مـدل پلتفـرم داده بـزرگ و نقـش آن در کیفیـت داده و هوشـمندی کسـب‌وکار

نجمه ملایی*: کارشناس ارشد مدیریت صنعتی، دانشگاه علامه طباطبایی، Email:n.mollaii@gmail.com
عباس طهماسبی: کارشناس ارشد مهندسی سیستمهای اقتصادی اجتماعی، دانشگاه جامع امام حسین (ع)، Email:abbas.tahmasebi@gmail.com

چکیده

هوش کسب‌وکار به استناد پایگاه‌های ذخیره‌سازی داده‌ها و تحلیل آن به تصمیم‌گیری در فعالیت‌های هوشمند تجاری و کسب‌وکار کمک می‌کند. در این پژوهش «فناوری‌های داده بزرگ» به مجموعه‌ای از زیرساخت‌های رایانش داده، ذخیره‌سازی داده، رویکرد تجزیه‌وتحلیل داده، مصورسازی داده، خودکارسازی داده، امنیت و حریم خصوصی تقسیم شده است. چالش‌های ناشی از عدم کیفیت داده نیز به دو دسته چالش‌های داخلی و خارجی تقسیم شده‌اند. چالش‌های داخلی بی‌کیفیتی داده‌ها با میزان آمادگی کسب‌وکار و فرهنگ سازمانی مرتبط بوده که استقرار پلتفرم داده بزرگ در ایجاد آمادگی و فرهنگ سازمان تأثیرگذار خواهد بود. در پلتفرم داده بزرگ پیشنهادی با توجه به دو فاکتور شناخت حوزه‌های داده و سنجش میزان آمادگی، از هر یک از فناوری‌های داده بزرگ، نرم‌افزار و سخت‌افزار متناسب انتخاب می‌گردد. جمع‌آوری داده متغیرهای مدل به وسیله پرسش‌نامه محقق ساخته صورت گرفته و حجم نمونه نیز به صورت نمونه‌گیری در دسترس تعداد ۳۷ نفر از مدیران و کارشناسان کسب‌وکار را شامل شده است. داده‌ها در روش معادلات ساختاری کمترین مربعات جزئی در فضای نرم‌افزاری اسمارت پی‌ال‌اس نگارش ۳ مدل‌سازی شده و نتایج حاکی از آن است که طراحی پلتفرم داده بزرگ و چالش‌های کیفیت داده نسبت به دیگر متغیرهای مورد بررسی دارای بالاترین میزان اثر بر هوشمندسازی کسب‌وکار هستند. همچنین با توجه به ماتریس اهمیت – عملکرد، طراحی پلتفرم داده بزرگ، فناوری‌های داده بزرگ و سنجش میزان پذیرش کسب‌وکار به ترتیب دارای بیشترین درصد اهمیت و عملکرد در مدل می‌باشند.

واژگان کلیدی: *کیفیت داده، پلتفرم داده بزرگ، هوشمندی کسب‌وکار، چالش‌های بی‌کیفیتی داده.*

Big Data Platform Model and its role in Data Quality and Business Intelligence

Mollaii Najmeh[1]*, Tahmasebi Abbas[2]

Abstract

Business intelligence contributes to decision making in intelligent business and business activities based on data storage and analysis. In this study, "Big Data Technologies" is divided into a set of infrastructures including computing, data storage, data analysis, data visualization, data automation, security, and privacy. The challenges of poor data quality are divided into two categories of internal and external challenges. The inadequacy of internal data challenges are related to the readiness level of business and organizational culture, and the deployment of a big data platform will impact the organization's readiness and culture. In the proposed big data platform, considering each of the two factors, acquisition of business data domains and business readiness assessment, software and hardware is selected proportional from each of the big data technologies. To collect the model data, a researcher-made questionnaire was used and a stratified random sampling method was used. In this research, 37 business managers and experts were selected as the sample for answering the questionnaire. Data was modeled in least squares modular structural equation modeling in SmartPlus Script 3 software environment and the results showed that designing a large data platform and data quality challenges comparing to other variables have the highest impact on business intelligence. Considering the importance of performance matrix, big data platform design, big data technologies, and business acceptance measurements have the highest percentages of importance and performance in the model, respectively.

Key words: *Data quality, big data platform, Business intelligence, Data quality challenges.of Crisis*

1 – Master of Industrial Management, Allameh Tabatabaei University, Email:n.mollaii@gmail.com
2 – Master of Science in Social–Economic Systems Engineering, Imam Hossein University, Email:abbas.tahmasebi@gmail.com

مقدمه

از زمانی که مفهوم «کیفیت»، میان‌رشته‌ای شده هیچ تعریف مورد توافقی برای کیفیت وجود ندارد [۱]. داده پایه‌ی اولیه در فعالیت‌های عملیاتی، فنی و تصمیم‌گیری است. همچنین داده‌ها منابع حیاتی در همه سازمان‌ها، کسب‌وکارها و کاربردهای دولتی است. کیفیت داده برای مدیران و فرآیندهای عملیاتی به‌منظور شناسایی عملکرد مربوطه بسیار اهمیت دارد [۲ و ۳]. کیفیت داده تعاریف متفاوتی در دوره‌ها و زمینه‌های مختلف داشته است. بر اساس مدیریت کیفیت، کیفیت داده، مناسب برای استفاده یا برآوردن نیازهای کاربر است یا کیفیت داده برآوردن نیازهای مشتری است [۴]. «مناسب برای استفاده» یک مفهوم چندبعدی است که هر دوی مفهوم ذهنی و معیار عینی بر اساس مجموعه داده‌های مورد درخواست را داراست [۵].

