داروها برای انواع بیماری ها مورد استفاده قرار می گیرد، بهتر تعیین نمایند.</p>	
--	--

سوال	اهداف اصلی در استفاده از تجزیه و تحلیل پیشرفته در صنعت کالاهای مصرفی و خرده فروشی چیست؟
جواب	<p>✓ بهینه سازی قیمت</p> <p>✓ بینش مشتری</p> <p>✓ برنامه ریزی و سازمان</p> <p>✓ برنامه ریزی فروش</p> <p>✓ بهینه سازی حجم و اندازه</p>

سوال	عوامل اصلی در رشد و توسعه فناوری تجزیه و تحلیل پیشرفته چیست؟
جواب	<ul style="list-style-type: none"> • استقرار راه حلها بر روی فناوری ابر • افزایش تقاضا برای مقابله با چالش های کلان داده ها. • افزایش نیاز به راه حل های عمودی عمودی • راه حل های تحلیلی پیشرفته در اقتصادهای نوظهور

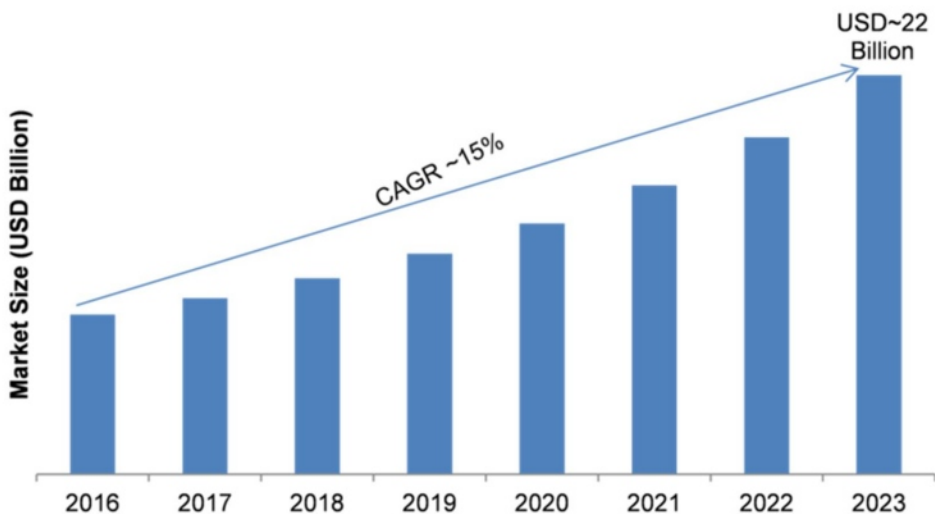
<p>فراهم بودن راه حل های تحلیلی پیشرفته بر مبنای سرویسهای ابری هزینه ها را کاهش می دهد و مشتریان جدیدی را جذب می کند که قبلاً بخاطر هزینه های بالا از آنها خودداری می کردند. به دلیل کاهش قیمت ها، شرکت های کوچک و متوسط نیز از تجزیه و تحلیل پیشرفته استفاده می کنند، که باعث افزایش تعداد مشتریانی - که می خواهند از این راه حل ها استفاده کنند- خواهد شد. از سوی دیگر، اقتصادهای نوظهور خواستار تجزیه و تحلیل پیشرفته برای پیشبرد کار خود و همچنین وفاداری مشتریان هستند.</p>	
--	--

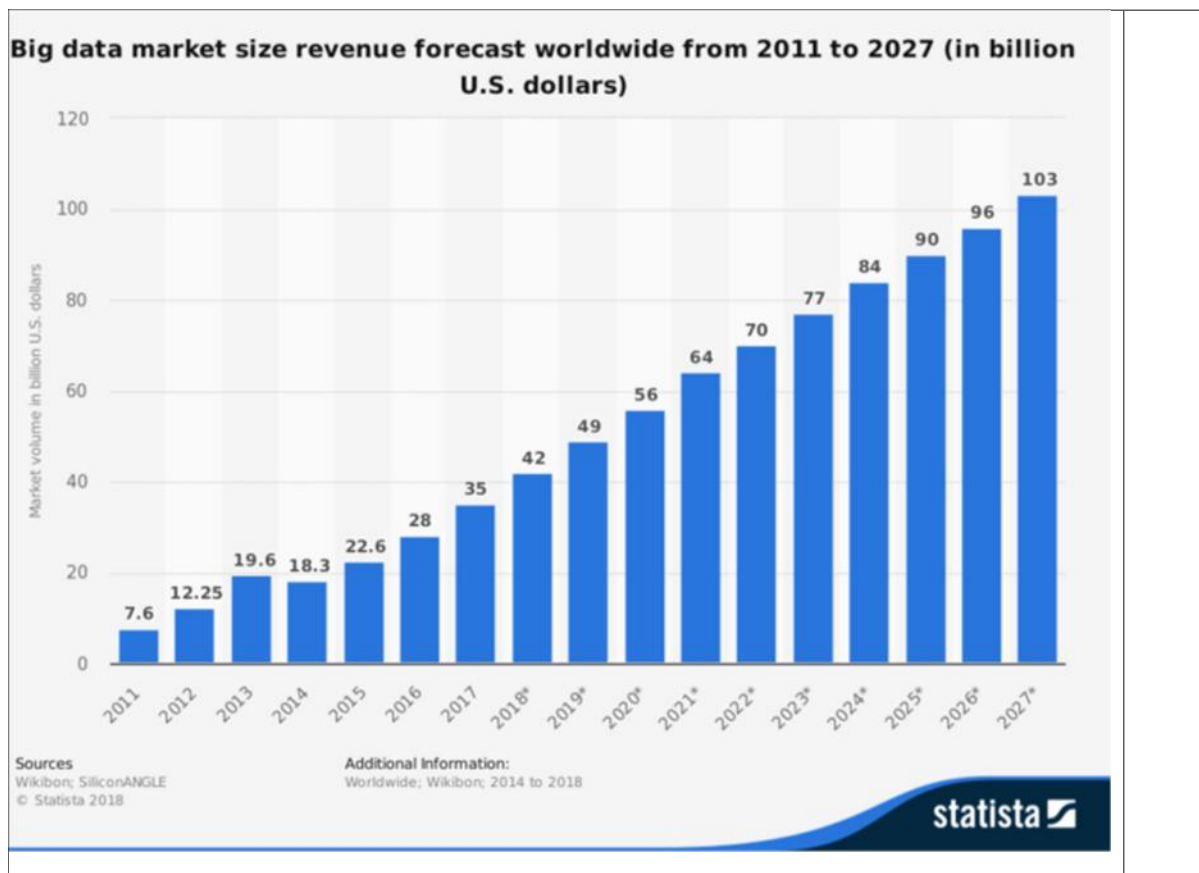
سوال	بازیگران اصلی در این صنعت چه کسانی هستند؟
جواب	<ul style="list-style-type: none"> • بازار تجزیه و تحلیل پیشرفته شاهد ظهور بازیگران جدیدی است که بر روی ارائه راه حل های خاص به این مشتریان متمرکز هستند. ظهور شرکت هایی مانند Statsoft ، آنگس، Knime، که راه حل های ارزانتری را نسبت به SAS، SAP ارائه می کنند. باعث شده است تا بازیگران اصلی نیز در تمرکز بر توسعه راه حل ها جدی تر شده و راه حل های سفارشی ارائه نمایند. بعضی از بازیگران برجسته در بازار تحلیلی پیشرفته عبارتند از: • Statsoft - www.statsoft.com • SAS موسسه (U.S) - https://www.sas.com/en_us/home.html • KNIME AG (سوئیس) - https://www.knime.com/

