

(استفاده از داده‌های بزرگ: تجزیه و تحلیل، چالش‌ها و نوآوری‌ها در صنایع مختلف)

یوسف رحیمی آخوندزاده

گروه مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه آزاد اسلامی، واحد قزوین، ایران

چکیده

این مقاله به بررسی تأثیر تحولات مربوط به تجزیه و تحلیل داده‌های بزرگ در صنایع مختلف، برجسته کردن نقش آنها در نوآوری‌ها، کارآیی عملیاتی و تصمیم‌گیری‌های آگاهانه می‌پردازد. در ابتدا با مروری بر اهمیت داده‌های بزرگ، به برنامه‌های کلیدی مربوط به حوزه‌های مراقبت‌های بهداشتی، مالی و شهرهای هوشمند می‌پردازیم و نشان می‌دهیم که چگونه بینش‌های داده محور در حال تغییر شکل این حوزه‌ها می‌باشند. با همه‌ی اینها، مسیر اجرای مؤثر تجزیه و تحلیل در این حوزه‌ها مملو از چالش‌هایی از جمله حریم خصوصی داده‌ها، پیچیدگی‌های ادغام فناوری‌ها در صنایع مختلف، و همچنین شکاف‌های مهارتی در حوزه‌های تخصصی داده‌های بزرگ می‌باشد. ما در مورد نوآوری‌های اخیر، همچون هوش مصنوعی و یادگیری ماشین، محاسبات ابر و محاسبات لبه و پتانسیل‌های محاسبات کوانتومی، که در حال پیشبرد قابلیت‌ها و دسترسی‌های داده‌های بزرگ می‌باشند نیز بحث خواهیم کرد. این مقاله همچنین چشم‌اندازهای آینده را نیز با تأکید بر اهمیت داده‌های، انطباق‌های نظارتی و افزایش تقاضا برای مهارت‌های تخصصی، در نظر می‌گیرد. با استفاده از داده‌های بزرگ، سازمان‌ها می‌توانند پیشرفت‌های معنادار و گسترده‌ای را در صنایع مختلف هدایت کنند و راه را برای آینده‌ای داده محور هموار کنند.

واژگان کلیدی: داده‌های بزرگ، پایگاه داده، سیستم‌های اطلاعاتی

۱- مقدمه

در عصر تحولات دیجیتال، داده‌های بزرگ به عنوان یکی از دارایی‌های حیاتی صنایع مختلفی چون مراقبت‌های بهداشتی، مالی و اقتصادی، و غیره ظهور کرده‌اند. داده‌های بزرگ که با حجم وسیعی از داده‌های تولید شده در واحد زمان تعریف می‌شوند، مجموعه‌های اطلاعاتی پیچیده‌ای را در بر می‌گیرند که ابزارهای مدیریت داده‌ی سنتی نمی‌توانند بطور کارآمد آنها را پردازش کنند. این تغییرات اساسی باعث بروز پیشرفت‌هایی سریع در زمینه‌ی تجزیه و تحلیل داده‌ها می‌شود و سازمان‌ها را قادر می‌سازد تا بینش‌های ارزشمندی را از آنها استخراج کنند که کمک فراوانی به تصمیم‌گیری‌های استراتژیک می‌کند و کارایی عملیاتی را نیز افزایش داده و نوآوری را تقویت می‌کند. با اینحال، در کنار این فرصت‌ها، داده‌های بزرگ چالش‌های خاص و مهمی را نیز ارائه می‌کند - از نگرانی‌های مربوط به حریم خصوصی داده‌ها تا نیازهای فنی پردازش و تجزیه و تحلیل داده‌ها در مقیاس بزرگ. این مقاله به بررسی چشم‌انداز چند وجهی تجزیه و تحلیل داده‌های بزرگ و بررسی تأثیرات دگرگون کننده، موانع و پیشرفت‌های تکنولوژیکی را که نقش آنها را در صنایع مختلف تعریف می‌کند، می‌پردازد. با نگاهی دقیق به بخش‌های مختلف، این تحقیق نشان می‌دهد که چگونه سازمان‌ها می‌توانند از داده‌های بزرگ برای هدایت تصمیمات آگاهانه و پرورش مزیت‌های رقابتی در دنیایی که به طور فزاینده‌ای مبتنی بر داده می‌باشد، استفاده کنند.

۲- درک تجزیه و تحلیل داده‌های بزرگ

درک تجزیه و تحلیل داده‌های بزرگ ضروری می‌باشد، زیرا سازمان‌ها را قادر می‌سازد تا حجم عظیمی از داده‌های تولید شده در دنیای دیجیتال امروزی را هدایت و از آنها استفاده کنند. با شناخت ویژگی‌های داده بزرگ (مانند حجم، سرعت، تنوع، صحت و ارزش) کسب و کارها می‌توانند به طور مؤثری اطلاعات را پردازش و تجزیه و تحلیل کنند و از آنها برای تولید بینش‌های عملی استفاده کنند. این درک تصمیم‌گیری‌های استراتژیک را امکان‌پذیر می‌سازد، کارایی عملیاتی را افزایش می‌دهد و راه‌هایی را نیز برای نوآوری در بخش‌هایی مانند سیستم‌های مراقبت‌های بهداشتی، مالی و تولید باز می‌کند.

۲-۱- ویژگی‌های داده‌های بزرگ

- حجم^۱: به حجم عظیمی از داده‌های تولید شده‌ی روزانه که توسط منابعی مانند رسانه‌های اجتماعی، دستگاه‌های اینترنت اشیا، حسگرها و تراکنش‌ها تولید می‌شوند اشاره دارد. مدیریت و پردازش این حجم از داده‌ها یک چالش و همچنین یک فرصت کلیدی می‌باشد.
- سرعت^۲: سرعت تولید داده‌ها، جمع‌آوری و تجزیه و تحلیل داده‌ها را برجسته می‌کند. با گسترش داده‌ها از منابعی مانند رسانه‌های اجتماعی و یا حسگرهای بلادرنگ، پردازش سریع آنها برای به دست آوردن بینش به موقع، ضروری می‌باشد.
- تنوع^۳: داده‌های برگ شامل فرمت‌های متنوعی از داده‌ها می‌باشد، از جمله داده‌های ساختار یافته، داده‌های نیمه ساختار یافته (مانند اسناد JSON و XML) و داده‌های بدون ساختار (مانند ایمیل‌ها، ویدیوها). تجزیه و تحلیل چنین داده‌های متنوعی به ابزارها و روش‌های خاص و انعطاف‌پذیری نیاز دارد.
- صحت^۴: به کیفیت و قابلیت اطمینان داده‌ها اشاره دارد. داده‌های متناقض، ناقص و یا مغرضانه می‌توانند منجر به ایجاد بینش‌های گمراه کننده‌ای شوند، بنابراین اطمینان از صحت و ثبات داده‌ها بسیار مهم می‌باشد.
- ارزش^۵: هدف نهایی استخراج و تجزیه و تحلیل داده‌های بزرگ، بینش‌های عملی می‌باشد که ارزش افزوده، امکان تصمیم‌گیری بهتر و مزایای استراتژیک را فراهم می‌کنند.

