



## چالش ها و مزایای استفاده از تجزیه و تحلیل داده های بزرگ در ابر برای هوش تجاری

شهناز محمدزاده،<sup>۱</sup> زهرا آزادی،<sup>۲</sup> آذر لواف پور نوری،<sup>۳</sup> فرشید صهبا،<sup>۴</sup>

۱ کارشناسی ارشد فناوری اطلاعات موسسه غیرانتفاعی غیاث الدین جمشید کاشانی. پست الکترونیکی:  
mohammadzadeh365@gmail.com

۲ کارشناسی ارشد فناوری اطلاعات موسسه غیرانتفاعی غیاث الدین جمشید کاشانی. پست الکترونیکی:  
Mehrnoosh.azadi@yahoo.com

۳ کارشناسی ارشد فناوری اطلاعات موسسه غیرانتفاعی غیاث الدین جمشید کاشانی. پست الکترونیکی:  
nouriazar1@gmail.com

۴ مدرس موسسه غیرانتفاعی غیاث الدین جمشید کاشانی. پست الکترونیکی:  
f\_sahba@yahoo.com

### چکیده

رایانش ابری و تجزیه و تحلیل داده های بزرگ بدون شک دو فناوری مهم برای ورود به صنعت فناوری اطلاعات در سال های اخیر است. به طرز شگفت انگیزی، این دو فناوری برای آینده در حال آمادگی هستند تا نتایج و مزایای بسیار مهمی را برای کسب و کار ارائه دهند. در حال حاضر محاسبات ابری در حال تغییر هستند و نحوه ارائه خدمات فناوری اطلاعات توسط شرکت های به اصطلاح ابر و نحوه تعامل کسب و کار و کاربران با منابع فناوری اطلاعات است. تجزیه و تحلیل داده های بزرگ نیاز به مقدار زیادی از محاسبات و هزینه های منابع تکنولوژی داده های بزرگ می باشد که برای مقاصد بسیاری از شرکت های کوچک و متوسط مقرون به صرفه نیست. در این مقاله، ما مزایا و چالش های موجود در راه اندازی تجزیه و تحلیل داده های بزرگ را از طریق محاسبات ابری ارائه می دهیم. ما استدلال می کنیم که محاسبات ابری می تواند از نیازهای ذخیره سازی و محاسبات تجزیه و تحلیل داده های بزرگ پشتیبانی کند. ما در مورد چگونگی ادغام این دو تکنولوژی غالب می توانیم فرآیند بزرگ داده کاوی را افزایش دهیم تا کسب و کار را قادر به بهبود فرایندهای تصمیم گیری کند.

### واژه های کلیدی

رایانش ابری، تجزیه و تحلیل داده های بزرگ، ابر تجزیه و تحلیل، امنیت، حریم خصوصی، هوش کسب و کار

### ۱- مقدمه

اصطلاح (BI) به تکنولوژی ها، برنامه ها و شیوه های جمع آوری، ادغام، تجزیه و تحلیل و ارائه اطلاعات کسب و کار اشاره دارد. هدف اصلی از هوش تجاری، پشتیبانی از تصمیم گیری بهتر و سریعتر کسب و کار است. سازمانها مجبور به گرفتن، و به دست آوردن اطلاعات خود برای حمایت از تصمیم گیری برای بهبود عملیات تجاری هستند [۱].

دنیای کسب و کار همیشه در حال تغییر است، امروزه بسیاری از شرکت ها با فشار رو به رشد برای توسعه و افزایش سرعت سعی در رقابت و کم کردن هزینه ها می باشند. محاسبات ابری که اخیرا به وجود آمده است، نحوه تغییر ارائه خدمات فناوری اطلاعات توسط شرکت ها و نحوه کسب و کار و کاربران را با منابع IT ارتباط می دهد. این نشان دهنده یک تغییر پارادایم است که مدل های خدمات انعطاف پذیر را ارائه می دهد که شرکت ها می توانند در مدل پرداختی به عنوان استفاده شما اشتراک نمایند. داده ها در جهان به طور نمادین رشد می کنند. داده های بزرگ یک اصطلاح تکاملی است که مقادیر زیادی از اطلاعات ساختار یافته، نیمه ساختار یافته و