<ul style="list-style-type: none"> • شرکت اوراکل (ایالات متحده) - https://www.oracle.com/corporate/ • شرکت IBM (ایالات متحده) - https://www.ibm.com/ • شرکت نرم افزار آنگس (کانادا) - https://www.anges.co.jp/en/company/index.html • SAP SE (آلمان) - https://www.sap.com/corporate/en/company.html • شرکت مایکروسافت آمریکا - https://www.microsoft.com/ • شرکت فیکو (ایالات متحده) - https://www.fico.com/en 	
---	--

سوال	چرا به یک پلتفرم پیشرفته به منظور پیاده سازی تجزیه و تحلیل داده های پیشرفته داریم؟
جواب	<p>واحدها فناوری اطلاعات در همه شرکت های بزرگ با چندین تصمیم مهم در خصوص انتخاب معماری جدید برای سیستمها مواجه هستند. به همین منظور سازمانها معمولاً با شرایط زیر مواجه هستند:</p> <ul style="list-style-type: none"> • در قدم اول، معمولاً در سازمانها یک زیرساخت وجود دارد که تعداد زیادی از کارشناسان که وظیفه نگهداری و پشتیبانی از آن را بر عهده دارند، البته این زیرساخت به شدت توسط تقاضای رو به رشد برای چهار Vs تحت فشار قرار میگیرد و همواره با سونامی کلان داده مواجه است. چهار V شامل موارد زیر است : <ul style="list-style-type: none"> ○ Volume : حجم داده ○ Variety : تنوع داده ○ Veracity : صحت داده

<p>○ Velocity : سرعت تولید داده</p> <p>لذا در اغلب موارد، سازمان به این نتیجه میرسد که ادامه سرمایه گذاری بروی زیرساخت فعلی به منظور پاسخگویی به نیازهای آینده مقرون به صرفه نیست.</p> <ul style="list-style-type: none"> • در قدم دوم و با توجه به فشار بازار بر سازمانها به منظور توسعه زیرساختها بر مبنای تجزیه و تحلیل داده ها ، آنها نیاز به انجام تغییرات بنیادینی در حوزه های مختلف - مربوط به بازاریابی، فروش، عملیات، مدیریت عملیات، مصرف کنندگان هوشمند ، رقابت سالم با رقبا، هوشمند سازی تامین – میباشند. • در قدم سوم، وجود زیرساختهای جدید، زیرساختهای متن باز، زیرساخت ابری تولید شده است که میتواند نیازهای روز را پوشش دهد. 	
---	--

سوال	نمودار رشد این صنعت چگونه است؟
پیش بینی اول :	 <p>Market Size (USD Billion)</p> <p>CAGR ~15%</p> <p>USD~22 Billion</p> <p>2016 2017 2018 2019 2020 2021 2022 2023</p>
جواب	پیش بینی دوم :



سوال	حوزه های کاربرد تحلیل پیشرفته اطلاعات در مدیریت عملکرد منابع انسانی چیست؟
جواب	<ul style="list-style-type: none"> داشتن اطلاعات دقیق و کاملی و نیز تحلیل اطلاعات مربوط به افراد و نیز تیمهای موفق و با عملکرد بالا در سازمان و مقایسه آن با تیم های ناموفق میتواند بینشهای بسیار ذیقیمتی را در اختیار سازمانها و مدیران قرار دهد.

سوال	معروفترین ابزارهای تجزیه و تحلیل پیشرفته اطلاعات کدام هستند؟
جواب	جدول زیر فهرستی از برخی محصولات در حوزه تجزیه و تحلیل پیشرفته را به همراه شرکت تولید کننده و همچنین اطلاعات کلی از قابلیت های آنها را ارائه می کند:

Product category/name	Company's website	Comment
Comprehensive BI suite Business Objects XI 3.0	www.businessobjects.com	Offers a broad family of BI tools and applications for different sizes of businesses
IBM Cognos 8 BI	www.cognos.com	Offers a wide range of BI capabilities on a single, service-oriented – architecture (SOA)
Data Mining Tool Clementine from SPSS	www.spss.com	Leading visual rapid modeling environment for data mining, available in several major languages
DBMiner 2.0 (Enterprise)	www.dbminer.com	Powerful and affordable tool to mine large databases; uses Microsoft SQL Server 7.0 Plato
Oracle Data Mining (ODM)	www.oracle.com	Uses data mining functionality embedded in Oracle Database 10 g Enterprise Edition
SAS Enterprise Miner	www.sas.com	Integrated suite, provides an user-friendly GUI frontend to the SEMMA (Sample, Explore, Modify, Model, Assess) process
Text Mining Tool Inxight	www.businessobjects.com	Enterprise software solutions to understand and analyze information contained in unstructured text, in all major languages
SAS Text Miner	www.sas.com	Provides a rich suite of tools for discovering and extracting knowledge from text documents
SPSS LexiQuest	www.spss.com	To mine the “unstructured” information contained in text documents, integrated with SPSS Clementine data mining tool

Web Mining Tool MicroStrategy Web Traffic Analysis Module	www.microstrategy.com	Built on MicroStrategy 7 platform, provides traffic highlights, contents analysis, and web visitor analysis reports
SAS Webhound	www.sas.com	Analyzes website traffic to answer questions like: Who is visiting? How long do they stay? What are they looking at?
SPSS Web Mining for Clementine	www.spss.com	Enables extraction of Web events, including online campaign results, uses this online behavior in predictive modeling
WebTrends	www.webtrends.com	A suite for data mining of web traffic information to learn and use the customer behavior to take targeted action

سوال	حوزه های کاربرد تحلیل پیشرفته اطلاعات چیست؟
جواب	<ul style="list-style-type: none"> • صنعت بانکی و مالی از تجزیه و تحلیل پیشرفته برای کاهش تقلبات، بهینه سازی فرایندهای دنبال شده توسط بانک و مدیریت ریسک ها استفاده میکنند. • صنعت مخابرات از تجزیه و تحلیل ها برای محافظت از مشتریان استفاده می کند و برای برنامه ریزی دقیق برنامه های بازاریابی برای مشتریان خاص استفاده میکند، • صنعت بهداشت و درمان با استفاده از تجزیه و تحلیل برای بهبود عملکرد بالینی، و تشخیص و کشف تقلب مرتبط با صنعت بیمه استفاده خواهد کرد. • شرکت های خرده فروشی نیز به منظور تحلیل رفتار خرید مشتریان و همچنین به خرده فروشان در برنامه ریزی و بهینه سازی تجربه مشتریان از این فناوری استفاده خواهند کرد. • مدیریت عملکرد و منابع سازمانی • مدیریت زندگی انسانها