بی‌کیفیت شدن داده‌ها علل گوناگونی مانند خطای انسانی، فقدان منابع داخلی، استراتژی نامناسب، عدم‌کفایت تکنولوژی مرتبط در حال استفاده، فقدان تکنولوژی مرتبط، فقدان مهارت مورد نیاز برای استفاده صحیح از تکنولوژی موجود، فقدان ارتباطات داخلی بین واحدها، ناکافی بودن حمایت مدیر ارشد و بودجه ناکافی دارد [۶]. با کامل شدن فناوری اطلاعات و تکنولوژی‌های مربوط به آن سازمان‌ها باید از این اهرم به‌عنوان ارتباطات و تراکنش‌های امن و دقیق با مشتریان و تأمین‌کنندگان و ذینفعان استفاده کنند [۷]. هم‌اکنون بسیاری از سازمان‌ها به آهستگی برای اتخاذ فناوری‌های داده عمل می‌کنند [۸].

برای کسب‌وکارها تبدیل داده‌ها به دانش مطلوبیت دارد، اگر داده‌ها کیفیت بالایی داشته باشند این مطلوبیت افزایش می‌یابد [۶]. هوش تجاری، نه به‌عنوان یک محصول و نه به‌عنوان یک سیستم، بلکه به‌عنوان یک معماری و رویکردی جدید موردنظر است که البته شامل مجموعه‌ای از برنامه‌های کاربردی و تحلیلی است که به استناد پایگاه‌های داده عملیاتی و تحلیلی به اخذ و کمک به تصمیم‌گیری برای فعالیت‌های هوشمند تجاری و کسب‌وکار می‌پردازند [۹].

داده بزرگ اصطلاحی است که برای توصیف حجم وسیعی از داده‌های ساختاریافته، بدون ساختار و داده‌های نیمه ساختاریافته استفاده می‌شود. چالش‌های داده بزرگ شامل ذخیره‌سازی، مدیریت، تجزیه‌وتحلیل و بصری‌سازی است. بر طبق نظر شرکت IBM، در مقایسه با داده‌های ساخت‌یافته در کاربردهای تجاری امروزی، داده بزرگ شامل سه ویژگی عمده تنوع، حجم و سرعت است [۱۰]. فرایند طراحی پلتفرم داده بزرگ چیست و چه تأثیری بر کیفیت داده دارد؟ آیا با طراحی پلتفرم داده بزرگ تمامی چالش‌های بی‌کیفیتی داده مرتفع می‌شوند؟ آیا هوشمند ساختن کسب‌وکار، بدون طراحی پلتفرم داده بزرگ امکان‌پذیر است؟

مبانی نظری و بیان مفاهیم پژوهش

در میان طرح‌های داده بزرگ، در سال‌های اخیر صنایع رسانه، مخابرات، خدمات مسافرتی، فناوری‌های نوین و خدمات مالی و بانکداری بیشترین سرمایه‌گذاری‌ها را به خود اختصاص داده‌اند. یکی از ویژگی‌های مشترک همه‌ی این صنایع داشتن مشتریان بی‌شمار برخط و برون‌خط است. نتایج مطالعه‌ای در سال ۲۰۱۳ بر روی ۱۲۱۷ شرکت از نه کشور جهان نشان می‌دهد [۱۱]:

- حدود نیمی از شرکت‌هایی که در داده‌های بزرگ سرمایه‌گذاری کرده‌اند، انتظار بازگشت سرمایه[۱] بیش از ۲۲ درصد را داشته‌اند.
- هرچند فعالیت‌های فروش و بازاریابی بیشترین سرمایه‌گذاری‌ها را در داده‌های بزرگ انجام داده‌اند، اما بخش‌های تدارکات و عملکردهای مالی بیشترین بازگشت سرمایه را خواهند داشت.
- چالش‌های فرهنگی و سازمانی به‌عنوان بزرگ‌ترین مشکل در طرح‌های داده بزرگ مطرح‌شده‌اند.
- طرح‌های مرتبط با داده‌های خارجی[۲] و بدون ساختار منجر به بازگشت سرمایه‌ی بیشتری برای شرکت‌ها شده است.

