



Namatek
True Education

Big Data

www.namatek.com

آشنایی با فناوری کلان
داده یا بیگ دیتا

فهرست مطالب

۱. بیگ دیتا یا کلان داده چیست؟ (Big Data)
۲. تاریخچه Big data
۳. انواع بیگ دیتا یا کلان داده
۴. خصوصیات بیگ دیتا
۵. تجزیه و تحلیل کلان داده
۶. فناوری بیگ دیتا چیست؟
۷. مزایای کلان داده چیست؟

اگر شما هم مشغول فعالیت در یک سازمان بزرگ باشید متوجه این شده اید که با انبوه زیادی از داده ها رو به رو هستید؛ اما به صورت دقیق مفهوم بیگ دیتا یا کلان داده چیست؟ داده های کلان را می توان برای بینش ها و آگاهی هایی که منجر به تصمیم گیری بهتر و حرکت های استراتژیک تجاری می شوند، تجزیه و تحلیل کرد. اگر به دنبال تحول در تجارت خود هستید و نیاز به دانش برای تصمیمات استراتژیک دارید، با ما همراه باشید تا در این مقاله به بررسی دیتای کلان و ویژگی های آن بپردازیم.

#۱ بیگ دیتا یا کلان داده چیست؟ (Big Data)



Big Data به مجموعه های متنوع و گسترده ای از اطلاعات اشاره دارد که ممکن است حجم زیادی داشته باشند و با سرعت فزاینده ای رشد کنند. بیگ دیتا ها معمولاً حاصل داده کاوی هستند و در قالب های مختلفی به دست می آیند. حجم این مجموعه داده ها به قدری زیاد است که مدیریت آن ها توسط نرم افزارهای سنتی پردازش داده امکان پذیر نیست؛ اما این داده های حجیم می توانند برای مرتفع کردن مشکلات تجاری مورد استفاده قرار گیرند که قبلاً این مشکلات قابل حل نبودند. کلان داده ها می تواند به عنوان مجموعه داده هایی تعریف شوند که اندازه یا نوع آن ها فراتر از توانایی پایگاه های ارتباطی سنتی برای ضبط، مدیریت و پردازش داده ها به صورت بلادرنگ است. از مشخصات بیگ دیتا ها می توان به حجم، سرعت و تنوع زیاد اشاره کرد.

منابع Big Data ها نسبت به داده های سنتی پیچیده تر می شوند؛ زیرا توسط هوش مصنوعی، دستگاه های تلفن همراه، رسانه های اجتماعی و اینترنت اشیاء هدایت می شوند. به عنوان مثال، انواع داده ها از حسگرها، دستگاه ها، فیلم/صدا، شبکه ها، پرونده های گزارش، برنامه های معاملاتی، وب و رسانه های اجتماعی نشأت می گیرند که بیشتر آن ها در زمان واقعی (real-time) و در مقیاس بسیار بزرگ تولید می شوند.

۲# تاریخچه Big data



اگرچه مفهوم کلان داده تقریباً نوین است؛ اما پیدایش Big Data به دهه ۱۹۶۰ و ۱۹۷۰ برمی گردد. زمانی که دنیای داده ها به مرور با ایجاد اولین دیتا سنترها (Data Centre) و پایگاه داده رابطه ای (Relational Database) آشنا می شد.