سوال	در فناوری تجزیه و تحلیل پیشرفته اطلاعات چگونه باید به چه نکاتی توجه نمود؟
جواب	<p>فن آوری های انبارداری داده شامل توابعی ETL (استخراج، دگرگونی داده ها و ارائه) برای انتقال داده های قدیمی و همچنین دیگر داده به داخل و خارج از یک انبار داده های جامع سازمانی میباشد. ممکن است به پاکسازی و تمیز کردن داده ها نیازمند باشیم که شامل تمیز کردن، استاندارد سازی و دریافت داده ها از سیستم های قدیمی باشد. ارتقاء کیفیت داده ها و در صورت نیاز شامل اضافه کردن داده های خارجی مانند اطلاعات دموگرافیک یا فضایی می شود.</p> <p>در صورت نیاز به شناسایی اطلاعات سودآوری مشتری باید ارزش تاریخی، فعلی و پیش بینی شده مشتریان و سپس استفاده از آن برای بهبود بخشیدن و اجرای استراتژی های مشتری اقدام نمود.</p> <p>الگوریتم های جستجوی وب، داده کاوی، جستجوی متن - برای داده های استخراج شده از انبار داده ها - انتخاب شده و برای تولید مدل های پیش بینی مورد استفاده قرار خواهند گرفت. به طور کلی، قبل از انجام عملیات فوق، ابتدا داده ها با افزودن ویژگی های جدید و اضافی به انبار داده غنی می شوند.</p>

برنامه های تجزیه و تحلیل پیشرفته باید به راحتی با سیستم های ارائه اطلاعات مجتمع سازی شوند تا نتایج تحلیل با مدیران ارشد و تصمیم گیران سازمان - به سادگی - به اشتراک گذاشته شوند.

این نوع تکنولوژیها شامل مجموعه گسترده ای از الگوریتم ها، تکنیک های نمایه سازی، کشف داده های انعطاف پذیر و قابل دستکاری است که مدل های دقیقی را تولید می کنند. نتایج باید به صورتی عرضه شوند که حتی کاربران غیر فنی نیز توان درک و تحلیل آنها را داشته باشند نتایج باید بطور کامل با استفاده از اصطلاحات تجاری روزمره ارائه گردد.

بنا براین برای پیاده سازی ایده های فوق، که شامل ایجاد پروفایل و تقسیم بندی اطلاعات کاربران میشود. یک تحلیلگر اطلاعات ممکن است ابتدا از خوشه بندی استفاده کند تا گروههای نسبتاً همگن از مشتریان را پیدا کند که رفتارهای خرید مشابهی را نشان می دهند که البته به این ترتیب، پایگاه داده مشتری تقسیم بندی خواهد شد. هنگامی که این بخش های اطلاعاتی از هم تفکیک می شوند، مدلهای پیش بینی یا آماری می توانند برای پیش بینی رفتار خرید اقدام نمایند. به عنوان مثال، برای پیش بینی رفتار خرید هر خوشه از کاربران، رگرسیون را اعمال کنید. سپس هر یک از این گروه ها محصول و خدمات مربوط به نمایه سازی خود را دریافت می کند که باعث صرفه جویی زیادی در هزینه ارسال کاتالوگ های ارسال شده به مشتریان می شود.

پیاده سازی تجزیه و تحلیل پیشرفته می تواند در یک محدوده از زیرساختهای فناوری انجام شود. برنامه های پیشرفته تجزیه و تحلیل را می توان در یک طیف گسترده از سخت افزارها - که از **Main Farme** شروع شده و تا یک کامپیوتر خانگی ادامه پیدا میکند - اجرا نمود. قیمت برخی از این سیستمها به چند هزار دلار تا چند میلیون دلار بالغ میگردد. قیمت یک زیرساخت فیزیکی جهت اجرای تجزیه و تحلیل پیشرفته معمولاً با معیارهای زیر تعیین میشود:

✓ اندازه و حجم انبار داده ها

✓ پیچیدگی پرس و جو هایی که باید از انبار داده پرسیده و پاسخ داده شوند.

<p>دو عامل دیگر که باید در هنگام تحلیلی پیشرفته مورد توجه قرار گیرند عبارتند از :</p> <p>✓ کیفیت داده های موجود در انبار داده ها ، اگر نتایج حاصل از تجزیه و تحلیل پیشرفته برای تصمیم گیری مورد استفاده قرار خواهد گرفت کیفیت داده ها باید بالا باشد. توجه به این موضوع وقتی بسیار مهم است که قرار است داده هایی را از فروشندگان خارجی تهیه و خریداری نمود. لذا در این حالت باید به بررسی کیفیت داده ها پرداخت.</p> <p>✓ استخدام تحلیلگران کسب و کار - که امکان پیاده سازی موثر و استفاده صحیح از تجزیه و تحلیل پیشرفته را داشته باشد - بسیار سخت و پرهزینه است. تجزیه و تحلیل پیشرفته داده ها نیازمند به مجموعه ای از مهارت ها - در مدیریت داده ها، تجزیه و تحلیل آماری و تصمیم گیری در کسب و کار - دارد که معمولاً پیدا کردن چنین افرادی بسیار سخت است.</p>	
---	--

نیازمندیهای انجام تجزیه و تحلیل چیست؟	سوال
<p>در ادامه لیستی از نیازهای مهم تجزیه و تحلیل کسب و کار به شرح ذیل است:</p> <ul style="list-style-type: none"> • کشف داده ها و اطلاعات : کشف الگوها و اطلاعات معنادار و مهم از منابع اطلاعاتی مختلف • نمایه سازی (مصور سازی) اطلاعات: مشخص کردن نحوه نمایش اطلاعات به فرمتهای گوناگون تصویری شامل تصویر، گراف، نمودار و • تجزیه و تحلیل اطلاعات : اشاره به تکنیک ها و فرآیندهای کیفی و کمی مورد استفاده برای افزایش بهره وری و کسب و کار است در این مرحله داده ها برای شناسایی و 	جواب

<p>تجزیه و تحلیل داده ها و الگوهای رفتاری استخراج و دسته بندی می شوند و با توجه به الزمات سازمانی تکنیک ها تحلیل اطلاعات متفاوت است.</p> <ul style="list-style-type: none"> تجزیه و تحلیل پیشرفته : یک اصطلاح است که مانند یک چتر بروی طیف گسترده ای از تکنیک های تجزیه و تحلیل است که استفاده از تکنیک های محاسبات پیشرفته مانند یادگیری ماشین ، شبکه های عصبی ، داده کاوی و یادگیری عمیق است. این تجزیه و تحلیل ها نسبت به روشهای قدیمی برتری دارد. هدف از تجزیه و تحلیل پیشرفته کشف ینش، پیش بینی یا ایجاد توصیه برای بهبود کسب و کار میباشد. گزارشگیری و مدیریت اطلاعات : فرایند جمع آوری، ذخیره، مدیریت و نگهداری اطلاعات در تمام اشکال آن است .مدیریت اطلاعات یک اصطلاح عمومی است که شامل سیاست ها و روش های مدیریت و به اشتراک گذاری اطلاعات در میان افراد، سازمان ها و / یا سیستم های اطلاعاتی در سراسر چرخه زندگی اطلاعات می باشد. 	
---	--