مک‌کینزی (۲۰۱۱)[۱۲] ارزش ایجاد شده توسط داده‌های بزرگ را چنین خلاصه کرد: اگر داده‌های بزرگ می‌توانست به‌صورت خلاقانه و مؤثر برای بهبود کیفیت و کارایی استفاده شود ارزش بالقوه‌ی حاصل از داده‌ها در صنعت دارویی ایالات‌متحده ممکن بود از ۳۰۰ میلیارد دلار عبور کند و هزینه‌ی صرف شده‌ی سیستم سلامت ایالات‌متحده را بیشتر از ۸ درصد کاهش دهد. گوگل چهل‌وپنج دسته مدخل جستجوی اطلاعات مرتبط با شیوع آنفلوانزا پیدا کرد و آن‌ها را در مدل‌های ریاضی خاصی ترکیب کرد تا شیوع آنفلوانزا را پیش‌بینی و حتی مکان‌هایی را پیش‌بینی کند که آنفلوانزا آنجا شیوع پیدا خواهد کرد [۱۳]. در سال ۲۰۱۴ تعداد ۷۳ درصد از ۳۰۲ شرکت مورد بررسی موسسه‌ی تحقیقات گارتنر در حال سرمایه‌گذاری در داده‌های بزرگ بوده‌اند که این میزان، ۶۴ درصد نسبت به تعداد آن در سال ۲۰۱۳ رشد داشته است [۱۴]. دلایلی مانند افزایش کارایی، بهبود رضایت مشتری، توانایی تصمیم‌گیری آگاهانه‌تر، کاهش هزینه‌ها، حفظ اعتبار سازمان‌ها، افزایش نوآوری در داده، کشف فرصت‌های بازار از طریق پروفایل مشتری و کاهش ریسک تقلب سازمان‌ها را وامی‌دارد تا برای ایجاد استراتژی داده‌های باکیفیت برنامه‌ریزی کنند [۶]. همان‌طور که در جدول ۱ مشاهده می‌شود، چالش‌های کیفیت داده را می‌توان به

جدول ۱: چالش‌های داخلی و خارجی تقابل بهبود کیفیت داده [۶]

چالش‌های داخلی سازمان	چالش‌های خارجی سازمان
مهارت‌ها / دانش	جمع‌آوری داده
نیروی انسانی	امنیت داده
شناخت ابزار کیفیت داده	مصورسازی داده
تدوین استراتژی بلندمدت	مدیریت حجم داده
بودجه / سرمایه	مدیریت تنوع داده
بازگشت سرمایه	مقررات حاکم
حسابرسی	آماده‌سازی داده
انسجام داده	عدم شناخت
عدم شناخت	-

جدول ۲- نمونه‌هایی از کاربرد داده‌ی بزرگ [۱۶]

کاربرد داده بزرگ	مشخصات	منابع داده
مانیتورینگ و بهینه‌سازی شبکه‌ی انرژی	گلوگاه داده	داده‌های حسگر حاصل از کنتورهای هوشمند
	گلوگاه محاسبات	
	حجم زیاد داده	
تشخیص تقلب‌ها در کارت‌های اعتباری	گلوگاه داده	داده‌های نقطه‌ی فروش یا کارت‌خوان‌ها (POS)
	گلوگاه محاسبات	پروفایل مشتریان/اطلاعات مشتریان
	حجم زیاد داده	سوابق تراکنش‌ها
	موازی‌سازی	مدل‌ها و الگوهای پیش‌بینی‌شده
	تنوع و گوناگونی داده‌ها	
پروفایل اطلاعات	حجم زیاد داده‌های موازی	منابع منتخب برای پیشنهادهای تجاری
خوشه‌بندی و طبقه‌بندی مشتریان	گلوگاه داده	پروفایل مشتریان/اطلاعات مشتریان
	گلوگاه محاسبات	سوابق تراکنش‌ها
	حجم زیاد داده	مجموعه داده‌های افزایش‌یافته/ارتقاءیافته
	موازی‌سازی	
	تنوع و گوناگونی داده‌ها	
موتورهای توصیه‌گر	گلوگاه داده	پروفایل مشتریان/اطلاعات مشتریان
	گلوگاه محاسبات	سوابق تراکنش‌ها
	حجم زیاد داده	مجموعه داده‌های افزایش‌یافته/ارتقاءیافته
	موازی‌سازی	داده‌های مربوط به شبکه‌های اجتماعی
	تنوع و گوناگونی داده‌ها	
مدل‌سازی قیمت	گلوگاه داده	داده‌های نقطه‌ی فروش یا کارت‌خوان‌ها (POS)
	گلوگاه محاسبات	پروفایل مشتریان/اطلاعات مشتریان
	حجم زیاد داده	سوابق تراکنش‌ها
	موازی‌سازی	مدل‌ها و الگوهای پیش‌بینی‌شده

جنبه‌ها و رویکردهای تکنولوژی در حوزه هوشمندی کسب‌وکار و تجزیه‌وتحلیل، به نقطه‌ای از رشد و تکامل رسیده است که بخش وسیعی از کسب‌وکارها می‌توانند از قابلیت‌های بی‌شماری استفاده نمایند که در گذشته قابلیت‌های مزبور تنها مربوط به سازمان‌های بزرگ با بودجه‌های بسیار بالا بوده است. هوش تجاری یا هوش کسب‌وکار که قالب عمده‌تری را مانند استفاده‌های تجاری و غیرتجاری (نظامی و غیرانتفاعی) در بر دارد، عبارت است از بُعد وسیعی از کاربردها و تکنولوژی برای جمع‌آوری داده و دانش جهت زایش پرس‌وجو در راستای آنالیز بنگاه برای اتخاذ تصمیمات تجاری دقیق و هوشمند [۱۵]. مهم‌ترین چالش برای سازمان‌های امروز، هماوردی در فضای جهانی شده می‌باشد و کمترین چشم‌داشت برای هماورد در چنین میدانی، دانستن این موضوع است که (دیگران) به چه کاری می‌پردازند که این امر همان اساس هوش کسب‌وکار را تشکیل می‌دهد [۹].