<sup>1</sup> business intelligence  
<sup>2</sup> Information Technology



غیر ساختاری ممکن است برای اطلاعات مفید استخراج شوند. داده های بزرگ، داده هایی هستند که بیش از ظرفیت پردازش پایگاه های سنتی هستند. الگوهای پنهان، همبستگی ها و سایر بینش ها را کشف می کنند. در این مقاله، هدف ما بررسی تاثیرات محاسبات ابری و داده های بزرگ در مورد کسب و کار و تجزیه و تحلیل مزایا و چالش های آن برای شرکت ها می باشد. ابتدا، ما مفاهیم، مسائل و فن آوری ابر رایانه و داده های بزرگ را به طور جداگانه مرور می کنیم. سپس یک چارچوب را ارائه می دهیم که این دو تکنولوژی را ترکیب می کند تا یک پلت فرم ایده آل برای تجارت الکترونیکی بسازند. همچنین ما در مورد نقش داده های بزرگ در ارتقاء بحث می کنیم. زمینه های کاربردی اصلی تجارت الکترونیک مانند مدیریت مشتری، بازاریابی، پرداخت، زنجیره تامین و مدیریت می باشد.

محبوبیت رایانش ابری باعث شده است تا چندین طرح علمی و صنعتی برای بررسی قابلیت ها و پیشرفت های محاسبات ابری صورت گیرد. گزاره ارزش محاسبات ابری در مقایسه با سرمایه گذاری های پیش فرض یکی از زمینه های تحقیق کلیدی است. طرح های متعددی برای مشخص کردن مسائل امنیتی و چالش های رایانش ابری وجود دارد. چندین برنامه آموزشی وجود دارد که جنبه های مدل الکترونیکی کسب و کار رایانش ابری را بررسی می کند [۲].

## ۲- پارامتر محاسبات ابری

۲.۱. رایانش ابری چیست؟

بسیاری از محققان ابر رایانه را به صورت متفاوتی تعریف کرده اند. یکی از مهمترین تعاریف پذیرفته شده، تعریف موسسه استاندارد ایالات متحده (NIST) است. "رایانش ابری یک مدل برای دسترسی به شبکه گسترده و بر اساس تقاضا برای دسترسی به یک استخر مشترک از منابع محاسباتی قابل تنظیم (مانند شبکه ها، سرورها، ذخیره سازی، برنامه ها و سرویس ها) است که می تواند به سرعت و با حداقل تلاش مدیریتی منتشر شود. این مدل ابر شامل پنج ویژگی اساسی، پنج مدل خدماتی، و چهار مدل استقرار می باشد [3]".

۲.۲. ویژگی رایانش ابری

ابر رایانه دارای پنج ویژگی اصلی است. آنها قابلیت های تقاضا، دسترسی گسترده به شبکه، جمع آوری منابع، انعطاف پذیری سریع و خدمات اندازه گیری شده را دارند. این ها ویژگی هایی هستند که از پارادایم های دیگر محاسباتی آن را تشخیص می دهند.

قابلیت ها بر روی تقاضا: مصرف کننده می تواند به طور یک جانبه قابلیت های محاسباتی مانند زمان سرور و ذخیره سازی شبکه را به صورت خودکار و بدون نیاز به تعامل انسان با هر ارائه دهنده خدمات فراهم کند. دسترسی به شبکه گسترده: از قابلیت های مهم رایانش ابری در دسترس بودن شبکه های آن می باشد. و از طریق مکانیزم های استاندارد قابل دسترسی می باشند، و از سیستم عامل های نازک یا ضعیف مشتری استفاده می شود (به عنوان مثال، تلفنهای همراه، رایانه لوحی و ایستگاههای کاری).

جمع آوری منابع: منابع محاسباتی ارائه دهنده برای جمع آوری از چندین مصرف کننده با استفاده از یک مدل چند مستاجر، با منابع مختلف فیزیکی و مجازی به صورت پویا اختصاص داده شده و مجدداً به تقاضای مصرف کننده متصل می شوند.

انعطاف پذیری سریع: در برخی موارد توانایی ها را می توان به صورت اتوماتیک تضمین کرده و به سرعت به سمت درون و بیرون، با توجه به تقاضا، مقیاس پذیری کرد.

<sup>1</sup> National Institute of Standards and Technology



سرویس اندازه گیری: سیستم های ابری به صورت خودکار کنترل منابع را با استفاده از قابلیت اندازه گیری در برخی سطوح انتزاعی متناسب با نوع سرویس (مانند ذخیره سازی، پردازش، پهنای باند و حساب های کاربری فعال) کنترل و بهینه سازی می کنند [۱].

۲,۳. مدل های توسعه دهی ابری

مدل های گسترش استقرار ابر به طور گسترده به چهار مدل تقسیم می شوند: ابر خصوصی، ابر عمومی، ابر جامعه و ابر ترکیبی.