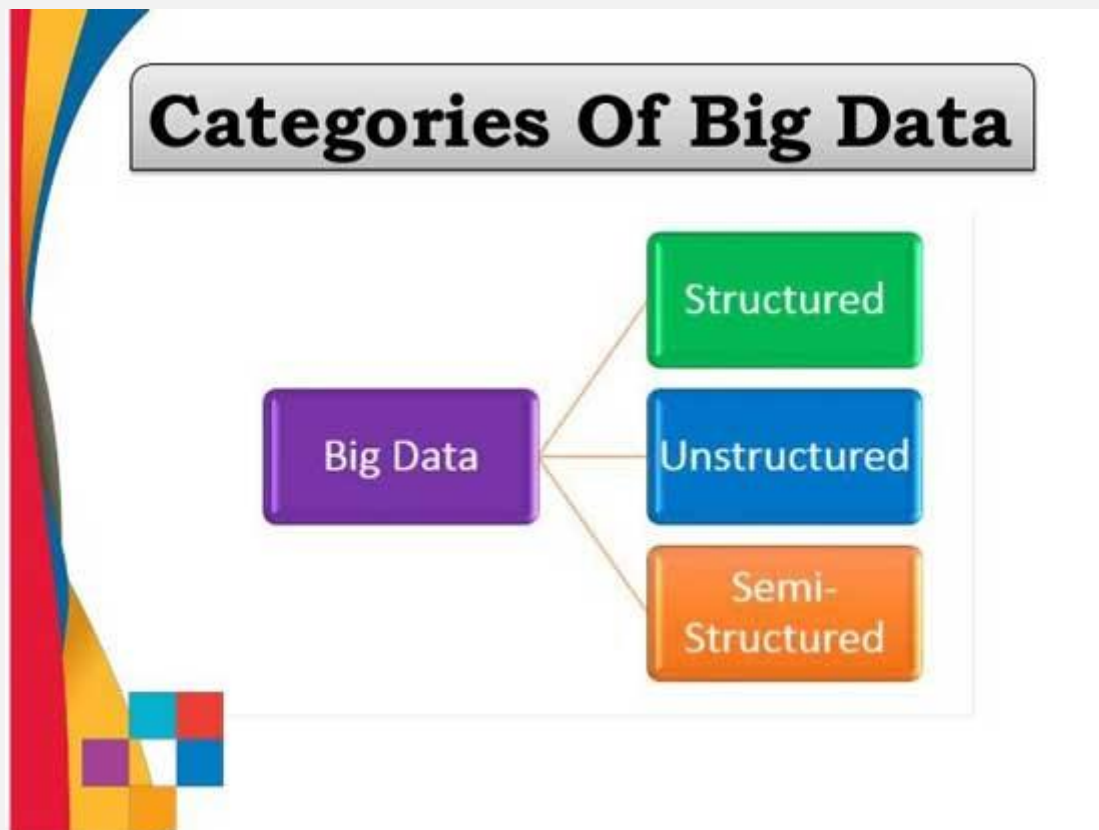
تقریباً از سال ۲۰۰۵، رفته رفته میزان تولید داده های کاربران از طرق مختلفی مانند فیس بوک، یوتیوب و سایر خدمات آنلاین توسط مردم درک شد. (Hadoop یک چهارچوب منبع باز که به طور خاص برای ذخیره و تجزیه و تحلیل مجموعه Big Data ایجاد شده است) در همان سال ساخته شد. محبوبیت NoSQL نیز در این مدت به مرور شروع شد. توسعه یافتن چهارچوب های متن باز مانند Hadoop (و جدیداً Spark) برای رشد و توسعه Big Data لازم و ضروری بود؛ زیرا کار با آن ها را سهل تر و ذخیره آن ها را مقرون به صرفه تر می کند.

در سال های پس از آن، حجم کلان داده ها سر به فلک کشیده است. کاربران هنوز هم مقدار زیادی داده تولید می کنند؛ اما این فقط افراد نیستند که این کار را انجام می دهند. با ظهور اینترنت اشیا، اشیا و دستگاه های بیشتری به اینترنت متصل می شوند و داده های مربوط به الگوی استفاده مشتری و عملکرد محصول را جمع می کنند. این در حالی است که ظهور یادگیری ماشین هنوز اطلاعات زیادی تولید نکرده است. درحالی که بیگ دیتا ها بسیار زیاد به دست آمده اند؛ اما سودمندی آن ها تازه آغاز شده است. رایانش ابری امکانات Big Data را حتی بیش از پیش گسترش داده است. ابر (Cloud) مقیاس پذیری قابل ارتجاعی را ارائه می دهد، جایی که توسعه دهندگان می توانند بخش های موردی را برای آزمایش زیرمجموعه ای از داده ها به صورت خوشه ای جمع کنند. همچنین پایگاه داده ای نموداری از اهمیت بیشتری برخوردار می شود.

#۳ انواع بیگ دیتا یا کلان داده

تا به اینجا به خوبی متوجه شدیم کلان داده چیست و از چه زمانی این مفهوم رایج شده است، در ادامه بررسی می کنیم که آیا انواع متنوعی از کلان داده ها وجود دارد؟

اکنون که در جستجوی کلان داده هستیم، بیایید نگاهی به انواع Big Data بیندازیم:



- ساختاریافته

ساختاریافته به اطلاعات بسیار سازمان یافته ای گفته می شود که می تواند به آسانی و یکپارچه از طریق پایگاه داده توسط الگوریتم های موتور جستجوی ساده ذخیره شده و به آن ها دسترسی پیدا کنند.