۲-۲- تکنیک‌های تجزیه و تحلیل داده‌های بزرگ

- تجزیه و تحلیل توصیفی: این تکنیک یک نمای کلی تاریخی ارائه می‌دهد و داده‌های خام را برای درک آنچه که اتفاق افتاده است آماده می‌کند.
 - تجزیه و تحلیل پیشگویانه: این تکنیک از داده‌های تاریخی برای پیش‌بینی آینده از طریق الگوریتم‌های یادگیری ماشین، روندهای پیش‌بینی، رفتار مشتریان و یا خرابی تجهیزات استفاده می‌کند.
 - تجزیه و تحلیل تجویزی: این تکنیک با تجزیه و تحلیل نتایج احتمالی و پیشنهاد بهترین مسیرهای اقدام، توصیه‌هایی برای تصمیم‌گیری ارائه می‌دهد. از این تکنیک اغلب در بهینه‌سازی، مدیریت ریسک و توصیه‌های شخصی استفاده می‌شود.
- برخی از ابزارها و چارچوب‌های رایج مورد استفاده در این تکنیک‌ها عبارتند از:
- Hadoop: یک چارچوب مقیاس‌پذیر و منبع باز برای ذخیره‌سازی و پردازش داده‌های توزیع شده می‌باشد که برای مدیریت حجم زیادی از داده‌های ساختاریافته و بدون ساختار ایده‌آل می‌باشد.
 - Apache Spark: به دلیل سرعت بالای خود شناخته شده می‌باشد و همچنین قابلیت‌هایی چون پردازش داده‌های بلادرنگ را نیز ارائه می‌دهد که آن را برای برنامه‌های جریان داده ارزشمند می‌سازد.
 - پایگاه‌داده‌های NoSQL: پایگاه‌داده‌هایی چون MongoDB, Cassandra و Couchbase از انواع مختلفی از داده‌ها پشتیبانی می‌کنند که برای مدیریت داده‌های بدون ساختار طراحی شده‌اند. این ویژگی آنها را برای برنامه‌های داده‌های بزرگ حیاتی می‌کند.

۲-۳- داده‌های بزرگ در اکوسیستم‌های صنعتی

- سازگاری با اهداف صنعتی: تجزیه و تحلیل داده‌های بزرگ از اهداف خاص صنعت مانند بهینه‌سازی عملیات تولید، افزایش رضایت مشتریان و همچنین افزایش نتایج بهبودی بیماران در سیستم‌های مراقبت‌های بهداشتی پشتیبانی می‌کند.
- تأثیر بر کارایی کسب و کار: با استفاده از بینش‌های بلادرنگ، سازمان‌ها می‌توانند فرآیندها را ساده‌تر، هزینه‌ها را کاهش داده و تصمیمات پیشگیرانه بگیرند. به عنوان مثال، تعمیر و نگهداری پیش‌بینی شده در فرآیند تولید می‌تواند از خرابی تجهیزات پرهزینه جلوگیری کند.
- تصمیم‌گیری مبتنی بر داده: داده‌های بزرگ پایه و اساس استراتژی‌های مبتنی بر شواهد می‌باشند و به کسب و کارها این امکان را می‌دهند تا روندها را بهتر درک کنند، تغییرات بازار را پیش‌بینی کنند، و محصولات یا خدمات را متناسب با نیازهای در حال تکامل مشتری تنظیم کنند.

۳- برخی کاربردهای کلیدی داده‌های بزرگ در صنایع مختلف

۳-۱- مراقبت‌های بهداشتی

- بهینه‌سازی مراقبت از بیماران: تجزیه و تحلیل داده‌های بزرگ با امکان درمان‌های پیش‌بینی شده و شخصی، سیستم‌های مراقبت‌های بهداشتی را متحول کرده است. همانطور که توسط (Batko et al, 2022) اشاره شده است، ارائه دهندگان خدمات مراقبت‌های بهداشتی می‌توانند با تجزیه و تحلیل سوابق بیمار، اطلاعات ژنتیکی و داده‌های بهداشتی بلادرنگ، تشخیص‌های دقیق‌تر را انجام داده و خطرات احتمالی بیماری را نیز پیش‌بینی کنند. همچنین (Dash et al. 2019) بر این نکته تأکید می‌کند که چگونه ابزارهای داده‌های بزرگ به شناسایی الگوها و پیش‌بینی شیوع بیماری‌ها کمک می‌کنند و امکان مداخله زود هنگام و تخصیص مؤثر منابع را فراهم می‌کنند. این می‌تواند به ویژه در زمان همه‌گیری‌ها، که در آنها داده‌های بلادرنگ از پاسخ‌های هدفمند پشتیبانی می‌کنند، بسیار مهم باشد.

- مدیریت کارایی و هزینه: تجزیه و تحلیل مجموعه داده‌های بزرگ به بهبود کارایی عملیاتی در سیستم‌های مراقبت‌های بهداشتی کمک شایانی می‌کند. برای مثال، با بهینه‌سازی زمان‌بندی، تخصیص منابع و زنجیره‌های تامین، سازمان‌های مراقبت‌های بهداشتی می‌توانند هزینه‌های خود را بدون به خطر انداختن کیفیت مراقبت‌های بهداشتی کاهش دهند.