ابر خصوصی امن ترین راه برای استفاده از ابر رایانه می باشد. زیرساخت ابر برای استفاده منحصر به فرد توسط یک سازمان واحد که شامل چندین مصرف کننده (به عنوان مثال، واحدهای تجاری) است، ارائه می شود.

ممکن است مالکیت و مدیریت آن توسط یک سازمان، شخص ثالث یا ترکیبی از آنها باشد، و ممکن است در محل یا خارج از محل باشد. برای استفاده منحصر به فرد توسط یک جامعه خاص از مصرف کنندگان و سازمان هایی که نگرانی های مشترک دارند، ارائه می شود.

ابر عمومی برای استفاده عمومی توسط عموم فراهم شده است. ابر ترکیبی، ترکیبی از دو یا چند زیرساخت ابری متمایز (خصوصی، عمومی) هستند اما منحصر به فرد نیستند، اما با تکنولوژی استاندارد یا اختصاصی، داده ها و قابلیت حمل را فراهم می کند [۱].

۲,۴. مدل تحویل خدمات ابر

خدمات مبتنی بر ابر بر سه دسته تقسیم می شوند: داده ها به عنوان یک سرویس (DaaS)، نرم افزار به عنوان یک سرویس (SaaS)، پلت فرم به عنوان یک سرویس (PaaS) و زیرساخت به عنوان یک سرویس (IaaS). نرم افزار به عنوان یک سرویس (SaaS): یک مدل است که کاربر را به دسترسی به برنامه هایی که در حال حاضر در حال اجرا در ابر هستند فراهم می کند. دسترسی توسط مشتری های ابر به دست می آید و کاربران ابر نمی توانند زیرساخت های کاربردی را مدیریت کنند و از این طریق راه را برای نصب و اجرای برنامه بر روی رایانه های شخصی ابر خود حذف می کنند.

پلت فرم به عنوان سرویس (PaaS): یک مدل است که به کاربران اجازه توسعه محیط عملیاتی خود را می دهد. کاربران می توانند برنامه های ساخته شده داخلی خود را توسعه و اجرا کنند. خدمات عبارتند از سیستم عامل، محیط اجرایی زبان برنامه نویسی، پایگاه داده ها و سرورهای وب.

زیرساخت به عنوان یک سرویس (IaaS): یک مدلی است که برای کاربران زیرساخت های مجازی، به عنوان سرورها و فضای ذخیره سازی داده، فراهم می کند. مجازی سازی در این حالت نقش مهمی را ایفا می کند، با اجازه دادن به ارائه دهندگان ابر IaaS به ارائه منابع بر اساس تقاضا، استخراج آنها از استخرهای بزرگ نصب شده را امکان پذیر می کند.

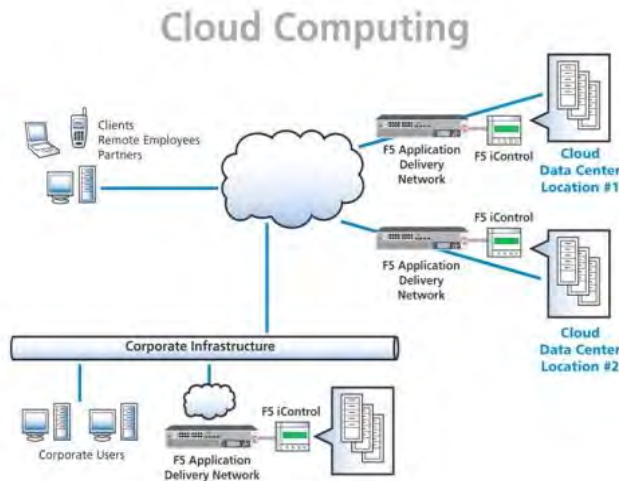
داده ها به عنوان یک سرویس (DaaS): یک مدلی است که در آن داده ها به آسانی از طریق پلتفرم مبتنی بر ابر قابل دسترسی هستند. DaaS به سادگی یک راه جدید برای دسترسی به اطلاعات مهم کسب و کار در یک مرکز داده را فراهم می کند [۲].

<sup>1</sup> Data as a service

<sup>2</sup> Software as a service

<sup>3</sup> Platform as a Service

<sup>4</sup> Infrastructure as a service



شکل ۱ معماری محاسبات کلی ابر را نشان می دهد.