۶ کاربرد از موارد کاربرد داده بزرگ را می‌توان در جدول ۲ مشاهده نمود. هریک از «موارد استفاده» (برنامه‌های کاربردی) اشاره شده در جدول ذیل مستلزم به‌کارگیری ملاحظات؛ ذخیره‌سازی (قابلیت مقیاس‌پذیری، قابلیت توسعه‌پذیری، قابلیت دسترسی، بازه خطا، قابلیت ورودی/ خروجی با سرعت بالا و قابلیت یکپارچگی)، تجهیزات داده بزرگ (تجهیزات سخت‌افزاری و ملزومات نرم‌افزاری)، گزینه‌های معماری، مشخصات عملکردی (حجم داده‌ها، قابلیت عملکرد و مقیاس‌پذیری، یکپارچگی داده، محدوده خطا، ناهمگونی و تحویل دانش) و چیدمان داده در پایگاه داده و عملکرد نرم‌افزار (قابلیت دسترسی، سرعت ترکیب و تجمیع، مناسب جهت فشرده‌سازی و سرعت بارگذاری داده) می‌باشد که در طراحی پلتفرم نهایی باید بررسی و لحاظ شوند.

قابلیت پذیرش داده بزرگ

میزان قابلیت پذیرش داده بزرگ از سوی سازمان با توجه به معیارهایی مانند امکان‌سنجی، توجیه‌پذیری و معقولیت، ارزش، قابلیت و یکپارچگی، نگه‌داری مستمر و توسعه پایدار مورد ارزیابی قرار می‌گیرد. هنگامی‌که بر اساس شرایط بازار، موانع کمتری برای پیاده‌سازی و استفاده از راه‌حل‌های داده بزرگ وجود دارد، به این معنا نیست که پیاده‌سازی و به‌کارگیری این تکنولوژی‌ها و فرآیندهای کسب‌وکار این حوزه کاری آسان و بدون مشکل

است. قبل از هر اقدامی به‌منظور سرمایه‌گذاری کلان در خصوص پیاده‌سازی و بهره‌برداری از داده بزرگ سؤالاتی به شرح ذیل مطرح می‌گردد تا میزان آمادگی سازمان جهت پذیرش و تعیین مشخصات پلتفرم ارزیابی شود و سپس با توجه به نوع و مشخصات داده‌های بزرگ ذخیره شده در پایگاه‌های داده و استراتژی صنعت مربوطه رویکردهای تحلیلی مناسب و به‌تبع آن نرم‌افزارهای مناسب بررسی و پیاده‌سازی می‌شود.

- سؤال ۱ (امکان‌سنجی): آیا کسب‌وکار، تکنولوژی‌های جدید و نوظهور را موردپذیرش، بررسی و آزمایش قرار می‌دهد؟
- سؤال ۲ (امکان‌سنجی): آیا کسب‌وکار، منابع سازمانی خود را برای رویارویی با چالش‌های پیش رو در حوزه تکنولوژی داده‌ی بزرگ در نظر می‌گیرد؟
- سؤال ۳ (توجیه‌پذیری و معقولیت داده): آیا کسب‌وکار، رویکرد داده‌محور [۳] (محوریت داده) و استفاده از منابع داده‌ای بیشتر جهت تحلیل و گزارش‌دهی را در آینده در دستور کار قرار خواهد داد؟
- سؤال ۴ (ارزش): آیا شاخص‌ها و روش‌های مشخصی برای اندازه‌گیری دستاوردهای حاصل از به‌کارگیری و پیاده‌سازی داده‌ی بزرگ تعیین و تعریف شده است؟
- سؤال ۵ (قابلیت یکپارچگی): چه مراحل و گام‌هایی برای سنجش یکپارچگی داده‌ی بزرگ با سایر اجزا در کسب‌وکار وجود دارد که باید طی شود؟
- سؤال ۶ (نگهداری مستمر و توسعه پایدار): چه برنامه‌ای برای شناسایی دقیق محدودیت‌ها و موانع به‌منظور مدیریت مستمر و نگهداری پایدار جهت پیاده‌سازی هرچه بهتر داده‌ی بزرگ می‌توان تدوین کرد؟ [۱۶].