سوال	این فناوری برای کدام سازمانها مناسب تر است؟ چرا؟
<p>سازمانهایی که دارای ویژگیهای زیر باشند:</p> <ul style="list-style-type: none"> وجود کلان داده در سازمان نیاز به تصمیم گیریهای متعدد که نیاز به اطلاعات متنوع و همه جانبه دارد 	جواب

فصل هفتم

پیوستها

پیوست شماره یک

فرصت‌های پیش روی صنعت تجزیه و تحلیل اطلاعات

۱. داده کاوی (جستجوی داده)

داده کاوی یکی از ابزارهای اصلی برای ایجاد مزیت رقابتی برای سازمانهای مشتری محور میباشد. شرکتهای پیشرو در بازار با تمرکز بر کاربردهای داده کاوی به صورت آهسته و پیوسته از تحلیل مشتری به سمت تحیل روابط تغییر یافته است. کسب و کارهای رقابتی و بازار مصرف کنندگان برای شرکت ها نه تنها برای جذب مشتری بلکه برای حفظ مشتریان بسیار سودآور آنها، حتی درصد کمی از ضروری است. استراتژی های مورد نیاز برای حفظ مشتریان ارزشمند به طور کلی بر انگیزه های مالی و / یا سطح خدمات در ارتقای وفاداری تاثیر گذار است. از آنجایی که تنها چند شرکت می توانند از اقتصاد برای حفظ تمایز رقابتی در برابر قیمت ها بهره مند شوند، بسیاری از شرکت ها به دنبال افزایش ارزش مشتری با ایجاد وفاداری از طریق برند و تمایز خدمات می باشند.

شرکتهای پیشرو در عرصه کسب و کار از تجزیه و تحلیل پیشرفته با داده کاوی به منظور بهینه سازی روابط خود با مشتریان استفاده می کنند. به عنوان نمونه میتوان به موارد ذیل اشاره نمود:

✓ بهبود کارآیی کمپین های بازاریابی و جذب مشتریان جدید

✓ به حداکثر رساندن میزان فروش به مشتریان موجود

- ✓ به حداقل رساندن از دست دادن مشتریان
- ✓ تعیین نمره ریسک اعتبار مشتریان
- ✓ تحلیل و مدلسازی چرخه عمر ارزش مشتریان

علاوه بر موارد فوق همچنین از تکنیک های داده کاوی میتوان - برای تحلیل و نظارت بر میزان رضایت مشتری و وفاداری آنها و تشخیص علل تغییرات در رضایت مشتریان - استفاده نمود.

لازم به ذکر است که حوزه استفاده از صنعت تجزیه و تحلیل پیشرفته به صورت نمایی در حال رشد است. بعد از حوادث 11 سپتامبر سال 2001، استفاده از تجزیه و تحلیل داده ها جهت شناسایی جرم و نیز کشف تقلب در بسیاری از سازمانها و موسسات آمریکایی اجباری شده است.

- ارائه بینشهای جامع که در جهت ایجاد فرصت های درآمدی، افزایش کانال های فروش و کاهش هزینه ها
- استفاده از اطلاعات و پیش بینی وضعیت مشتریان امکان ارائه خدمات با کیفیت به مشتریان فراهم خواهد آمد.

۲. متن کاوی (جستجوی متن)

همانطور که در مورد تکنولوژی داده کاوی میدانیم، یکی از زمینه های کاربردی اولیه این فناوری استخراج متن، جمع آوری و ادغام واقعیت ها به عنوان پایه ای برای کمک به تصمیم گیری است. استفاده از فناوری متن کاوی سه مزیت اصلی برای یک کسب و کار بوجود خواهد آورد:

- ✓ توانایی زیاد برای سرعت بخشیدن به خواندن و جستجوی مقادیر زیادی از داده های متنی توسط ماشین نسبت به سرعت خوانندگان انسانی.
- ✓ عینیت بخشی و سفارشی سازی فرآیندها به عنوان مثال: نتایج منحصرأ وابسته است به الگوریتم های پردازش زبانی و محاسبات آماری ارائه شده توسط این فناوری
- ✓ امکان خودکارسازی برخی از کارهای روزمره به ابزارهای داده کاوی و ارائه نتایج به صورت خودکار. با استفاده از این روش میتوان کارهای بیشتری را به کاربران سپرد.

با استفاده از مزایای فوق، برنامه های کاربردی داده کاوی و استخراج متن امروزه و به طور معمول در حوزه های زیر مورد استفاده قرار میگیرند:

- ✓ استخراج اطلاعات مربوطه از یک سند (خلاصه، استخراج ویژگی، و غیره)؛
- ✓ به دست آوردن بینش در مورد روندها، روابط بین افراد / مکان ها / سازمان ها و غیره به طور خودکار و همچنین جمع آوری و مقایسه اطلاعات استخراج شده از انواع خاصی از اطلاعات و یا اسناد (مانند ایمیل های دریافتی، نامه های مشتری، سیم خبری، و غیره)
- ✓ طبقه بندی و سازماندهی اسناد با توجه به محتوای آنها به صورت خودکار، به عنوان مثال، پیش انتخاب گروهی از اسناد با یک موضوع خاص و انتساب آنها به شخص مناسب به صورت خودکار
- ✓ سازماندهی مخازن مربوط به اطلاعات توصیفی مربوط به اسناد برای جستجو و بازیابی سریع اسناد.

فهرست فوق نشان می دهد که تکنولوژی استخراج متن دو حوزه کاربرد اصلی را پوشش می دهد:

- ✓ کشف دانش و اطلاعات مناسب از داده های متنی زیاد
- ✓ تقطیر اطلاعات (عصاره گیری از اطلاعات)

فناوری های مورد استفاده در فرآیند داده کاوی و استخراج متن عبارتند از:

- ✓ استخراج اطلاعات،
- ✓ ردیابی موضوع،
- ✓ خلاصه سازی،
- ✓ طبقه بندی (شناسایی موضوع اصلی یک سند)،
- ✓ خوشه بندی،
- ✓ پیوند مفهوم (ابزار مرتبط با اسناد مرتبط با شناسایی مفاهیم مشترک آنها، کمک به کاربران برای پیدا کردن اطلاعاتی که احتمالاً از طریق روش های جستجوی سنتی (تجسم اطلاعات) پیدا نکرده اند.
- ✓ قرار دادن منابع متنی گسترده در یک سلسله مراتب بصری یا نقشه
- ✓ فراهم کردن قابلیت های مرور اطلاعات،
- ✓ امکان انجام جستجوهای ساده و سوال و جواب

یک روند جدید شامل ادغام داده کاوی و استخراج متن در یک سیستم واحد، ترکیبی معروف به duo-mining است. این ترکیب به خصوص در مدیریت ارتباط با بانکداری و مدیریت کارت اعتباری مشتریان مورد استفاده قرار میگیرد. در این مثال، به جای اینکه قادر به تجزیه و تحلیل اطلاعات ساختار یافته از تراکنشها باشند، آنها می توانند وقایع مربوط به سرویسها و خدمات مشتری را ثبت کرده و الگوهای مربوط به روش خرج کردن مشتریان را با استفاده فناوری استخراج متن پیدا کنند - البته فراتر از روشهای معمول جستجوی اطلاعات -

شرکت هایی که مجموعه های گسترده ایی از اسناد را دارند باید بروی برنامه های کاربردی استخراج متن سرمایه گذاری کنند تا امکان تجزیه و تحلیل اسناد و اطلاعات خود را داشته باشند و امکان ارائه بازخورد اطلاعاتی به ذینفعان را داشته باشند.