## ۲.۵. مزایای رایانش ابری

هزینه بهره وری - این یکی از بزرگترین مزیت های محاسبات ابری می باشد که با حذف سرمایه گذاری بر روی نرم افزار یا سرویس های مستقل به دست می آید. با استفاده از توانایی های ابر، شرکت ها می توانند در هزینه های صدور مجوز و همچنین هزینه های اضافی مانند هزینه ذخیره سازی داده ها و به روز رسانی نرم افزار، مدیریت و غیره صرفه جویی کنند. اجاره یک زیر ساخت می تواند حس خوب مالی را به شما بدهد. در دسترس بودن مداوم - ابرهای عمومی قادر هستند خدمات خود را حتی به نتهایی ترین کاربر خود ارائه دهند. این رویکرد دسترسی آسان به اطلاعات را در مناطق مختلف زمانی و مکانی فراهم می کند و نیازهای کاربران را برآورده می سازد. به عنوان یک مزیت جانبی، با رونق همکاری، مشاهده و تغییر اسناد و فایل های مشترک ساده تر از همیشه در دسترس می باشد. علاوه بر این، یکی از موارد تضمینی بارگذاری سرویس می باشد و همچنین دسترسی مستمر منابع را فراهم می سازد. فروشنندگان ابرهای مختلف معمولاً از سرورهای متعدد برای حداکثر انحراف استفاده می کنند. در صورت شکست سیستم، موارد جایگزین به طور خودکار بر روی ماشین های دیگر ایجاد می شوند.

مقیاس پذیری و انعطاف پذیری - مقیاس پذیری یک ویژگی داخلی ساخته شده برای توسعه دهندگان است. نمونه های ابری بطور اتوماتیک تنها در زمان مورد نیاز مستقر می شوند و بنابراین شما تنها برای برنامه ها و ذخیره داده ها هزینه پرداخت می کنید. همچنین ابرها می توانند با درخواست سیستم در برابر تغییرات فناوری اطلاعات مقیاس پذیر تر باشند.

استقرار سریع و سهولت ادغام - یک کسب و کار مبتنی بر ابر می تواند در عرضیتن ها چند ساعت به راه چیده و چند ماه و بدون صرف مقیاس زیاد از پول بگازار چشوده و ارزش خود طی هیه از میله اچ د مدچیلر ملت. در ه مان لب، عرفهیه یک بار بر اود درس بقابل فطره و به هذف دوره هاچرات هارنگه امهی افند

انعطاف پذیری و افزونگی - استقرار ابری معمولاً بر روی یک معماری قوی انجام می گیرد و انعطاف پذیری و افزونگی را به کاربران فراهم می آورد. ابر به طور خودکار شکست بین پلتفرم های سخت افزاری را ارائه می دهد و این در حالی است که خدمات بازبازی فاجعه را نیز اغلب شامل می شود.

افشای ظرفیت ذخیره سازی از چ - ابر میتواند اطلاعات بیشتری نسبت به یک رایانه شخصی ذخیره کند و ظرفیت بودن فضای ذخیره سازی از دیدها باعث نگرانی است اما



در عین حال سخت افزار رایانه ی خود را در جهت نیاز های کسب و کاری ارتقا می دهد و هزینه کلی فناوری اطلاعات را کاهش می دهد.

### ۳. تجزیه و تحلیل داده های بزرگ

۳.۱. *Big Data* چیست؟

اصطلاح *Big Data* به مجموعه ای از داده ها اشاره می کند که به قدری بزرگ و پیچیده است که به روش های سنتزی و ابزارهای سنتزی قابل درک برای استفاده بیشتر است. فرایندهای اصلی داده های بزرگ شامل گرفتن، نگهداری، ذخیره سازی، جستجو، اشتراک گذاری، انتقال، تجزیه و تحلیل و تجسم است [۲].

به تازگی اهمیت این حوزه توجه زیادی را به خود جلب کرده است، زیرا کسب و کار اطلاعات مفید و بینش بهتر از داده های ساختاری و غیر ساختاری را می دهد که ممکن است منجر به تصمیم گیری بهتر اطلاعات شود. در زمینه کسب و کار، تجزیه و تحلیل داده های بزرگ روند بررسی مجموعه داده های بزرگ برای کشف الگوهای پنهان، همبستگی ناشناخته، روند بازار، ترجیحات مشتری و دیگر اطلاعات کسب و کار مفید است. پیشرفت های امروز در تکنولوژی همراه با پیشرفت های اخیر در الگوریتم های تجزیه و تحلیل داده ها و رویکردها، سازمان ها را قادر به استفاده از داده های تجزیه و تحلیل شده داده های بزرگ ساخته اند. بعضی از مسائل مهم در استفاده از تجزیه و تحلیل داده های بزرگ شامل کیفیت داده، ذخیره سازی، تجسم و پردازش می شود.