۲-۲- تشخیص تقلب

در امور مالی، تجزیه و تحلیل داده‌های بزرگ در شناسایی و پیشگیری از کلاهبرداری‌ها مؤثر می‌باشد. با تجزیه و تحلیل داده‌های مربوط به تراکنش‌های مالی، الگوریتم‌های یادگیری ماشین می‌توان الگوهای مرتبط با فعالیت‌های متقلبانه را شناسایی کرد و به موسسات در زمان‌های واقعی هشدار داد. تکنیک‌هایی که توسط (Gulia et al. 2016) مورد بحث قرار گرفته‌اند نیز تشخیص چنین ناهنجاری‌هایی را که کلید کاهش ریسک‌های مالی می‌باشند را بهبود می‌بخشد.

- مدیریت ریسک: شرکت‌های مدیریت ریسک از داده‌های بزرگ برای ارزیابی ریسک‌های اعتباری، ریسک‌های بازار و ریسک‌های عملیاتی استفاده می‌کنند. تجزیه و تحلیل‌های پیشرفته بینش‌هایی را در مورد خطرات بالقوه ارائه می‌دهد که مؤسسات را قادر می‌سازد تا به طور فعالانه‌ای این ریسک‌ها را مدیریت کنند و در عین حال سبد سرمایه‌گذاری و استراتژی‌های قیمت‌گذاری را نیز بهینه کنند.
- مشتری‌مداری و شخصی‌سازی: مؤسسات مالی از داده‌های بزرگ برای تجزیه و تحلیل رفتار مشتریان استفاده می‌کنند و خدمات شخصی‌سازی شده‌تری را برای آنها ممکن می‌سازند. این شامل تطبیق محصولات با ترجیحات فردی، حفظ مشتری و ایجاد استراتژی‌های بازاریابی کارآمدتر می‌باشد (Kune et al. 2015).

۳-۳- تجزیه و تحلیل رفتار مصرف‌کنندگان در خرده‌فروشی و تجارت الکترونیک

در خرده‌فروشی، تجزیه و تحلیل داده‌های بزرگ منجر به جمع‌آوری و تجزیه و تحلیل ترجیحات مشتری می‌شود و این شرکت‌ها را قادر می‌سازد تا استراتژی‌های بازاریابی و توصیه‌های محصولات خود را اصلاح کنند. با بررسی الگوهای خرید و جمعیت‌شناسی مشتری، شرکت‌ها تبلیغات هدفمندتری را ایجاد می‌کنند و تعاملات مشتریان را بهبود می‌بخشند.

- بهینه‌سازی موجودی: خرده‌فروشان را تجزیه و تحلیل می‌کند تا موجودی را در زمان واقعی نظارت کنند و از موجود بودن اقلام محبوب اطمینان حاصل کنند و در عین حال مازاد و ضایعات را به حداقل می‌رساند. این امر به ویژه برای شرکت‌هایی با زنجیره‌های تامین گسترده مفید است، جایی که بینش‌های مبتنی بر داده فرآیند مدیریت موجودی را ساده می‌کنند.
- توصیه‌های شخصی: در تجارت الکترونیک، داده‌های بزرگ موتورهای توصیه‌گر را فعال می‌کند و براساس رفتار گذشته‌ی مشتریان محصولات متناسبی را پیشنهاد می‌دهند. بر اساس گفته‌های (Galetsi et al. 2019)، این الگوریتم‌ها به طور قابل توجهی تجربه‌ی خرید مشتریان را بهبود می‌دهند و میزان فروش را با توجه به اولویت‌های مشتریان افزایش می‌دهند.

۳-۴- تولید

تحلیل‌های پیش‌بینی کننده به طور گسترده‌ای در فرآیندهای تولید برای پیش‌بینی خرابی ماشین آلات قبل از وقوع آن، استفاده می‌شوند. با تجزیه و تحلیل داده‌های مربوط به حسگرهای تجهیزات، سازندگان می‌توانند تعمیر و نگهداری را به‌طور مؤثرتری برنامه‌ریزی کنند و زمان خرابی و هزینه‌های تعمیر و نگهداری تجهیزات را کاهش دهند (Wang et al. 2019).

- بهینه‌سازی زنجیره‌ی تامین: تولیدکنندگان از تجزیه و تحلیل زنجیره‌ی تامین خود برای نظارت بر در دسترس بودن مواد خام، پیش‌بینی نوسانات عرضه و تقاضا و همچنین شناسایی اختلالات احتمالی استفاده می‌کنند. داده‌های بزرگ به ایجاد یک زنجیره‌ی تامین انعطاف‌پذیرتر کمک می‌کند، هزینه‌ها را کاهش می‌دهد و در دسترس بودن محصولات را نیز تضمین می‌کند.

- کنترل کیفیت: تجزیه و تحلیل داده‌های به دست آمده از فرآیندهای تولید، تولیدکنندگان را قادر می‌سازد تا مسائل کیفی را سریع‌تر تشخیص دهند. با شناسایی ناهنجاری‌ها و نظارت بر معیارهای کیفیت، تولیدکنندگان می‌توانند استانداردهای بالایی را حفظ کرده و فراخوانی محصولات را کاهش دهند که به نوبه‌ی خود باعث صرفه‌جویی در هزینه‌ها و محافظت از شهرت برند نیز می‌شود.

۵-۳- انرژی و تاسیسات

- شبکه‌های هوشمند و مدیریت انرژی: داده‌های بزرگ با اجرای شبکه‌های هوشمند، که میزان مصرف انرژی را در زمان واقعی نظارت می‌کنند و افزایش تقاضا را پیش‌بینی می‌کنند، بخش انرژی را متحول می‌کند. بر اساس گفته‌های (Zhang et al. 2023)، این داده‌ها به شرکت‌ها کمک می‌کنند تا بارها را به‌طور مؤثرتری مدیریت کنند و از قطع برق جلوگیری کنند.
- پیش‌بینی تعمیر و نگهداری زیرساخت‌ها: تجزیه و تحلیل داده‌ها به پیش‌بینی نیازهای تعمیر و نگهداری برای زیرساخت‌های انرژی کمک شایانی می‌کند و همچنین اختلالات خدمات را به حداقل می‌رساند. این امر به ویژه در تاسیسات حیاتی، که در آنها قطعی‌های برنامه‌ریزی نشده می‌تواند تأثیرات گسترده‌ای داشته باشد، بسیار مهم می‌باشد.
- پایداری و تأثیرات زیست محیطی: شرکت‌های انرژی نیز از داده‌های بزرگ برای نظارت و کاهش ردپای زیست محیطی خود استفاده می‌کنند. با ردیابی الگوهای مصرف انرژی و داده‌های منتشر شده، آنها می‌توانند زمینه‌های قابل بهبود را شناسایی کنند و به آنها کمک کنند تا استانداردهای نظارتی را رعایت کنند و با اهداف پایداری هماهنگ شوند.