کاربرد دیگری از استخراج متن - که دارای محبوبیت زیادی است با نام کشف الکترونیکی نامگذاری شده است. مقوله کشف الکترونیکی ریشه در حوزه حمایت از دعاوی مدنی داشته و با سازماندهی فایل های الکترونیکی با استفاده از متادیتای متصل به آن فایلها، انجام میپذیرد. در این روش با توجه به حجم بسیار بالای فایلها، این فایل ها معمولاً در یک سیستم بازیابی دعاوی قرار می گیرند تا امکان بررسی و تولید اطلاعات با یک روش آسان فراهم آید. از این روش در مدیریت داده های قانونی استفاده میشود.

۳. وب کاوی (جستجوی وب)

امروزه انواع مختلفی از ابزارهای نرم افزاری به منظور جستجوی محیط وب وجود دارد. یکی از ابزار استخراج و تجزیه و تحلیل وب سایتها ، به منظور تحلیل الگوهای رفتاری کاربران در وبسایتها مورد استفاده قرار میگیرد، گزارش هایی که به مدیران وبسایتها کمک میکند تا ساختار وب سایت و ناوبری آن را بهبود ببخشند.

همچنین ابزارهایی هستند که تجزیه و تحلیل رفتار آنلاین از کاربران ارائه می دهند، با این حال، ابزارهای دیگری نیز به منظور تجزیه و تحلیل وقایع نگاری دسترسی جامع به وب را ارائه می دهد.

این ابزارها اجازه می دهد تا ردگیری فعالیت های انجام شده توسط کاربران در سایت را بازه های زمانی مختلف پیگیری نموده و تعداد کل بازدیدها، حجم اطلاعات جابجا شده و میزان نمایش صفحات محبوب سایتها را مشخص نمایند. مطالعات اخیر نشان داده است که محتوای و ساختار یک وب سایت می تواند نقش مهمی در کیفیت توصیه های ارائه شده توسط یک سیستم جستجوی وب ایفا نماید.

سیستمهای جستجوی وب به طور سنتی بر روی موضوع تجارت الکترونیک متمرکز شده اند، جایی که به کاربران وب محصولاتی نظیر کتاب، فیلم، و یا سایر محصولات تجاری را از یک فروشگاه بزرگ از میان هزاران قلم جنس توصیه می کنند. در واقع هدف از چنین سیستمهایی، بهره گیری از اطلاعات از ارائه شده توسط منابع مختلف به منظور ایجاد توصیه های برای خرید توسط کاربر مربوطه است.

سیستم های پیشنهاد کننده در حال حاضر از حوزه های فعال در تحقیق و توسعه میباشند، که عمدتاً به دلیل کاربرد آنها در شرایطی عملی است که کاربر نیازمند به انتخاب یک گزینه از مجموعه ای از گزینه های مناسب است که برای وی منفعت به دنبال دارد. سیستم های پیشنهاد دهنده براساس رویکرد به چهار دسته طبقه بندی شده است:

✓ **رویکرد مبتنی بر محتوا :** در این رویکرد، توصیه ها معمولاً از شباهت محصولاتی است که کاربر با محصولات دیگر که کاربر قبلاً خریداری کرده - و یا به آنها مراجعه کرده و یا بطور کلی مورد توجه وی قرار گرفته است - سرچشمه میگیرد.

- ✓ **رویکرد مبتنی بر تعامل** : در این رویکرد، توصیه ها معمولاً از اندازه گیری شباهت ها بین محصولات انتخاب شده توسط یک کاربر با دیگر کاربران مشابه از نظر رفتاری سرچشمه میگیرد.
- ✓ **رویکرد استفاده از توصیه های رفتاری** : در این رویکرد، توصیه ها بر اساس شباهت رفتار مصرفی کاربران خواهد بود. در این روش معمولاً توصیه ها مربوط به رفتارهای کاربران در مشاهده صفحات و نیز بررسی گزارش رخدادهای مربوط به سرورهای وب مربوط به کاربران میباشد.
- ✓ **رویکرد ترکیبی Hybrid** : همانطور که از نام این روش پیداست میتواند از ترکیب هر یک از سه روش فوق میتواند باشد. باید دقت داشت که برای اجرای صحیح رویکردهای مبتنی بر محتوا و مبتنی بر تعامل نیاز به در دسترس بودن پروفیل های کاربر و نظرات صریح آنها درباره محصولات وجود دارد.

پیوست شماره دو

تهدیدهای پیش روی صنعت تجزیه و تحلیل اطلاعات

۱. نداشتن مشتری

در حالیکه داده کاوی و انبارداری داده ها فناوری بسیار قدرتمند میباشند، اما نمونه شکست در سازمانها نیز کم نیست. برای اطمینان از این که نتایج تجزیه و تحلیل داده به طور مناسب مورد استفاده ذینفعان (مشتریان) قرار گیرد بسیار مهم است که از همه واحدهای کسب و کار "خریدار اطلاعات" وجود داشته باشید. برای اطمینان یافتن از این موضوع میبایست تعدادی از مراحل به شرح زیر انجام شود:

- ✓ برای اجرای یک پروژه تجزیه و تحلیل پیشرفته اطلاعات در سازمان نیاز به ایجاد یک تیم منسجم از واحدهای مختلف سازمان با توانمندیهای مختلف میباشد.

✓ در حالیکه اعضای فنی تیم، نقش مهمی ایفا می کنند اما مدیران بخش هسته کسب و کار باید مالک و حامی پروژه بوده و تیم را هدایت نمایند.

✓ فرایندها باید با مجموعه ای مناسب از معیارهای قابل قبول مورد اندازه گیری و سنجش قرار گیرند تا اطمینان حاصل شود که در تمام مراحل پروژه با موفقیت در حال انجام است.

۲. پیاده سازی تجزیه و تحلیل پیشرفته

معرفی و استفاده از تجزیه و تحلیل پیشرفته در یک سازمان باید با دقت انجام شود. با در نظر گرفتن هزینه اولیه این پروژه و همچنین اعمال تغییرات قابل توجه در فرایندهای سازمانی، احتمال بالا رفتن هزینه ها در صورت عدم توجه کافی به تعریف و اجرای درست پروژه وجود دارد.

عدم وجود مزایا و کاربردهای کافی برای تجزیه و تحلیل اطلاعات میتواند به دلسردی در سازمان بوجود آورد. بهتر است بجای یک روش تغییرات بزرگ و سریع از رویکرد تغییرات کم و پیوسته استفاده نماید. زیرا با مشخص شدن مزایا و منافع بخشهای اولیه از اجرای پروژه، برای تکمیل و ادامه فعالیتهای به عنوان اعتبار استفاده نمود.

این پروژه باید در یک حوزه هایی تعریف شود که بتواند بالقوه اثر گذار باشد به عبارت دیگر در جایی مورد استفاده قرار گیرد که روندها و رویه های فعلی بسیار بی اثر است.