برخی از نمونه های کسب و کار از داده های بزرگ شامل محتوای رسانه های اجتماعی، جزئیات تلفن همراه، داده های معاملات، پررند، های سلامت، استاد مالی، اینترنت اشیا و اطلاعات آب و هوا است [۴].

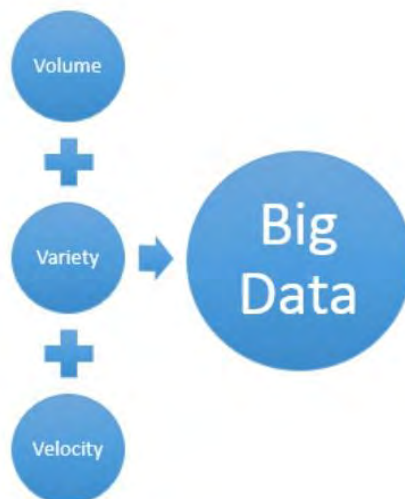
۳.۲. فناوری های *Big Data*

برای پشتیبانی از تجزیه و تحلیل داده های بزرگ، یک پلت فرم محاسباتی باید ۳ معیار زیر را داشته باشد، همانطور که  $V^3$  ها در شکل ۲ نشان داده شده است.

تنوع: این پلت فرم از طیف وسیعی از داده ها پشتیبانی می کند و شرکت ها را قادر می سازد تا این داده ها را به صورت فرمت اصلی و با ابزارهای گسترده استفاده کنند تا آن را به سایر فرمت های مورد نظر تبدیل کنند.

سرعت: این پلت فرم می تواند داده ها را در هر سرعت، یا جریان های کم تاخیر، مانند داده های سنسور یا سهام یا حجم زیادی داده های دسته ای، اداره کند.

حجم: این پلت فرم می تواند حجم زیادی از اطلاعات در حالت استراحت یا جریان را مدیریت کند [۲].





استخراج داده های سنتی شامل یافتن الگوهای جالب از مجموعه داده ها است در حالی که تجزیه و تحلیل داده های بزرگ شامل ذخیره سازی و پردازش گسترده ای از مجموعه داده های بزرگ است. به طور سنتی *Hadoop* و *MapReduce* از دو تکنولوژی محبوب برای تجزیه و تحلیل داده های بزرگ استفاده می کنند. ابزارها و فن آوری های بیشتری برای پردازش داده های بزرگ در دسترس هستند. نمونه هایی از آن ها عبارتند از ذخیره سازی داده های BI میزبان *Redshift* آمازون، سرویس تجزیه و تحلیل داده های *BigQuery* گوگل، پلتفرم *Cloud Bluemix* آی بی ام و سرویس پردازش اطلاعات *Kinesis* آمازون. چشم انداز آینده داده های بزرگ، ترکیبی از مکان و ابر است، جایگزینی برای پایگاه های داده های رابطه ای مبتنی بر *SQL*، به نام پایگاه داده های *NoSQL* (نه تنها *SQL*)، به سرعت به عنوان ابزاری برای استفاده در انواع خاصی از تحلیل داده های بزرگ در برنامه های کاربردی به دست می آید [۴].

۳،۳. مزایای داده بزرگ

واقعیت این است که داده های سازمانی ارزشمند خارج از فایروال شرکتی قرار می گیرند و نگرانی های جدی را به وجود می آورند. بعضی از چالش های رایج در زیر مورد بحث قرار می گیرند کاهش هزینه- فن آوری های داده های بزرگ مانند *Hadoop* و تجزیه و تحلیل مبتنی بر ابر می تواند منجر به کاهش قابل توجه هزینه ها باشد. در حالی که مقایسه بین تکنولوژی داده های بزرگ و معماری های سنتی (انبار داده ها و مارک ها) به علت تفاوت در عملکرد، دشوار است، اما مقایسه قیمت تنها می تواند به منظور بهبود مقیاس در نظر گرفته شود. برای مثال، شرکت ها به جای پردازش و ذخیره مقادیر وسیعی از داده های جدید در یک انبار داده، از این کار برای استفاده از خوشه های *Hadoop* استفاده می کنند و اطلاعات را به نبارهای سازمانی به صورت مورد نیاز برای برنامه های کاربردی تحلیلی تولید می کنند. تصمیم گیری سریع تر و بهتر- تجزیه و تحلیل ها همیشه در جهت بهبود تصمیم گیری ها تلاش می کنند، و داده های بزرگ آن ها را تغییر نمی دهند. دنبال کردن تجزیه و تحلیل داده های بزرگ باعث می شود مدیران کسب و کار تصمیم گیرندگان خوبی باشند. سازمان های بزرگ به دنبال تصمیم های سریع و بهتر با داده های بزرگ هستند و آنها را پیدا می کنند. چندین شرکت بر اساس سرعت *Hadoop* و تجزیه و تحلیل حافظه ای، بر سرعت بخشیدن به تصمیمات موجود تمرکز دارند.