۴- چالش‌های موجود در تجزیه و تحلیل داده‌های بزرگ

تجزیه و تحلیل داده‌های بزرگ، درحالی‌که تحول آفرین می‌باشد، با چالش‌های مهمی نیز مواجه می‌باشد که اجرای آن را در برخی از صنایع پیچیده می‌کند. حفظ حریم خصوصی و امنیت داده‌ها، به ویژه در مورد اطلاعات حساس در بخش‌هایی همچون سیستم‌های مراقبت‌های بهداشتی و سیستم‌های مالی که در آنها مقرراتی مانند GDPR و HIPAA دستورالعمل‌های سختگیرانه‌ای را اعمال می‌کنند، یکی از نگرانی‌های اساسی می‌باشد. یکپارچه‌سازی و نگهداری داده‌های با کیفیت بالا از منابع متنوع و متناقض از نظر فنی پیچیده می‌باشد، همانطور که زیرساخت مورد نیاز برای پردازش حجم وسیعی از داده‌ها در زمان واقعی پیچیده می‌باشد. علاوه بر این، کمبود متخصصین داده‌ی ماهر چنین چالش‌هایی را تشدید می‌کند، زیرا سازمان‌ها در تلاش برای یافتن پرسنلی می‌باشند که هم در تجزیه و تحلیل و هم در چنین حوزه‌ی خاصی مهارت داشته باشند. نگرانی‌های اخلاقی، مانند سوء استفاده از داده‌ها نیز پذیرش داده‌های بزرگ را پیچیده‌تر می‌کند و سازمان‌ها را ملزم می‌کند تا از شیوه‌های احراز هویت و اعمال سطوح دسترسی به داده‌ها اطمینان حاصل کنند. پیروی از استانداردهای نظارتی در مناطق مختلف نیز تنها بر این پیچیدگی‌ها می‌افزاید، و استقرار موفقیت‌آمیز تجزیه و تحلیل داده‌های بزرگ را به یک تلاش چالش برانگیز اما حیاتی برای کسب و کارهایی تبدیل می‌کند که قصد دارند از بینش‌های مبتنی بر داده استفاده کنند.

۱-۴- حریم خصوصی و امنیت داده‌ها

- نگرانی‌های مربوط به حفاظت از داده‌ها: با مجموعه‌ی گسترده‌ای از اطلاعات حساس، به ویژه در بخش‌هایی مانند مراقبت‌های بهداشتی و مالی، اطمینان از حفظ حریم خصوصی و امنیت داده‌ها و پیروی از مقرراتی مانند GDPR و HIPA، چالش برانگیز می‌باشد. به ویژه صنعت مراقبت‌های بهداشتی با موانع قابل توجهی در بحث حفاظت از داده‌های بیماران مواجه می‌باشد، زیرا بروز هرگونه نقض قوانین می‌تواند منجر به نقض حریم خصوصی و از دست رفتن اعتماد عمومی شود (Batko et al. 2022).

- خطرات امنیت سایبری: همانطور که (Wang et al. 2019) اشاره کردند، سیستم‌های داده‌های بزرگ به دلیل ارزش بالای داده‌هایی که دارند، مدام هدف حملات سایبری می‌باشند. حفاظت از آنها در برابر نقض داده‌ها به اقدامات امنیتی پیشرفته‌ای همچون رمزگذاری، احراز هویت و تشخیص تهدیدها در زمان واقعی نیاز دارد تا از دسترسی‌های غیرمجاز جلوگیری شود.

۴-۲- یکپارچه‌سازی و کیفیت داده‌ها

- پیچیدگی یکپارچه‌سازی منابع مختلف داده: تجزیه و تحلیل داده‌های بزرگ اغلب شامل جمع‌آوری داده‌ها از منابع مختلفی همچون پایگاه‌داده‌ها، حسگرهای اینترنت اشیا، رسانه‌های اجتماعی و تراکنش‌های مشتریان می‌باشد. (Dash et al. 2019) تاکید می‌کند که ادغام این منابع متفاوت در یک سیستم یکپارچه از نظر فنی چالش برانگیز می‌باشد، زیرا داده‌ها اغلب ناهمگن، ناسازگار و یا ناقص می‌باشند.
- کیفیت داده‌ها: داده‌های با بی کیفیت یا با کیفیت پایین می‌توانند منجر به نتایج تحلیلی گمراه کننده شوند. (Kune et al. 2015) بر اهمیت فرآیندهای پاکسازی داده‌ها برای اطمینان از قابلیت اطمینان داده‌ها که خود لایه‌ای از پیچیدگی و نیاز به منابع را به مراحل آماده‌سازی داده‌ها می‌افزاید، تاکید می‌کنند.

۴-۳- پیچیدگی‌های فنی و محاسباتی

- زیرساخت: زیرساخت سخت‌افزاری و نرم‌افزاری مورد نیاز برای تجزیه و تحلیل داده‌های بزرگ اغلب پرهزینه می‌باشد. همانطور که توسط (Gulia et al. 2016) مشخص شده است، ذخیره‌سازی داده‌ها در مقیاس بزرگ، پردازش داده‌های بزرگ با کارایی بالا و همچنین مقیاس‌پذیری برای تطبیق با حجم رو به رشد داده‌ها، ضروری اما گران می‌باشند.
- مقیاس‌پذیری و پردازش بلادرنگ: پردازش حجم وسیعی از داده‌ها در زمان واقعی چالش‌هایی را برای بسیاری از سازمان‌ها ایجاد می‌کند، به ویژه در زمانیکه حجم و سرعت تولید داده‌ها به سرعت افزایش می‌یابد. بر طبق گفته‌ی (Galetsi et al. 2019)، حفظ یک زیرساخت مقیاس‌پذیر که بتواند رشد داده‌ها را در حالیکه بینش‌های به موقعی را ارائه می‌دهد مدیریت کند، به ویژه برای شرکت‌هایی با منابع تکنولوژیک محدود پیچیده می‌باشد.