فناوری‌های داده بزرگ

گروه «بررسی داده بزرگ»، یک طبقه‌بندی شش‌بعدی برای معرفی فناوری‌های داده، شامل داده‌ها، زیرساخت رایانش، زیرساخت ذخیره‌گاه، رویکرد تحلیلی، مصورسازی، امنیت و حریم خصوصی معرفی کرده است. در طبقه «داده‌ها» حوزه‌های مختلفی که داده‌های بزرگ در آن‌ها به وجود آمده و رشد می‌کنند بررسی می‌شود. الزامات نهفتگی [۴] و ساختار [۵] داده‌های کسب‌وکار در این طبقه بررسی می‌شوند. دلیل دسته‌بندی این حوزه‌ها این است که انتخاب زیرساخت و الزاماتی که برای انواع ویژه داده‌ها نیاز است درک شوند [۱۷]. ایجاد زیرساخت‌های مطرح‌شده، مستلزم ارزیابی، انتخاب، مهار فناوری و همچنین ترکیب، یکپارچه‌سازی، تولید و بومی‌سازی فناوری‌هاست [۱۸] و حوزه ویژه‌ای که در آن داده‌ها به وجود می‌آیند، انواعی از معماری را مشخص می‌کنند که برای ذخیره‌سازی، پردازش و انجام تحلیل‌ها روی آن نیاز است [۱۷]. با توجه به نوع و مشخصات داده‌های بزرگ ذخیره شده در کسب‌وکار مربوطه رویکردهای تحلیلی مناسب و به‌تبع آن نرم‌افزارهای مناسب بررسی و پیاده‌سازی می‌شود.

طراحی و ارائه مدل پلتفرم داده بزرگ و تأثیر آن بر چالش‌های بی‌کیفیتی داده‌ها

با توجه به ادبیات موضوع در این پژوهش فرآیند طراحی پلتفرم داده بزرگ مطابق شکل ۱ ترسیم شده است. در این فرآیند ابتدا حوزه‌های داده شناسایی می‌شوند. شناسایی ساختار داده‌ها و الزامات نهفتگی در کسب‌وکار مربوطه توسط متخصصان مربوطه قبل از سنجش میزان آمادگی سازمان معرفی شده است. در حقیقت شناخت حوزه‌های داده بزرگ مقدمه‌ای است که در هر سازمانی قبل از سنجش میزان آمادگی سازمان باید انجام گیرد.

در مرحله بعد میزان سنجش پذیرش داده بزرگ با ۶ شاخص؛ امکان‌سنجی، توجیه‌پذیری و معقولیت داده، ارزش، قابلیت یکپارچگی و نگهداری مستمر و توسعه‌ی پایدار آمده است.

در مرحله پایانی با توجه به حوزه‌های داده در کسب‌وکار نرم‌افزارهای متناسب از ۵ زیرساخت فناوری داده بزرگ شامل زیرساخت رایانش، زیرساخت ذخیره‌گاه، الگوهای تحلیل، امنیت و حریم خصوصی داده و بصری‌سازی انتخاب می‌شوند.

چالش‌های بی‌کیفیتی داده‌ها به دو دسته خارجی و داخلی تقسیم می‌شود (دو سنجه) که چالش‌های داخلی بافرهنگ سازمان مرتبط است و اهمیت بیشتری دارند و مطابق با ادبیات موضوع دارای ۹ سنجه می‌باشند. این چالش‌ها را می‌توان با میزان قابلیت پذیرش داده بزرگ از سوی کسب‌وکار برطرف کرد. چالش‌های خارجی نیز با ۸ سنجه معرفی شده‌اند که با توجه به ۳ ویژگی تنوع، سرعت و حجم داده بزرگ قطعاً با طراحی پلتفرم داده بزرگ می‌توان بر چالش‌های خارجی «مدیریت حجم داده» و «مدیریت تنوع داده» فائق آمد. همچنین «فناوری‌های داده بزرگ» از «طراحی پلتفرم داده بزرگ» شامل زیرساخت‌های رایانش و ذخیره‌گاه، الگوهای تحلیل، امنیت و حریم خصوصی داده‌ها و بصری‌سازی می‌باشد که با «جمع‌آوری داده»، «امنیت داده» و «مصورسازی داده» از چالش‌های خارجی بی‌کیفیتی داده‌ها انطباق دارد.

مطابق با «هوشمندسازی کسب‌وکار» می‌توان گفت از ملزومات هوشمند ساختن کسب‌وکار ذخیره‌سازی و تحلیل داده‌ها و پاسخگویی آنلاین از طریق داده‌ها است؛ که در شکل ۱ به سه سنجه نمایش داده شده است.

در شکل ۱ تأثیر چالش‌های کیفیت داده بر طراحی پلتفرم داده بزرگ با فلش‌ها در مدل مشخص شده است. با این توضیح که در فرآیند طراحی پلتفرم داده بزرگ چالش‌های بی‌کیفیتی داده مرتفع می‌شوند، هریک از چالش‌های داخلی و خارجی با قسمتی از اجزای فرآیند طراحی پلتفرم داده بزرگ در ارتباط بوده و درنهایت با طراحی پلتفرم داده بزرگ می‌توان حجم و تنوع داده را مدیریت (چالش‌های خارجی) نموده و ابزار کیفیت داده (چالش داخلی) را شناسایی و طراحی و باعث افزایش هوشمندی کسب‌وکار گردید.