۳. مقررات و حفظ حریم خصوصی

حریم خصوصی غالباً با حفاظت از اطلاعات شخصی مانند نام، آدرس، شماره امنیت اجتماعی، شماره کارت اعتباری، پرونده پزشکی و پرونده های مالی مرتبط است. در زمینه تجزیه و تحلیل کسب و کار، تجاوز به حریم خصوصی افراد زمانی رخ می دهد که در حین تجزیه و تحلیل اطلاعات و یا ارائه گزارشات احتمالی، هویت یک فرد یا اطلاعات خاص، با ارزش و حساس یک یا چند فرد به نحوی آشکار شود که موقعیت اجتماعی و یا مالی و ... به خطر بیفتد.

بطور کلی آشکار شدن هر گونه اطلاعاتی که مطابق با قانون و عرف جامعه مناسب نباشد نقض حریم خصوصی محسوب میگردد.

افشای حریم خصوصی در سازمانها هنگامی رخ می دهد که اطلاعات مربوط به عملیات تجاری شرکت مثلاً اطلاعات مربوط به مشتریان، قیمتها، محدوده محصولات، سیاستها و استراتژی های سازمانها برای طرفهای تجاری غیر مجاز (مانند رقبا و ...) افشا گردد.

در حال حاضر، نیاز به اطلاعات دقیق تر و صیقل داده شده در حوزه اطلاعات مالی و پزشکی - که باید توسط قوانین مورد حفاظت قرار گیرند - در حال گسترش میباشد.

دفاتر مجاز و غیر مجاز دولتی یا غیر دولتی محدودیتهایی را تنظیم کردهاند و در انتظار نرم افزارهای ضد هرزنامه میخواهند همان کار را در سطح الکترونیکی انجام دهند.

با توجه به تهدیدها و سرقت هویت و مسائل مربوط به امنیت اجتماعی، مشتریان از پیام های بازاریابی - که ظاهراً بر علایق شخصی افراد متمرکز هستند- هراسان بوده و چندان رضایت به مشارکت در فرآیندهای تجزیه و تحلیل اطلاعات ندارند.

۴. مشکلات ناشی از فناوریهای تجزیه و تحلیل پیشرفته

- پیچیدگی مفاهیم و همچنین نحوه استفاده از ابزارها و روشهایی است که منجر به تولید ارزش برای مشتریان و ذینفعان گردد.
- نبود نیروهای متخصص در بسیاری از سازمانها، با توجه به اینکه استفاده از اطلاعات یک تخصص و یک علم جدید بوده و استفاده از ابزارهای فناوری کامپیوتری یک موضوع جداگانه است لذا باید افراد خاصی تعلیم و آموزش ببینند که بتوانند هر دو مفهوم را درک کنند.
- به عبارت دیگر دانستن علم تحلیل داده به تنهایی ضامن استفاده از ابزارها نبوده و تسلط بر ابزارها نیز لزوماً امکان تولید ارزش برای مشتریان را نخواهد داشت.
- چالش مربوط به تغییرات سریع در ابزارها و نیز متغیر بودن آینده این صنعت است که باعث شده است تا پیش بینی روند آن کمی سخت شود.
- مشخص کردن فرمت خروجی تحلیل پیشرفته اطلاعات دیگر مشکل این حوزه است. ورودی تجزیه و تحلیل پیشرفته معمولاً حجم بسیار زیادی از اطلاعات است اما خروجی آن باید خلاصه، مفید، قابل خواندن و قابل فهم برای ذینفعان باشد.
- طراحی و مدلسازی یک سیستم که داده های ورودی بسیار زیادی را بگیرد و یک خروجی ارزشمند برای مشتریان ایجاد کند بسیار اهمیت دارد. استفاده از ابزارهای پشتیبان شامل داشبوردها، گزارشات و سیستمهای مصور سازی اطلاعات در تهیه خروجیهای موثر میتوانند کارساز باشند.

۵. فراهم سازی و به اشتراک گذاری داده ها بین سازمانها

در واقع، داده ها بلاکهای کلیدی برای سیستمهای که بر مدلسازی و تحلیل تمرکز دارند. داده های داخلی - یا اصطلاحاً داده های خانگی - نکته اصلی در تمایز در فرآیند هوش تجاری میباشد. ابزارهای تحلیلی بسیار پیشرفته نیز میتوانند نمایشی غیر موثر از اطلاعات داشته باشند اگر داده مناسب فراهم نشده باشد و یا فقر اطلاعاتی در سازمان وجود داشته باشد.

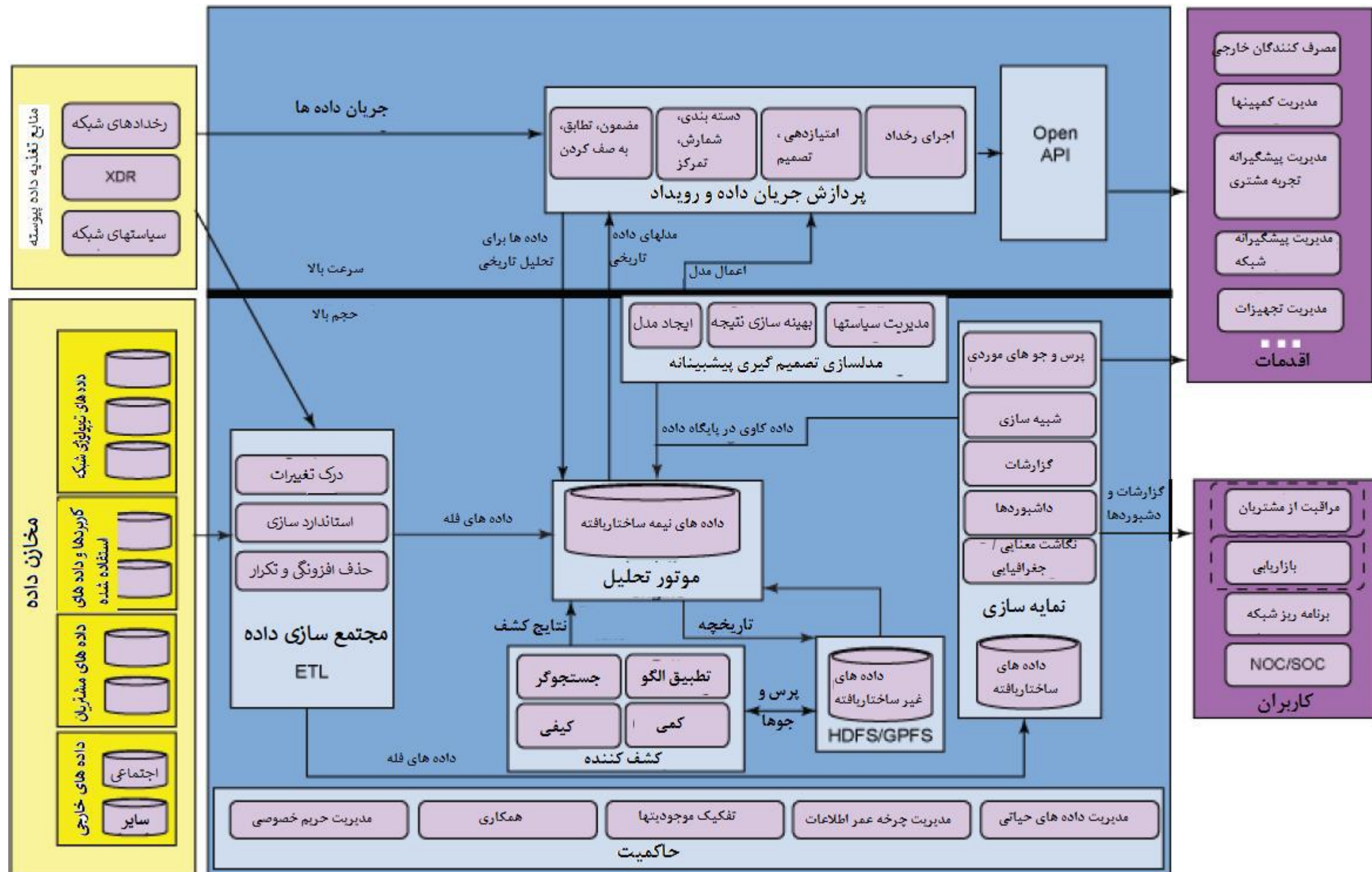
به منظور استفاده واقعی از امکانات تجزیه و تحلیل پیشرفته، سازمانها باید اطلاعات دقیقی از نیازها، ارزشها و نیز خواسته های مشتریان خود داشته باشند.