۴-۴- شکاف‌های مهارتی

- کمبود متخصصین ماهر: تقاضاهای زیادی برای دانشمندان، مهندسان و تحلیلگران داده‌ای وجود دارد که هم مهارت‌های فنی و هم دانش خاص یک صنعت را دارا باشند، وجود دارد. کمبود متخصصان ماهر یک گلوگاه قابل توجه در حوزه‌ی داده‌های بزرگ می‌باشد که (Wang et al. 2018) با اشاره به اینکه بسیاری از سازمان‌ها برای یافتن پرسنل ماهر در ابزارهای تجزیه و تحلیل داده‌ها، یادگیری ماشین و هوش مصنوعی تلاش می‌کنند، آن را به عنوان یکی از چالش‌های موجود در این حوزه مطرح می‌کنند.
- توسعه‌ی مستمر مهارت‌ها: با پیشرفت‌های سریع در زمینه‌ی فناوری‌های تحلیلی، سازمان‌ها به متخصصانی نیاز دارند که بتوانند به طور مداوم مهارت‌های خود را به‌روز کنند. این نیاز به آموزش مداوم، هزینه‌ها و پیچیدگی‌های بیشتری را به خصوص برای شرکت‌های کوچکتر که ممکن است فاقد منابع آموزشی باشند، اضافه می‌کند.

۵- نوآوری‌ها و روندهای نوظهور در حوزه‌ی داده‌های بزرگ

نوآوری‌ها و روندهای نوظهور در حوزه‌ی داده‌های بزرگ، نحوه‌ی تجزیه و تحلیل سازمان‌ها و استفاده از داده‌ها در بخش‌های مختلف را تغییر می‌دهند. هوش مصنوعی^۱ (AI) و یادگیری ماشین^۱ (ML) تجزیه و تحلیل‌های پیش‌بینی کننده را تقویت می‌کنند و امکان

تشخیص دقیق‌تری را به عنوان مثال در سیستم‌های مراقبت‌های بهداشتی و تصمیم‌گیری خودکار در امور مالی را فراهم می‌کند. راه‌حل‌های مبتنی بر ابر نیز میزان دسترسی‌ها را افزایش داده‌اند و به سازمان‌ها این امکان را می‌دهند که قابلیت‌های پردازش داده‌های خود را بدون انجام سرمایه‌گذاری بر روی زیرساخت‌های سنگین، در حالیکه مدل‌های ترکیبی و چند ابری انعطاف‌پذیری را اضافه می‌کنند، مقیاس‌بندی کنند. محاسبات لبه، تجزیه و تحلیل‌های بلادرنگ را به خط مقدم این حوزه می‌آورد، به‌ویژه در برنامه‌های اینترنت اشیا^۲ (IoT) که در آنها بینش‌های فوری حیاتی می‌باشند، مانند فرآیندهای تولید و یا وسایل نقلیه‌ی خودران. بلاک چین نیز به عنوان ابزاری برای مدیریت ایمن و غیر متمرکز داده‌ها و حصول اطمینان از یکپارچگی داده‌ها در زمینه‌هایی همچون سیستم‌های مالی و مراقبت‌های بهداشتی، مورد توجه قرار گرفته است. در همین حال، پیشرفت‌ها در محاسبات کوانتومی نوید بخش پردازش سریع‌تر داده‌ها می‌باشد، که می‌تواند تحولی شگرف در تحلیل‌های پیچیده‌ی داده‌های بزرگ در زمینه‌هایی همچون ژنتیک و آب و هوا ایجاد کند. این نوآوری‌ها در مجموع بر چشم‌انداز در حال تکامل داده‌های بزرگ تأکید می‌کنند و ابزارها و رویکردهای جدیدی را برای مهار ارزش داده‌ها به روش‌های گوناگون ارائه می‌دهند.

۱-۵- ادغام هوش مصنوعی و یادگیری ماشین

- قابلیت پیش‌بینی پیشرفته: ادغام هوش مصنوعی و یادگیری ماشین با تجزیه و تحلیل داده‌های بزرگ، صنایع را با ارائه‌ی ابزارهای پیش‌بینی قدرتمندی متحول می‌کند. (Wang et al. 2018) در مورد اینکه چگونه تجزیه و تحلیل‌های مبتنی بر هوش مصنوعی می‌توانند مجموعه داده‌های گسترده‌ای را در حوزه‌ی مراقبت‌های بهداشتی برای پیش‌بینی نتایج بیمار، بهبود تشخیص و شخصی سازی درمان‌ها پردازش کنند، بحث می‌کنند. مدل‌های یادگیری ماشین همچنین سازمان‌ها را قادر می‌سازد تا الگوهای پیچیده را شناسایی کرده و فرآیندهای تصمیم‌گیری را خودکار کنند.
- یادگیری عمیق برای داده‌های بدون ساختار: پیشرفت‌ها در یادگیری عمیق توانایی پردازش داده‌های بدون ساختار همچون تصاویر و متن را بهبود بخشیده است و امکانات جدیدی را برای بخش‌هایی مانند مراقبت‌های بهداشتی و مالی فراهم می‌کند (Dash et al. 2019). به عنوان مثال، الگوریتم‌های تشخیص تصاویر به تشخیص تصویربرداری پزشکی کمک می‌کنند، درحالیکه تکنیک‌های NLP (پردازش زبان طبیعی) برای تجزیه و تحلیل احساسات در امور مالی مورد استفاده قرار می‌گیرند.

۲-۵- راه حل‌های مبتنی بر ابر برای داده‌های بزرگ

- مقیاس‌پذیری و هزینه‌ی کارایی: رایانش ابری به عنوان یک مؤلفه‌ی حیاتی برای داده‌های بزرگ پدیدار شده است و ذخیره‌سازی و قدرت پردازش مقیاس‌پذیر را بدون نیاز به هیچ زیرساخت گسترده‌ای در محل ارائه می‌دهد. (Zhang et al. 2023) نشان می‌دهند که پلتفرم‌های ابری همچون AWS, Google Cloud و Microsoft Azure تجزیه و تحلیل داده‌های بزرگ را به ویژه برای سازمان‌های کوچک قابل دسترس‌تر کرده‌اند. با کاهش هزینه‌های زیرساخت، راه‌حل‌های ابری سازمان‌ها را قادر می‌سازند تا مجموعه داده‌های بزرگ را به طور کارآمدتری تجزیه و تحلیل و ذخیره کنند.
- مدل‌های ابری چندگانه و ترکیبی: بسیاری از سازمان‌ها مدل‌های ابری چندگانه و ترکیبی را برای افزایش انعطاف‌پذیری و امنیت داده‌های خود اتخاذ می‌کنند. این رویکرد به شرکت‌ها اجازه می‌دهد تا داده‌ها را در میان ارائه‌دهندگان مختلف ابری توزیع کنند، هزینه‌ها را متعادل کنند و در صورت قطع شدن سرویس، افزونگی را ارائه دهند.