روش پژوهش

پژوهش حاضر با توجه به هدف آن جزء پژوهش‌های کاربردی طبقه‌بندی می‌شود؛ زیرا نتایج حاصله از این پژوهش دارای موارد

جدول ۳: مشخصات جمعیت شناختی نمونه آماری

ردیف	سمت		میزان آشنایی با پلتفرم داده بزرگ					تحصیلات			آشنایی با هوشمندی کسب‌وکار				
	مدیر	کارشناس	خیلی کم	کم	متوسط	زیاد	خیلی زیاد	کارشناسی	ارشد	دکترا	خیلی کم	کم	متوسط	زیاد	خیلی زیاد
فراوانی	۱۱	۲۶	۳	۶	۹	۱۵	۴	۱۴	۱۷	۶	۶	۵	۶	۱۶	۴
درصد	۳۰	۷۰	۸	۱۶	۲۵	۴۰	۱۱	۳۸	۴۶	۱۶	۱۶	۱۳	۱۶	۴۴	۱۱

جدول ۴- بررسی معیارهای پایایی شاخص در مدل

متغیرها	Cronbach Alpha	Communality ۰/۷۱	CR
افزایش هوشمندسازی کسب‌وکار	۰/۸۵۰	۰/۲۲۰	۰/۹۰۶
سنجش میزان پذیرش کسب‌وکار	۰/۷۶۷	۰/۳۵۴	۰/۸۵۶
شناخت حوزه‌های داده: الزامات نهفتگی، ساختار داده‌ها	۰/۸۵۷	۰/۵۲۱	۰/۹۰۰
طراحی پلتفرم داده بزرگ	۰/۸۳۰	۰/۰۲۱	۰/۸۸۰
فناوری‌های داده بزرگ	۰/۸۵۶	۰/۱۲۴	۰/۹۰۲
چالش‌های خارجی بی‌کیفیتی داده‌ها	۰/۹۴۲	۰/۶۳۲	۰/۹۷۲
چالش‌های داخلی بی‌کیفیتی داده‌ها	۰/۷۵۴	۰/۴۵۵	۰/۸۹۰
چالش‌های کیفیت داده	۰/۷۴۳	۰/۳۱۸	۰/۹۰۱

کاربردی در حوزه‌های مختلف به‌ویژه صنعت بانک، بیمه و مخابرات می‌باشد و همچنین با توجه به نحوه گردآوری اطلاعات که از طریق پرسشنامه و در یک مقطع زمانی مشخص انجام گرفته است، از نوع پژوهش‌های پیمایشی - مقطعی است. جامعه آماری و نمونه آماری پژوهش حاضر مدیران و کارشناسان صنایع بانک، بیمه و مخابرات می‌باشند که به صورت نمونه‌گیری در دسترس تعداد ۳۷ نفر از ۵۰ نفر به‌عنوان حجم نمونه آماری انتخاب شد.

توصیف خصوصیات نمونه آماری

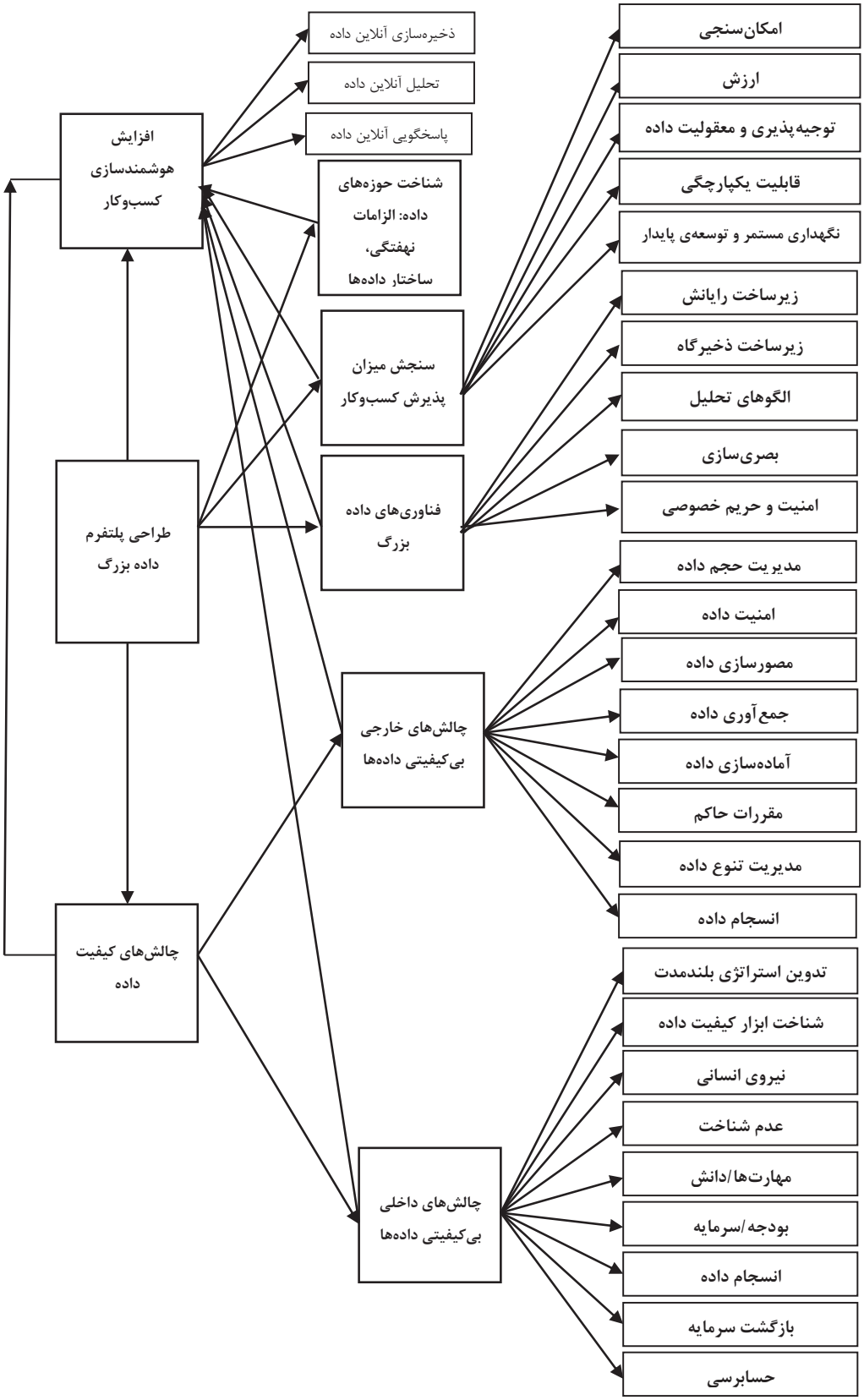
به‌منظور توصیف داده‌ها و اطلاعات جمعیت شناختی نمونه آماری از رایج‌ترین شاخص‌های آماری چون فراوانی و درصد برای بررسی چگونگی توزیع نمونه آماری از حیث متغیرهایی چون تأهل، سن، تحصیلات و میزان آشنایی با هوشمندی کسب‌وکار، آشنایی با پلت فرم داده بزرگ استفاده می‌شود. با توجه به اطلاعات گردآوری‌شده از ۳۷ پرسشنامه مطابق با جدول (۳) بیش از نیمی، یعنی در حدود ۷۰ درصد نمونه آماری کارشناس بوده، بیش از نیمی (۵۱ درصد) از نمونه آماری با پلتفرم داده بزرگ آشنا بوده و درصد بالایی ۶۲ درصد از نمونه آماری دارای سطح تحصیلات تکمیلی (ارشد و دکترا) بوده همچنین حدود ۵۵ درصد از نمونه آماری، آشنایی زیادی با هوشمندی کسب‌وکار داشتند.