محصولات و خدمات جدید- شاید استفاده جالب از تجزیه و تحلیل داده های بزرگ، ایجاد محصولات و خدمات جدید برای مشتریان است. شرکت های اینترنتی این کار را برای یک دهه یا بیشتر انجام داده اند، اما اکنون اغلب شرکت های آنلاین نیز این کار را انجام می دهند.

پیشنهاد محصول- بدیهی است که پذیرش داده های بزرگ و تجزیه و تحلیل ها ثابت شده است که یک استراتژی بسیار قدرتمند برای کسب و کارهای آنلاین است. تأثیر داده های بزرگ مشتریان در کسب و کار و تبدیل شدن به ابزار بسیار مهم و اقتصادی برای تقویت یک کسب و کار است. ذخیره سازی و کار بر روی داده های بزرگ همیشه یک چالش برای هر تجارت بوده است. داده های بزرگ جاده ای را برای مدیریت چنین اطلاعات بزرگ ساخته اند و کسب و کار را بسیار ساده و سودآور کرده اند.

تشخیص تقلب- تجزیه و تحلیل با کارایی بالا صرفاً تکنولوژی دیگری نیست. این نشان دهنده یک تغییر انقلابی در نحوه سازماندهی داده ها است. با استفاده از گزینه های جدید محاسباتی توزیع شده مانند پردازش در حافظه بر روی سخت افزار کالا، کسب و کار می تواند به یک راه حل تجزیه و تحلیل داده های انعطاف پذیر و مقیاس پذیر در زمان واقعی با هزینه معقول دسترسی داشته باشد. این حقیقت دارد که شرکت های بیمه در کسب و

خخیص تقلب [۱]، [۳].



## ۴. تجزیه و تحلیل داده های بزرگ در ابر

تجزیه و تحلیل داده های بزرگ مبتنی بر ابر یک مدل خدماتی است که در آن عناصر فرآیند تجزیه و تحلیل داده های بزرگ از طریق ابر عمومی یا خصوصی ارائه می شوند. این ابزار از طیف وسیعی از ابزار و تکنیک های تحلیلی برای کمک به کسب و کارها از داده های عظیم استفاده می کند و آن را به گونه ای ارائه می دهد که به راحتی دسته بندی شده و به راحتی از طریق یک مرورگر وب قابل دسترسی است. چنین برنامه های کاربردی و خدمات تجزیه و تحلیل داده مبتنی بر ابر معمولاً تحت یک مدل قیمت گذاری مبتنی بر اشتراک یا هزینه (پرداخت برای استفاده) ارائه می شود. مدل سرویس Cloud Analytics به عنوان سرویس (CLAsS) نامیده می شود. در این مدل، تجزیه و تحلیل به راحتی از طریق یک پلت فرم محاسبات ابری قابل دسترسی است. چنین سرویس تجزیه و تحلیل داده مبتنی بر ابر، کسب و کار را قادر می سازد تا فرآیندهای خودکار را در هر زمان و هر کجا بر اساس آن فعال کند. نمونه هایی از این محصولات و خدمات تجزیه و تحلیل مبتنی بر ابر شامل ابزار داده های میزبانی، هوش تجاری نرم افزار به عنوان یک سرویس (SaaS BI) و تجزیه و تحلیل رسانه های اجتماعی مبتنی بر ابر است. داده های ذخیره شده در یک پایگاه داده مبتنی بر ابر می تواند به کسب و کار با فرآیندهای تصمیم گیری خود کمک کند.

با داده های بزرگ مبتنی بر ابر، تحلیل گران نه تنها داده های بیشتری را برای کار با آن، بلکه توان پردازشی برای اداره تعداد زیادی از پرونده ها با ویژگی های متعددی را دارند. [dataacademy.ir](http://dataacademy.ir) وب سایت های بازدید شده یا مکان را به صورت روزانه بررسی کنند [۳].