۳-۵- محاسبات لبه و تجزیه و تحلیل در زمان واقعی

- کاهش تأخیر و بینش سریع‌تر: محاسبات لبه، که شامل پردازش داده‌ها در نزدیکی منبع خود می‌باشد، انقلابی در تجزیه و تحلیل داده‌های بلادرنگ ایجاد کرده است. با توجه تحقیقات انجام شده توسط (Gulia et al. 2016)، این رویکرد تأخیر را

به حداقل می‌رساند و زمان پاسخگویی سریع‌تری را ممکن می‌سازد. محاسبات لبه در بخش‌هایی مانند تولید و انرژی، که در آنها بینش‌های فوری برای تصمیم‌گیری خودکار، تعمیر و نگهداری پیش‌بینی‌کننده و کارایی عملیاتی ضروری می‌باشد، بسیار مهم می‌باشد.

- برنامه‌های پیشرفته‌ی اینترنت اشیا: با تجزیه و تحلیل داده‌ها در لبه، برنامه‌های کاربردی اینترنت اشیا در بخش‌های مربوط به مراقبت‌های بهداشتی، خودرو و تدارکات از پردازش داده‌ها در زمان واقعی، بهبود نظارت بر بیمار، رانندگی مستقل، و مدیریت زنجیره‌ی تامین سود می‌برند.

۴-۵- استفاده از بلاک چین برای امنیت داده‌ها

- مدیریت غیرمتمرکز داده‌ها: فناوری بلاک چین به طور فزاینده‌ای به عنوان وسیله‌ای برای ایمن سازی و مدیریت داده‌های بزرگ، به ویژه در صنایعی که یکپارچگی و شفافیت داده‌ها ضروری می‌باشد، مورد بررسی قرار می‌گیرد. (Kune et al. 2015) در مورد اینکه چگونه ماهیت غیرمتمرکز بلاک چین، مدیریت داده‌های غیر قابل دستکاری را تضمین می‌کند، بحث می‌کنند. این به ویژه در امور مالی و مراقبت‌های بهداشتی مفید می‌باشد. با بلاک چین، هر تراکنش داده در یک دفتر کل توزیع شده ثبت می‌شود و خطرات همچون نقض داده‌ها و تغییرات غیرمجاز را کاهش می‌دهد.
- قراردادهای هوشمند برای اشتراک‌گذاری ایمن داده‌ها: عملکرد قراردادهای هوشمند بلاک چین، اشتراک‌گذاری ایمن و شفاف داده‌ها را در بین سازمان‌ها امکان‌پذیر می‌کند و همچنین راهی برای اجرای خودکار خط‌مشی‌های حفظ حریم خصوصی داده‌ها را فراهم می‌کند. این برای حوزه‌ی مراقبت‌های بهداشتی و سایر بخش‌هایی که به اشتراک‌گذاری داده‌های حساس نیاز دارند، ارزشمند می‌باشد.

۵-۵- پتانسیل محاسبات کوانتومی

- قدرت پردازش داده‌ها: محاسبات کوانتومی، درحالی‌که هنوز در مراحل اولیه خود می‌باشد، پتانسیلی خوب برای بهبود تصاعدی قابلیت‌های پردازش داده‌های بزرگ دارد. (Zhang et al. 2023) نشان می‌دهند که چگونه الگوریتم‌های کوانتومی می‌توانند محاسبات پیچیده را به طور قابل توجهی سرعت بخشند و تجزیه و تحلیل مجموعه داده‌های عظیم را سریعتر از محاسبات سنتی ممکن سازند. اگرچه کاربردهای عملی محاسبات کوانتومی در حوزه‌ی داده‌های بزرگ هنوز در حال بررسی می‌باشد، اما روزی می‌تواند پیشرفت‌های عظیمی را در صنایعی با حجم بالای تقاضای برای داده‌ها، مانند داروسازی، مدل‌سازی آب و هوا، و ژنتیک امکان‌پذیر کند.
- پیامدهای آینده برای هوش مصنوعی و یادگیری ماشین: محاسبات کوانتومی ممکن است هوش مصنوعی و الگوریتم‌های یادگیری ماشین را با تسریع زمان آموزش برای مدل‌های پیچیده بهبود بخشد. این پتانسیل می‌تواند شبیه‌سازی‌ها و تحلیل‌های پیشرفته‌ای را که در حال حاضر توسط سرعت‌های محاسباتی سنتی محدود شده‌اند، امکان‌پذیر کند.

۶- پیشنهادات آتی

آینده‌ی داده‌های بزرگ فرصت‌های متحول‌کننده‌ای را در تمامی صنایع ایجاد کرده است. در حوزه‌ی مراقبت‌های بهداشتی تجزیه و تحلیل‌های و پیش‌بینی‌های دقیق مبتنی بر داده‌ها روند تشخیص بیماران را تسریع بخشیده و هزینه‌ها را کاهش داده‌است. در حوزه‌های مالی نیز، داده‌های بزرگ از مدیریت ریسک بلادرنگ و تشخیص تقلب پشتیبانی می‌کند. توسعه شهرهای هوشمند، با استفاده از داده‌های اینترنت اشیا، نوید مدیریت منابع کارآمدتر و بهبود خدمات عمومی را می‌دهد، درحالی‌که محاسبات کوانتومی، که هنوز در حال ظهور می‌باشد، می‌تواند به طور تصاعدی تجزیه و تحلیل داده‌های پیچیده را به ویژه در حوزه‌های کشف دارو و مدل‌سازی آب و هوا، سرعت بخشد. علاوه بر این، تقاضا برای متخصصان داده و هوش مصنوعی رو به رشد می‌باشد و این، نیاز به مهارت‌های بین رشته‌ای را برجسته

می‌کند که تخصص داده را با دانش خاص یک صنعت ترکیب می‌کند. این پیشرفت‌ها و برنامه‌ها با هم بر نقش رو به گسترش داده‌های بزرگ در ایجاد نوآوری، کارایی و ارزش در صنایع مختلف، تأکید می‌کنند.