سنجش مدل و تجزیه‌وتحلیل داده‌ها

در این پژوهش به‌منظور سنجش و ارزیابی مدل ارائه‌شده، داده‌های گردآوری‌شده با رویکرد معادلات ساختاری و روش کمترین مربعات جزئی (PLS-SEM) در فضای نرم‌افزاری Smart PLS نگارش ۳ تحلیل شده است. از این روش به‌عنوان مدل علی و تحلیل ساختار کوواریانس نیز یاد می‌شود. از طریق این روش می‌توان، قابل‌قبول بودن مدل‌های نظری را در جامعه‌های خاص با استفاده از داده‌های همبستگی غیرآزمایشی و آزمایشی آزمود [۱۹]. به این منظور سازه‌های «طراحی پلتفرم داده بزرگ» با یک سؤال (سنجه)، «شناخت حوزه‌های داده در کسب‌وکار» با دو سؤال (سنجه)، « فناوری‌های داده بزرگ» با پنج سؤال (سنجه)، «سنجش میزان پذیرش کسب‌وکار» با شش سؤال (سنجه)، « افزایش هوشمندسازی کسب‌وکار» با سه سؤال (سنجه)، «چالش‌های کیفیت داده» با دو سؤال (سنجه)، «چالش‌های خارجی بی‌کیفیتی داده‌ها» با هشت سؤال (سنجه) و «چالش‌های داخلی بی‌کیفیتی داده‌ها» با هشت سؤال (سنجه) به‌عنوان سازه‌ها در مدل کمترین مجذور مربعات تعریف و وارد نرم‌افزار شده‌اند. پیش از استفاده از نتایج مدل باید برازش‌های لازم در سطح مدل اندازه‌گیری و مدل ساختاری قابل‌قبول ارائه کرد که در ادامه ارائه می‌شود.

الف) برازش مدل اندازه‌گیری: برازش مدل اندازه‌گیری خود به سه بخش پایایی شاخص، روایی همگرا و روایی واگرا تقسیم می‌شود که در ادامه مطرح می‌شود.

الف-۱) پایایی شاخص؛ در بررسی پایایی شاخص لازم است کلیه بارهای عاملی شاخص‌ها بزرگ‌تر از مقدار ۰/۷۰ بوده و حداقل در سطح ۰/۰۵ معنی‌دار باشند [۲۰] و مقدار ضریب آلفای کرونباخ[۶] با مقادیر بالاتر از ۰/۷۰ پذیرفته و مقادیر کمتر از ۰/۶۰ نامطلوب ارزیابی می‌گردد [۲۱] و همچون آلفای کرونباخ پایایی ترکیبی[۷] نیز در مقادیر بالاتر از ۰/۷۰ پذیرفته شده و مقادیر کمتر از ۰/۶۰ نامطلوب ارزیابی می‌گردد [۲۲] نیز مقادیر اشتراکی[۸] باید اعدادی مثبت باشند که در ادامه و در جدول (۴) به تفکیک درج شده‌اند.