۴.۱. مزایای عمده برای سازمان های تجاری  
بر اساس سابقه این پژوهش: همانطور که از نام آن پیداست، سازمانها می توانند ذخیره سازی یا سرویس را با یک کلیک بر دکمه بدون هیچ گونه کمک انسانی گسترش دهند. سازمان ها می توانند زیرساخت داده های بزرگ را با بیشترین سرعت ممکن ایجاد کنند.

داده و اطلاعات در شبکه: اطلاعات در سراسر شبکه در دسترس است و می تواند در هر زمان از طریق شبکه و دستگاه های مختلف مانند لپ تاپ، تلفن همراه، تبلت و غیره قابل دسترس باشد.

جمع آوری منابع: منابع ارائه دهنده گروه بندی شده و به طور موثر توسط مدل چند مستاجر استفاده می شود. منابع ذخیره سازی از چ، ف، VM و غیره

کشش سریع: منابع (هر دو سخت افزار و نرم افزار) می توانند در مدت زمان کوتاهی کارآمد و مؤثر باشند. مشتریان می توانند منابع را برای هر مقدار و در هر زمان خریداری کنند.

هزینه موثر: استفاده از منابع می تواند نظارت شود و براساس نیاز استفاده می شود. این سیستم بسیار شفاف است که باعث می شود ارائه دهنده و کاربر راحت تر آن را بپذیرند. فن آوری های داده های بزرگ مانند

Hadoop و تجزیه و تحلیل مبتنی بر ابر هزینه قابل توجهی را در هنگام ذخیره سازی داده های بزرگ کاهش می دهند به علاوه آنها می توانند روش های کارآمدتر برای انجام کسب و کار را شناسایی کنند [۱].

۴.۲. داده های بزرگ و چالش های رایانه محاسباتی

واقعیت این است که داده های سازمانی ارزشمند خارج از فایروال شرکتی قرار می گیرند و نگرانی های جدی را به وجود می آورند. بعضی از چالش های رایج در زیر مورد بحث قرار می گیرند:

ذخیره سازی داده - ذخیره و تجزیه و تحلیل حجم زیادی از اطلاعات که برای یک شرکت ضروری است، نیاز به یک زیرساخت سخت افزاری گسترده و پیچیده دارد. با رشد مداوم داده ها، دستگاه ذخیره سازی داده ها به طور

شرکت های ابری ظرفیت ذخیره سازی زیادی را برای



مزیت رقابتی دنبال می کنند.

کیفیت داده- دقت و دسترسی به موقع داده ها برای تصمیم گیری بسیار مهم است. داده های بزرگ تنها هنگامی مفید هستند که فرایند مدیریت اطلاعات برای تضمین کیفیت داده ها اجرا شود.

امنیت و حریم خصوصی- امنیتی از نظر اچ ای داده ها چ بی ریلت. سازمانها باید از داده های بزرگ حس بیشتری داشته باشند، بخشی از داده های حساس خود را به داده های بزرگتر تقسیم می کنند. برای انجام این کار، شرکت ها باید شروع به ایجاد سیاست های امنیتی کنند که خود تنظیم باشند. این سیاستها باید روابط اعتماد موجود را ارتقا دهند و داده ها و به اشتراک گذاری منابع را در سازمان ها ارتقا دهند، در حالی که اطمینان حاصل شود که تجزیه و تحلیل داده ها بهینه و محدود به چنین سیاست هایی نیست.

هکرها و حملات مختلف به زیرساخت ابرها حتی اگر فقط یک سایت مورد حمله قرار گیرند، بر چندین مشتری تأثیر می گذارد. این خطرات را می توان با استفاده از برنامه های امنیتی، سیستم های فایل رمزگذاری شده، نرم افزار از بین بردن داده ها و خرید سخت افزار امنیتی برای ردیابی رفتار غیر معمول در سرورها کاهش داد.

تحول خدمات و صورتها با توجه به ماهیت تقاضای خدمات، ارزیابی هزینه های انجام شده دشوار است. بودجه بندی و ارزیابی هزینه ها بسیار دشوار خواهد بود مگر اینکه ارائه دهنده دارای معیارهای مناسب و قابل مقایسه برای ارائه باشد. توافقنامه سطح خدمات (SLA) ارائه دهنده کافی برای تضمین دسترسی و مقیاس پذیری نیست. شرکت ها تمایلی به تغییر ابری بدون تضمین کیفیت خدمات قوی نخواهند داشت.