سازمانهای پیشرو، اطلاعات مشتریان را از مجاری مختلف و منابع خارجی تهیه میکنند.

چالش دیگر، به اشتراک گذاری اطلاعات بین سازمانها میباشد زیرا امنیت اطلاعات برای همه سازمانها مهم است. حفظ حریم خصوصی و محرمانگی اطلاعات در مورد برخی داده های حساس بسیار مهم است. لذا بکار بردن داده هایی که دارای نگارش و مفاهیم گوناگونی هستند علاوه بر مسائل امنیتی ، مسائلی از قبیل همگن سازی داده ها نیز یک چالش دیگر میباشد.

پیوست سوم - معماری و اجزای مربوط به پلتفرم تجزیه و تحلیل پیشرفته داده

معماری سطح بالای این پلتفرم در شکل زیر نشان داده شده است که چگونه اجزاء در یک پلت فرم پیشرفته تجزیه و تحلیل جمع می شوند.



این پلتفرم شامل اجزای اصلی زیر است:

○ زیرسیستم پردازش جریان داده :

این ماژول به صورت بلادرنگ جریان داده های ورودی را مورد پردازش و تحلیل قرار میدهد. در ادامه یک نمونه از کمپینهای هوشمند را با استفاده از اطلاعات موقعیت مکانی خواهید دید. دستگاههای موبایل در حدود یک میلیارد تراکنش در روز تولید میکنند به منظور انجام یک کمپین تبلیغاتی برای یک فروشگاه خاص میتوان از این اطلاعات استفاده نمود. پردازش جریان داده با استفاده از سه ماژول زیر انجام میشود:

■ ماژول اول برای درک مضمون ، شناخت و تنظیم داده های ورودی نیاز است. این قابلیت ها باعث میشود اطلاعات مربوط به داده های تراکنش های به مشتریان یا رویدادهای شناخته شده پیوند داده شوند، علاوه بر آن میتوان از پردازش تراکنش ها برای شناسایی زمینه ها و فعالیت های کاربر نیز استفاده کرد.

■ ماژول دوم به منظور طبقه بندی ، شمارش و تمرکز مورد استفاده قرار میگیرد، این ماژول به صورت بلادرنگ ویژگیها را به داده های ورودی تخصیص میدهد. به عنوان مثال، " دسته بندی تماسهایی مشتریانی که بیش از دو بار در یک ساعت از دسته رفته اند" و یا کاربرانی که بروی یک تبلیغ کلیک کرده و به سایت اصلی آگهی دهنده مراجعه نموده اند.

این ماژولها به منظور طبقه بندی و شمارش اطلاعات از مجموعه ای از پارامترهای پویا استفاده میکنند. علت پویا بودن این پارامترها به این دلیل است که به طور مداوم و بر اساس تجزیه و تحلیل عمیق بر روی داده های گذشته – بروز رسانی می شوند. به عنوان مثال، داده های گذشته ممکن است کشف کنند که احتمال لغو تماس مشتریان پس از دو تماس، کاهش پیدا کرده و یا افزایش یافته است.

■ ماژول سوم برای امتیاز دهی و تصمیم گیری مورد استفاده قرار میگیرد. یک مجموعه از مدل های امتیاز دهی جدید معرفی شود، این مدل های امتیازدهی از طریق مدلسازی پیش بینانه و با استفاده از داده های گذشته کشف و معرفی میشوند. این مدل های امتیاز دهی میتوانند بروی داده های ورودی به صورت بلادرنگ اعمال شده و به منظور تصمیم گیری از آنها استفاده گردد. علاوه بر این، درخت های تصمیم گیری پیچیده و یا سایر استراتژی های مبتنی بر قاعده میتوانند از این مدل های امتیاز دهی در زمان اجرا و از طریق یک سیستم مدیریت قواعد تجاری دریافت نموده و از آنها استفاده کنند.

○ زیر سیستم مدلسازی تصمیم گیری پیشبینانه :

این ماژول در واقع یک موتور مدلسازی آماری است که یکسری مدل‌های تصمیم گیری را با استفاده از تحلیل داده های گذشته خلق میکند. سپس این ماژول از این مدل‌ها برای ردگیری مدل‌های موفقیت و شکست در نتایج پیش بینی شده استفاده میکند. پس از بررسی نتایج تصمیم گیری‌های قبلی، این ماژول مدل‌های تصمیم گیری نا موفق را با مدل‌های بهتر جایگزین میکند. در این زیر سیستم، یک ماژول دیگر بنام ایجاد مدل وجود دارد که برای توسعه مدل‌های پیش بینی - با استفاده از داده های گذشته - در معماری تعبیه شده است. از این مدل‌های پیش بینی میتوان در مورد پردازش جریان داده های بلادرنگ و یا داده های فله استفاده کرد.

نتایج حاصل از اجرای مدل با استفاده از یک ماژول بنام بهینه سازی نتیجه، مدل های پیش بینی را با یکدیگر مقایسه نموده سپس بهترین آنها را انتخاب و از آن برای موارد بعدی استفاده خواهد نمود.

در تمام زمانها، ماژول مدلساز پیش بینی میتواند صدها مدل پیش بینی را ایجاد نماید و بطور مداوم این مدل‌ها را در فرآیندهای واقعی مورد آزمایش قرار دهد و سپس با بهینه سازی این مدل‌ها بهترین نتایج را بدست بیاورد.

○ زیر سیستم موتور تجزیه و تحلیل :

یک انبار داده موجود در پلتفرم (که در بخش موتور تجزیه و تحلیل به عنوان داده های نیمه ساختار یافته شناخته میشود) می تواند پرس و جو های پیشرفته یی را اجرا کند، بنابراین می توان تمام مدل های پیش بینی و توابع نمایه سازی (مصور سازی) در این زیر سیستم انجام داد.