مثلا ساختار جدول اطلاعات کارمندان موجود در پایگاه داده یک شرکت به گونه ای است که مشخصات کارمند، موقعیت شغلی وی، حقوق و غیره به صورت سازمان یافته در آن وجود دارد.

- بدون ساختار

داده های بدون ساختار به داده هایی گفته می شود که فاقد هرگونه فرم یا ساختار خاصی باشند. این امر پردازش و تحلیل داده های بدون ساختار را

بسیار دشوار و وقت گیر می کند. ایمیل نمونه ای از داده های بدون ساختار است. ساختاریافته و غیرساختاری دو نوع مهم داده کلان هستند.

- نیمه ساختاریافته

نیمه ساختاریافته سومین نوع Big Data است. داده های نیمه ساختاری مربوط به داده های حاوی هر دو قالب ذکر شده در بالا، یعنی داده های ساختاریافته و غیرساختاری هستند. به طور دقیق، به داده هایی اشاره دارد که اگرچه در یک مخزن خاص (مانند پایگاه داده) طبقه بندی نشده اند؛ اما حاوی اطلاعات بسیار مهم یا برجسته هایی هستند که عناصر جداگانه را در داده ها تفکیک می کنند.

#۴ خصوصیات بیگ دیتا

در سال ۲۰۰۱، داگ لنی، تحلیلگر گارتنر، سه خصوصیت Big Data که به عنوان ۷۳ شناخته می شوند را فهرست کرد. این خصوصیات تنوع، سرعت و حجم بودند. بیایید درباره خصوصیات کلان داده ها بحث کنیم.



• تنوع

تنوع در Big Data به داده های بدون ساختار، ساختاریافته و نیمه ساختاری گفته می شود که از چندین منبع مختلف جمع آوری شده اند. در حالی که در گذشته داده ها فقط از صفحه گسترده و پایگاه داده جمع آوری می شدند، امروزه داده ها به صورت آرایه ای از جمله ایمیل، PDF، عکس، فیلم، صدا، پست و موارد دیگر ارائه می شوند. تنوع یکی از خصوصیات مهم کلان داده ها است.

• سرعت

مفهوم سرعت در اصل به سرعت ایجاد داده ها در زمان واقعی اشاره دارد. در چشم انداز وسیع تر، این موضوع شامل تغییر، پیوند دادن مجموعه داده های دریافتی با سرعت های مختلف و انفجارهای فعالیت است.

• حجم

حجم یکی از ویژگی های کلان داده است. ما از قبل می دانیم که Big Data نشان دهنده حجم عظیمی از داده ها است که به طور روزانه از منابع مختلف

مانند سیستم عامل های رسانه های اجتماعی، فرآیندهای تجاری، ماشین آلات، شبکه ها، تعاملات انسانی و غیره تولید می شود. چنین مقدار زیادی از داده ها در انبارهای داده ذخیره می شوند.

#۵ تجزیه و تحلیل کلان داده



تجزیه و تحلیل کلان داده، استفاده از تکنیک های پیشرفته تحلیلی در برابر مجموعه های کلان داده بسیار متنوع است که شامل داده های ساختاریافته، نیمه ساختاریافته و بدون ساختار، از منابع مختلف و در اندازه های مختلف از ترابایت تا زتابایت است. با تجزیه و تحلیل بیگ دیتا ها، در نهایت می توانید تصمیم گیری بهتر و سریع تر داشته باشید، مدل سازی بهتری انجام دهید، پیش بینی نتایج دقیق تری داشته باشید و هوش تجاری را تقویت کنید.

همان طور که استراتژی کلان داده خود را ایجاد می کنید، نرم افزارهای منبع باز مانند Apache Hadoop، Apache Spark و کل اکوسیستم Hadoop را به عنوان ابزارهای انعطاف پذیر و مقرون به صرفه ذخیره سازی و پردازش داده ها در نظر بگیرید که برای حجم داده های تولیدشده امروزی طراحی شده اند.

#۶ فناوری بیگ دیتا چیست؟

در پاسخ به فناوری کلان داده چیس می توان آن را به عنوان یک نرم افزار کمکی تعریف کرد که برای تجزیه و تحلیل، پردازش و استخراج اطلاعات از یک مجموعه داده بسیار پیچیده و بزرگ طراحی شده است که نرم افزار پردازش داده های سنتی هرگز نمی تواند با آنها سر و کار داشته باشد. ما برای تجزیه و تحلیل این مقدار عظیم از داده های زمان واقعی و ارائه نتیجه گیری و پیش بینی برای کاهش خطرات در آینده، به فناوری های پردازش Big Data احتیاج داریم.