سازگاری و قابلیت حمل و نقل باید هر زمان که می خواهند، از مهاجرت به داخل و خارج از ابر و ارائه دهندگان خدمات تعویض استفاده کنند، و نباید هیچ قبل در آن وجود داشته باشد. سرویس های ابر محاسباتی باید توانایی همگام سازی با IT را داشته باشند [۱۲].

لها ت اطمینان و ریک هستند رس های ارائه دهندگان ابر هنوز سرویس خسته و خاموس ندارند؛ این منجر به قطع شدن مکرر می شود. مهم است که خدمات با استفاده از ابزار داخلی یا شخص ثالث نظارت شود. حیاتی است که برنامه ای برای نظارت بر استفاده، SLA، عملکرد، قابلیت اطمینان و وابستگی تجاری این سرویس ها باشد.

عمی رد و پق اچ لده هکهنه کسب و کار باید در سخت افزارها کاهش یابد، و بیشتر برای پهنای باند صرف گردد. این می تواند هزینه کم برای برنامه های کاربردی کوچکتر باشد، اما می تواند برای برنامه های کاربردی با شدت بسیار زیاد باشد. ارائه اطلاعات فشرده و پیچیده در مورد شبکه نیاز به پهنای باند کافی دارد.

تمام این چالش ها نباید به عنوان بلوک های جاده در جستجوی محاسبات ابر محسوب شوند. این مهم نیست که قبل از اتخاذ تکنولوژی، این مسائل و راه های ممکن را در نظر بگیرید [۱].

## نتیجه گیری

کسب و کارها از تجزیه و تحلیل داده ها استفاده می کنند تا به راه انداختن استراتژی خود برای به حداکثر رساندن سود و حمایت از فرایندهای تصمیم گیری خود کمک کنند. امروزه به طور گسترده ای پذیرفته شده است که محاسبات ابری و فناوری های داده بزرگ، دو تکنولوژی غالب هستند که جهان تجارت را شکل می دهند. ابر تنها حرف زدن نیست - این یک واقعیت زندگی است که بر هر جنبه ای از صنعت فناوری تاثیر می گذارد. فن آوری های داده های بزرگ ارائه شده از طریق ابر رایانه اجازه می دهد تا کسب و کار تصمیم گیری فعالانه، مبتنی بر دانش داشته باشند و اجازه می دهند تا روند و رفتار های آینده را پیش بینی کنند. شرکت ها قادر به ذخیره داده های خود از راه دور و دسترسی به داده ها و خدمات از هر کجا و در هر زمان خواهند بود. علاوه بر این، تجزیه و تحلیل داده ها مبتنی بر ابر زیرساخت هایی را فراهم می کنند که شرکت ها در غیر این صورت باید خود آن را از ابتدا می ساختند [۲].

حفظ کسب و کار و رقابت با ارائه مزایای بسیاری از جمله





بهره وری هزینه، جمع آوری منابع، خدمات به صورت تقاضا، الاستیسیته سریع و سهولت مدیریت است. علی رغم این مزایا، برخی از چالش‌ها و خطاها، به خصوص در ارتباط با حفظ حریم خصوصی و امنیت وجود دارد. قبل از سرمایه‌گذاری در تجزیه و تحلیل داده‌های بزرگ مبتنی بر ابر، سازمان باید به طور کامل آنچه را که درگیر هستند، درک کنند. سرمایه‌گذاری در تجزیه و تحلیل‌های ابر می‌تواند برای سازمان سودمند باشد، اما برنامه‌ریزی مناسب لازم است تا اطمینان حاصل شود که تمام مراحل عناصر تحلیلی پوشش داده می‌شود [۱].

## مراجع

- [1] B. M. Balachandran and S. Prasad, "Challenges and Benefits of Deploying Big Data Analytics in the Cloud for Business Intelligence," *Procedia Comput. Sci.*, vol. 112, pp. 1112–1122, 2017.
- [2] C.-W. Tsai, C.-F. Lai, H.-C. Chao, and A. V. Vasilakos, "Big data analytics: a survey," *J. Big Data*, vol. 2, no. 1, p. 21, 2015.
- [3] M. Kubina, G. Koman, and I. Kubinova, "Possibility of Improving Efficiency within Business Intelligence Systems in Companies," *Procedia Econ. Financ.*, vol. 26, no. 15, pp. 300–305, 2015.
- [4] S. Williams, "The Strategic Importance of Business Intelligence," *Bus. Intell. Strateg. Big Data Anal.*, pp. 51–68, 2016.

dataacademy.ir