۱-۶- گسترش نقش داده‌های بزرگ در صنایع مختلف

- پیشرفت‌های حوزه‌ی مراقبت‌های بهداشتی: صنعت مراقبت‌های بهداشتی قرار است شاهد رشد مداوم در برنامه‌های مربوط به داده‌های بزرگ باشد. (Batko et al. 2022) پیش‌بینی می‌کنند که ادغام بیشتر هوش مصنوعی و داده‌های بزرگ منجر به پیشرفت‌هایی در حوزه‌ی پزشکی می‌شود، تا جاییکه درمان‌ها براساس مشخصات ژنتیکی فرد و داده‌های سلامت بلندمدت وی تنظیم می‌شوند. تجزیه و تحلیل پیش‌بینی کننده‌ی زود هنگام، فرآیند تشخیص و مراقبت‌های پیشگیرانه را تسریع می‌بخشد، که در نهایت می‌تواند هزینه‌های مربوط به مراقبت‌های بهداشتی را کاهش دهد.
- امور مالی و مدیریت ریسک بلندمدت: خدمات مالی به احتمال زیاد به طور فزاینده‌ای برای مدیریت ریسک بلندمدت و کشف تقلب به داده‌های بزرگ متکی خواهند بود. همانطور که توسط (Gulia et al. 2016) اشاره شده است، پیشرفت‌ها در تجزیه و تحلیل داده‌های بزرگ به اصلاح الگوریتم‌های مورد استفاده برای تشخیص ناهنجاری‌ها، افزایش توانایی موسسات مالی برای واکنش سریع به تغییرات بازار و تهدیدات بالقوه ادامه خواهد داد.

۲-۶- نوآوری‌های مبتنی بر داده و اینترنت اشیاء در شهرهای هوشمند

- توسعه‌ی شهرهای هوشمند: داده‌های بزرگ برای آینده‌ی شهرهای هوشمند بنیادی می‌باشند، تا جاییکه داده‌های دستگاه‌های متصل به اینترنت مدیریت کارآمد منابع را امکان پذیر می‌کنند، ازدحام ترافیک را کاهش می‌دهند و خدمات عمومی را بهینه می‌کنند. (Wang et al. 2018) نشان می‌دهند که داده‌های بزرگ به طور فزاینده‌ای در زمینه‌ی بهبود زیرساخت‌ها، از داده‌ها برای مدیریت موثر توزیع انرژی، زباله و برنامه‌ریزی شهری پشتیبانی می‌کنند.
- گسترش اینترنت اشیاء و تصمیم‌گیری در زمان واقعی: با رشد و گسترش اینترنت اشیاء، به ویژه در بخش‌هایی مانند لجستیک و کشاورزی، تجزیه و تحلیل داده‌های بزرگ از تصمیم‌گیری در زمان واقعی از طریق داده‌های جمع‌آوری شده از طریق شبکه‌ی رو به رشد دستگاه‌های هوشمند پشتیبانی می‌کند. تجزیه و تحلیل بلندمدت در مقیاس بالا، مدیریت پویا موجودی‌ها، نگهداری تجهیزات پیش‌بینی کننده و کشاورزی دقیق را امکان‌پذیر می‌سازد. این می‌تواند باعث افزایش بهره‌وری و کاهش هزینه‌ها شود.

۳-۶- مدیریت داده‌های پیشرفته و هوش مصنوعی

- حفظ حریم خصوصی و انطباق داده‌ها: با افزایش حجم داده‌ها، نیاز به چارچوب‌های قوی حاکمیت داده افزایش می‌یابد. (Dash et al. 2019) فشار مداوم برای قوانین سختگیرانه‌تر حفظ حریم خصوصی و امنیت داده‌ها را برجسته می‌کنند. این سازمان‌ها را به اتخاذ تدابیر انطباق پیشرفته و شیوه‌های شفاف سازی داده‌ها سوق می‌دهد. حاکمیت قوی داده‌ها در شرکت‌ها برای حفظ اعتماد کاربران و رعایت مقررات در بخش‌هایی مانند مراقبت‌های بهداشتی و مالی ضروری می‌باشد.
- هوش مصنوعی در تجزیه و تحلیل داده‌ها: سازمان‌ها باید به پیامدهای اخلاقی داده‌های بزرگ و هوش مصنوعی نیز بپردازند. (Zhang et al. 2023) پیشنهاد می‌کنند که پیشرفت‌های آینده بر توسعه‌ی الگوریتم‌هایی متمرکز خواهد بود که نه تنها مؤثر، بلکه اخلاقی و بی‌طرفانه نیز می‌باشند. این تلاش‌ها برای حفظ اعتماد عمومی واجب خواهد بود، به ویژه زمانی که سیستم‌های هوش مصنوعی در فرآیندهای تصمیم‌گیری ادغام می‌شوند.

۴-۶- رایانش کوانتومی و داده‌های بزرگ

- تجزیه و تحلیل و شبیه سازی شتاب یافته: محاسبات کوانتومی، اگرچه هنوز در حال ظهور می‌باشد اما نوید تغییر فرآیند تجزیه و تحلیل داده‌های بزرگ با انجام محاسباتی را دارد که در حال حاضر با روش‌های کلاسیک غیرممکن می‌باشد. همانطور که توسط (Simic et al. 2023) بیان شده است، محاسبات کوانتومی می‌تواند پیشرفت‌هایی را در زمینه‌هایی چون کشف دارو، که در آن مجموعه داده‌های عظیمی برای شبیه‌سازی فعل و انفعالات مولکولی تجزیه و تحلیل می‌شوند، امکان‌پذیر کند. این همچنین می‌تواند برای صنایع مالی و لجستیک نیز مفید باشد، تا جایگاه تجزیه و تحلیل داده‌های پیچیده و سریع باعث ایجاد مزایای رقابتی می‌شوند.
- پیامدهای یادگیری ماشین: محاسبات کوانتومی ممکن است توسعه‌ی مدل‌های یادگیری ماشین را تسریع کند و همچنین آموزش سریع‌تر الگوریتم‌های یادگیری عمیق را در مجموعه داده‌های بزرگ ممکن می‌سازد. این تغییر بالقوه می‌تواند راه‌های جدیدی را برای تجزیه و تحلیل داده‌های با ابعاد بالا در زمینه‌هایی مانند ژنتیک و اقلیم‌شناسی باز کند، تا جایگاه الگوهای پیچیده‌ی داده برای پیشرفت تحقیقات و نوآوری حیاتی می‌باشند.