داده های ذخیره سازی شده معمولاً آنقدر بزرگ و حجیم هستند که نمیتوان آنها را برای مدلسازی پیشبینانه و یا مصور سازی به سیستمهای خارجی انتقال داد. لذا در برخی موارد این موتور میتواند فعالیتهای مربوط به مصور سازی و مدلسازی پیشبینانه را با استفاده از دستوراتی که از دیگر ابزارها دریافت میکند نیز انجام دهد.

این دستورات معمولاً به صورت توابع پایه ایی و شناخته شده برای همه ابزارها به عنوان مثال، دستورات SQL منتقل می شوند، و در محیط سخت افزاری تخصصی برای مواجهه با داده های حجم بالا اجرا می شوند.

موتورهای تحلیلی به منظور ELT از توابع معمولی استفاده میکنند. ELT شامل سازماندهی داده های وارد شده به سیستم با استفاده از تبدیل داده ها به فرمتهای دیگر و همچنین اجرای مدل‌های پیش بینی و گزارشات است. علاوه بر این موتورهای تحلیلی برای سایر کارهای دیگر نیز - مانند تحلیل‌های جغرافیایی - مورد استفاده قرار میگیرد.

○ زیر سیستم کشف :

این زیرسیستم در واقع داده های با تنوع بسیار بالا را دریافت نموده و در جستجوی الگوهای کمی و یا کیفی موجود در اطلاعات است. این ابزارهای کشف شامل جستجو عمومی یا خاص در سراسر داده های غیر ساخت یافته هستند. آنها از ابزارهای تخصصی یادگیری ماشین برای تشخیص الگو استفاده میکنند و همچنین میتوانند علم هستی شناسی را در حوزه های مختلف به کار گیرند تا وظیفه خود را بطور هوشمندانه انجام دهند.

در عصر انفجار اطلاعات، این ابزارها پیشرفتهای خارق العاده ایی در تجزیه و تحلیل کیفی داده های غیر ساخت یافته مانند شبکه های اجتماعی و یا وبلاگها داشته اند.

نتایج تجزیه و تحلیل داده ها میتوانند شامل تعیین مشخصات یک داده باشد البته پس از کشف الگوها و ویژگی داده ها، آنها را برای تجزیه و تحلیل بیبشتر به موتور تجزیه و تحلیل منتقل خواهند شد. علاوه بر این، نتایج کشف شده در این زیر سیستم میتواند در گزارش های کیفی مانند نقشه برداری معنایی و تجسم جغرافیایی مورد استفاده قرار گیرد .

○ زیر سیستم نمایه (مصور) سازی:

سیستمهای تجزیه و تحلیل همواره از یک ماژول به منظور نمایش دادن نتایج تحلیلی اطلاعات در فرمتهای مختلف استفاده میکنند. این ماژول شامل گزارشات ساختار یافته، داشبوردهای اطلاعاتی، نمایش جغرافیایی یا معنایی اطلاعات و یا شبیه سازی آنها است.

تکنیک های مصور سازی امکانات قابل توجهی را برای تحلیلگران داده فراهم می کنند، بطوریکه آنها میتوانند داده ها را به قطعات کوچک تر تبدیل نمایند که بر مبنای پارمترهای از پیش تعیین شده و یا موردی تعریف شده اند.

○ زیر سیستم فایل توزیع شده HADOP/ سیستم فایل موازی:

کلان داده ها شاهد تکامل سیستم فایل توزیع شده Hadoop (HDFS) بوده اند. این سیستم مدیریت فایل توزیع شده ، یک مکانیزم ذخیره سازی فایل متن باز است که توسط چندین ارائه دهنده بزرگ خدمات کامپیوتری از جمله IBM پشتیبانی میشود.

یکی از ارزشهای ایجاد شده توسط HDFS ، توانایی آن در استفاده از معماری زیرساخت تجزیه و تحلیل داده ها، به منظور ذخیره سازی داده های دارای افزونگی در تعدادی بسیار زیادی از زیرساختها و ابزارهای دیگر و همچنین امکان اجرای پرس و جوهای پیچیده است، از دیگر مزایای این زیر سیستم میتوان به کشف داده های بدون ساختار بصورت بسیار موثر اشاره نمود.

HDFS یک جایگزین جدی برای انبار داده های ساخت یافته برای ذخیره سازی، بازیابی و تجزیه و تحلیل داده های بزرگ ارائه می کند. IBM در حال کار بر روی یک پلتفرم مدیریت فایل با کارایی بالا به نام GPFS است که جایگزینی برای HDFS است

○ زیر سیستم مجتمع سازی داده و حاکمیت:

پلتفرم تجزیه و تحلیل داده ها، دارای تعدادی از محلهای ذخیره سازی - داده های با ساختار و بدون ساختار که بصورت بلادرنگ وارد پلتفرم میشوند - است. علاوه بر موارد فوق، این پلتفرم، به یک مجموعه یکپارچه از ابزارها برای مجتمع سازی داده ها متنوع و گوناگون و همچنین یک بخش مجزا برای اجرای حکم در زمینه های کلیدی نیاز دارد. در نهایت چنین پلتفرمهایی، باید ماژولهای استانداردی به شرح ذیل نیز داشته باشند:

- مدیریت داده های حیاتی
- مدیریت چرخه عمر داده ها
- مدیریت دسترسی به اطلاعات
- حفظ حریم خصوصی
- بهبود کیفیت داده ها

فصل هشتم

منابع

<https://www.gartner.com/it-glossary/advanced-analytics/>

<https://bi-survey.com/predictive-analytics>

<https://rapidminer.com/glossary/advanced-analytics-vs-bi/>

<https://searchbusinessanalytics.techtarget.com/definition/advanced-analytics>

<https://bi-survey.com/predictive-analytics>

https://www.marketsandmarkets.com/Market-Reports/advanced-analytics-market-58104148.html?gclid=EAlaQobChMItpXft9eW4gIVAhUYCh3q5QC_EAAYASAAEgLRHfD_BwE

<https://www.ibm.com/developerworks/library/ba-adv-analytics-platform1/index.html>

<https://www.dataversity.net/advanced-analytics-exploration-transformative-trends/#>

<https://www.cio.com/article/2451762/11-market-trends-in-advanced-analytics.html>

<https://medux.com/advanced-analytics-are-the-future-of-the-telecom-industry/>

<https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf>

Advanced analytics:opportunities and challenges - Ranjit Bose
Anderson School of Management, University of New Mexico, Albuquerque, New Mexico, USA

Adomavicius, G. and Tuzhilin, A. (2005), "Towards the next generation of recommender systems: a survey of the state-of-the-art and possible extensions", IEEE Transactions of Knowledge and Data Engineering, Vol. 17.

Apte, C.V., Hong, S.J., Natarajan, R., Pednault, E.P.D., Tipu, F.A. and Weiss, S.M. (2003), "Data-intensive analytics for predictive modeling", IBM Journal of Research & Development, Vol. 47 No. 1, pp. 17-23.

<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.678.6717&rep=rep1&type=pdf>

<https://www.bobsguide.com/guide/news/2015/Jul/21/big-data-challenges-risks-and-solutions/>

https://www.researchgate.net/publication/303363742_A_SWOT_analysis_of_big_data

<http://www.bigdatanalysis.com/what-set-of-skills-you-need-to-do-business-analysis/>

<http://openbudgets.eu/assets/deliverables/D2.3.pdf>