حال بیایید نگاهی به دسته بندی هایی بیندازیم که فناوری های Big Data در آن ها طبقه بندی شده اند.



انواع فناوری های بیگ دیتا عبارتند از:

#۶-۱ فناوری های عملیاتی Big Data

اولاً، عملیات بیگ دیتا ها مربوط به داده های عادی هستند که ما روز به روز تولید می کنیم. این نوع فناوری می تواند مربوط به معاملات آنلاین، رسانه های اجتماعی یا داده های یک سازمان خاص و غیره باشند. حتی می توانید نتیجه این فناوری را نوعی داده خام در نظر بگیرید که برای تغذیه فناوری های تجزیه و تحلیل Big Data استفاده می شود.

چند نمونه از فناوری های عملیاتی Big Data به شرح زیر است:

- رزرو بلیط آنلاین که شامل بلیط قطار، بلیط پرواز، بلیط فیلم و غیره است.

- خرید آنلاین که معاملات آمازون، Walmart، Flipkart، Snapp و بسیاری موارد دیگر است.
- داده های سایت های رسانه های اجتماعی مانند فیس بوک، اینستاگرام، برنامه های کاربردی و موارد دیگر.
- مشخصات کارمندان هر شرکت چند ملیتی.

بنابراین اجازه دهید با استفاده از این فناوری به بررسی فناوری های تحلیلی Big Data بپردازیم.

#۶-۲ فناوری تحلیلی Big Data

فناوری تحلیلی Big Data مانند نسخه پیشرفته فناوری Big Data می باشد و کمی پیچیده تر از فناوری عملیاتی Big Data است. به طور خلاصه، فناوری تحلیلی Big Data جایی است که قسمت واقعی عملکرد به تصویر در می آید و تصمیمات مهم و واقعی بیزینس با تجزیه و تحلیل عملیاتی بیگ دیتاها اتخاذ می شوند.

چند نمونه از فناوری های تجزیه و تحلیل Big Data به شرح زیر است:

- بازاریابی سهام
- انجام مأموریت های فضایی با اطلاعات مهم موجود در آن ها
- اطلاعات پیش بینی آب و هوا
- زمینه های پزشکی کنترل کننده وضعیت سلامت بیماران خاص

#۷ مزایای کلان داده چیست؟



احتمالا با توجه به اینکه توضیح داده شد مفهوم کلان داده چیست و چه انواعی دارد به خوبی این را متوجه خواهید شد که یکی از بزرگترین مزایای Big Data آنالیز پیش بینی است. ابزارهای تجزیه و تحلیل کلان داده ها می توانند نتایج را به طور دقیق پیش بینی کنند. بدین ترتیب به کسب و کار ها و سازمان ها اجازه تصمیم گیری بهتر را می دهند. در عین حال کارایی عملیاتی خود را بهینه می کنند و خطرات را کاهش می دهند.

با استفاده از داده های جمع آوری شده از سیستم عامل های رسانه های اجتماعی که توسط ابزار تجزیه و تحلیل Big Data انجام می شوند، مشاغل در سراسر جهان، استراتژی های بازاریابی دیجیتال خود را برای افزایش تجربه کلی مصرف کننده ساده می کنند.

Big Data اطلاعاتی را درباره نقاط درد مشتری ارائه می دهد و به شرکت ها امکان می دهد محصولات و خدمات خود را بهبود بخشند. تجزیه و

تحلیل Big Data می تواند به شرکت ها کمک کند تا منجر به فروش بیشتر آن ها شود که به طور طبیعی به معنای افزایش درآمد است. مشاغل از ابزارهای تجزیه و تحلیل Big Data استفاده می کنند تا درک کنند که محصولات یا خدمات آن ها در بازار چگونه کار می کنند و مشتریان چگونه به آن ها پاسخ می دهند؛ بنابراین بهتر می توان درک کرد که بهتر است کجا وقت و پول خود را سرمایه گذاری کنند.

با بینش حاصل از Big Data، همیشه می توانید یک قدم جلوتر از رقبا باشید. شما می توانید بازار را بشناسید تا بدانید رقبای شما چه نوع تبلیغات و پیشنهادی را ارائه می دهند و سپس می توانید پیشنهادهای بهتری برای مشتریان خود ارائه دهید.