۵-۶- تکامل بازار کار و توسعه‌ی مهارت‌ها

- تقاضا برای متخصصان داده و هوش مصنوعی: رشد داده‌های بزرگ باعث افزایش تقاضا برای متخصصان ماهر در علم داده، یادگیری ماشین و همچنین هوش مصنوعی می‌شود. (Galetsi et al. 2019) تأکید می‌کنند که با اولویت دادن به تصمیم‌گیری‌های مبتنی بر داده، سازمان‌ها به طور فزاینده‌ای به دنبال افرادی می‌گردند که بتوانند تحلیل‌های پیچیده‌ای را در زمینه‌های تجاری تفسیر و اعمال کنند. این تقاضاها احتمالاً برنامه‌های آموزشی را شکل می‌دهد که منجر به توسعه دوره‌ها و گواهینامه‌های تخصصی علوم داده می‌شود.
- فرصت‌هایی برای تخصص‌های بین رشته‌ای: با ادغام برنامه‌های داده‌های بزرگ در صنایع مختلف، نیازهای روز افزونی به تخصص‌های بین رشته‌ای بوجود خواهد آمد که مهارت‌های داده‌های فنی را با دانشی خاص در یک حوزه ترکیب می‌کند. به عنوان مثال، متخصصان مراقبت‌های بهداشتی با آموزش‌های علوم داده، و یا کارشناسان مالی مسلط به هوش مصنوعی ارزشمندتر خواهند بود زیرا صنایع به دنبال ادغام تجزیه و تحلیل در زمینه‌های تخصصی می‌باشند.

۷- نتیجه‌گیری

تجزیه و تحلیل داده‌های بزرگ با ارائه بینشی عمیق، بهینه‌سازی عملیات‌ها و تقویت نوآوری‌ها، صنایع را متحول می‌کند. همانطور که سازمان‌ها به طور فزاینده‌ای از استراتژی‌های داده محور استفاده می‌کنند، پتانسیل قابل توجهی نیز برای بهبود فرآیندهای تصمیم‌گیری، کارایی و مزیت رقابتی به وجود خواهد آمد. با اینحال، سفر به سمت استفاده‌ی موثر از داده‌های بزرگ با چالش‌هایی اساسی، از جمله نگرانی‌های مربوط به حریم خصوصی داده‌ها، نیازهای زیرساختی، و نیاز به متخصصین ماهر همراه می‌باشد. نوآوری‌هایی چون ادغام هوش مصنوعی، محاسبات ابری، پردازش لبه، و فناوری‌های کوانتومی در حال ظهور، راه را برای راه‌حل‌های تحلیلی پیچیده‌تر و مقیاس‌پذیرتر هموار می‌کنند. حرکت رو به جلو، حاکمیت داده‌ها و تخصص‌های بین رشته‌ای از موارد حیاتی خواهند بود، زیرا چشم‌انداز آنها همچنان در حال تکامل می‌باشد. با استفاده‌ی مسئولانه و کارآمد از قدرت داده‌های بزرگ، صنایع این فرصت را دارند که تغییرات تاثیرگذاری را در ارتقای خدمات، محصولات و تجربیات آینده را ایجاد کنند.

منابع

Hemlata and Preeti Gulia, "Big Data Analytics", In: Research Journal of Computer and Information Technology Sciences, Vol. 4(2), pp. 1-4, February 2016.

- Jin Zhang, Dietmar Wolfram, Feicheng Ma, “The impact of big data on research methods in information science”, In: Data and Information Management 7 (2023) 100038, DOI:10.1016/j.dim.2023.100038.
- Kornelia Batko, Andrzej Slezak, “The use of Big Data Analytics in Healthcare”, In: Journal of Big Data, (2022)9:3, DOI:10.1186/s40537-021-00553-4.
- Lidong Wang, Cheryl Ann Alexander, “Big Data Analytics in Healthcare Systems”, In: International Journal of Mathematical, Engineering and Management Sciences, Vol. 4, No. 1, pp. 17-26, 2019.
- Panagiota Galetsi, Korina Katsaliaki, Sameer Kumar, “Big data analytics in health sector: Theoretical framework, techniques and prospects”, In: International Journal of Information Management”, DOI:10.1016/j.ijinfomgt.2019.05.003.
- Raghavendra Kune, Pramod Kumar Konugurthi, Arun Agarwal, Raghavendra Rao Chillarige, and Rajkumar Buyya, “The Anatomy of Big Data Computing”, 2015, DOI:10.1002/spe.2374.
- Sabyasachi Dash, Sushil Kumar Shakyawar, Mohit Sharma, Sandeep Kaushik, “Big Data in Healthcare: Management, Analysis and Future Prospects”, In: Journal of Big Data, (2019)6:54, DOI:10.1186/s40537-019-0217-0.
- Srdan Daniel Simic, Nikola Tankovic, Darko Etinger, “Big data BPMN workflow resource optimization in the cloud”, In: Parallel Computing (2023), DOI:10.1016/j.parco.2023.103025.
- Yichuan Wang, LeeAnn Kung, Terry Anthony Byrd, “Big Data Analytics: Understanding its Capabilities and Potential Benefits for Healthcare Organizations”, In: Technological Forecasting & Social Change 126, pp. 3-13, 2018, DOI:10.1016/j.techfore.2015.12.019.

Using Big Data: Analysis, Challenges, and Innovations in Different Industries

Youse Rahimy Akhondzadeh

Department of Computer and Information Technology Engineering, Qazvin Branch, Islamic Azad University, Qazvin, Iran

Abstract

This article examines the impact of developments related to the analysis of big data in various industries, highlighting their role in innovation, operational efficiency and conscious decisions. First, by reviewing the importance of large data, we discuss key programs for health care, financial and smart cities, and show how data driven insights are transforming these areas. However, the effective implementation of the analysis in these areas is full of challenges such as data privacy, the complexities of technology integration into various industries, as well as skill gaps in the specialized areas of big data. We will also discuss recent innovations, such as artificial intelligence and machine learning, cloud computing and edge calculations and quantum computing potential, which are advancing big data capabilities and access. This article also considers future prospects by emphasizing the importance of data, regulatory adaptations, and increasing demand for specialized skills. Using big data, rganizations can guide significant progress in various industries and pave the way for a data driven future.

Keywords: Big Data, Database, Information Systems