شکل ۱- مدل طراحی پلتفرم داده بزرگ و تأثیر آن بر کیفیت داده‌ها و هوشمندسازی کسب‌وکار [۶،۱۵،۱۶ و ۱۷]

جدول ۵- روایی همگرا در مدل پژوهش

متغیرها	AVE
افزایش هوشمندسازی کسب‌وکار	۰/۷۶۳
سنجش میزان پذیرش کسب‌وکار	۰/۷۴۸
شناخت حوزه‌های داده: الزامات نهفتگی، ساختار داده‌ها	۰/۶۹۴
طراحی پلتفرم داده بزرگ	۰/۶۵۲
فناوری‌های داده بزرگ	۰/۶۹۹
چالش‌های خارجی بی‌کیفیتی داده‌ها	۰/۹۴۵
چالش‌های داخلی بی‌کیفیتی داده‌ها	۰/۸۰۲
چالش‌های کیفیت داده	۰/۷۴۰

جدول ۶- ماتریس فرونل و لارکر مدل

	افزایش هوشمندسازی کسب‌وکار	سنجش میزان پذیرش کسب‌وکار	شناخت حوزه‌های داده: الزامات نهفتگی، ساختار داده‌ها	طراحی پلتفرم داده بزرگ	فناوری‌های داده بزرگ	چالش‌های خارجی بی‌کیفیتی داده‌ها	چالش‌های داخلی بی‌کیفیتی داده‌ها	چالش‌های کیفیت داده
افزایش هوشمندسازی کسب‌وکار	۰/۸۷۳							
سنجش میزان پذیرش کسب‌وکار	۰/۲۰۹	۰/۸۶۵						
شناخت حوزه‌های داده: الزامات نهفتگی، ساختار داده‌ها	۰/۴۱۲	۰/۰۷۳	۰/۸۳۳					
طراحی پلتفرم داده بزرگ	۰/۱۷۴	۰/۱۳۰	۰/۴۴۰	۰/۸۰۷				
فناوری‌های داده بزرگ	۰/۴۲۷	۰/۱۰۷	۰/۵۰۶	۰/۴۰۹	۰/۸۳۶			
چالش‌های خارجی بی‌کیفیتی داده‌ها	۰/۱۹۰	۰/۱۴۵	۰/۴۲۶	۰/۶۹۳	۰/۴۴۴	۰/۹۷۲		
چالش‌های داخلی بی‌کیفیتی داده‌ها	۰/۰۱۹	۰/۰۰۴	۰/۳۲۸	۰/۴۰۵	۰/۳۶۹	۰/۲۱۰	۰/۸۹۶	
چالش‌های کیفیت داده	۰/۵۲۰	۰/۰۱۲	۰/۳۱۵	۰/۰۸۷	۰/۲۱۰	۰/۱۲۳	۰/۶۰۶	۰/۷۶۰

همچنین با توجه به نمودار ۱ (نمودار ضرایب مسیر و بارهای عاملی مدل) کلیه بارهای عاملی شاخص‌ها بزرگ‌تر از مقدار ۰/۵۰۰ می‌باشد که نشان‌دهنده پایایی شاخص‌ها در مدل است. الف-۲) روایی همگرا؛ یا اعتبار همگرایی، همبستگی زیاد شاخص‌های یک سازه را در مقایسه با همبستگی شاخص‌های سازه‌های دیگر نشان می‌دهد که در مدل باید ارزیابی شود. به‌منظور ارزیابی اعتبار همگرایی در نرم‌افزار Smart PLS از AVE استفاده می‌شود. مقدار این ضریب نیز، از ۰ تا ۱ متغیر است که مقادیر بالاتر از ۰/۵۰ پذیرفته شده است [۲۲] که در ادامه و در جدول (۵) بیان شده است.

الف-۳) روایی واگرا؛ بیانگر وجود همبستگی‌های جزئی بین شاخص‌های یک سازه و شاخص‌های سازه‌های دیگر است که در مدل باید ارزیابی شود [۲۰] و بدین‌منظور می‌توان از معیار For-nell-Larcker [۲۲] استفاده نمود. معیار Fornell-Larcker اشاره به این مسئله دارد که ریشه دوم مقادیر واریانس شرح داده شده (AVE) هر سازه، بزرگ‌تر از مقادیر همبستگی آن سازه با سازه‌های دیگر باشد و مقادیر موجود در روی قطر اصلی ماتریس (جذر AVE)، باید از کلیه مقادیر موجود در ستون مربوطه بزرگ‌تر باشد [۲۳].

در ادامه و در جدول (۶) ماتریس فرونل و لارکر به‌عنوان شاخص روایی واگرای مدل بیان شده است.

ب) برازش مدل ساختاری

در بررسی برازش ساختاری مدل لازم است به ۴ معیار توجه داشت که در ادامه مطرح‌شده‌اند:

ب-۱) ضرایب معناداری z یا همان مقادیر t-values: با مقایسه مقدار t محاسبه شده برای هر مسیر می‌توان به بررسی تأیید و یا عدم‌تأیید فرضیه پژوهش پرداخت. بدین‌سان اگر مقدار آماره t بزرگ‌تر از قدر مطلق ۱/۹۶ گردد، در سطح اطمینان ۹۵٪ رابطه مورد نظر قابل‌تأیید می‌باشد [۲۵].

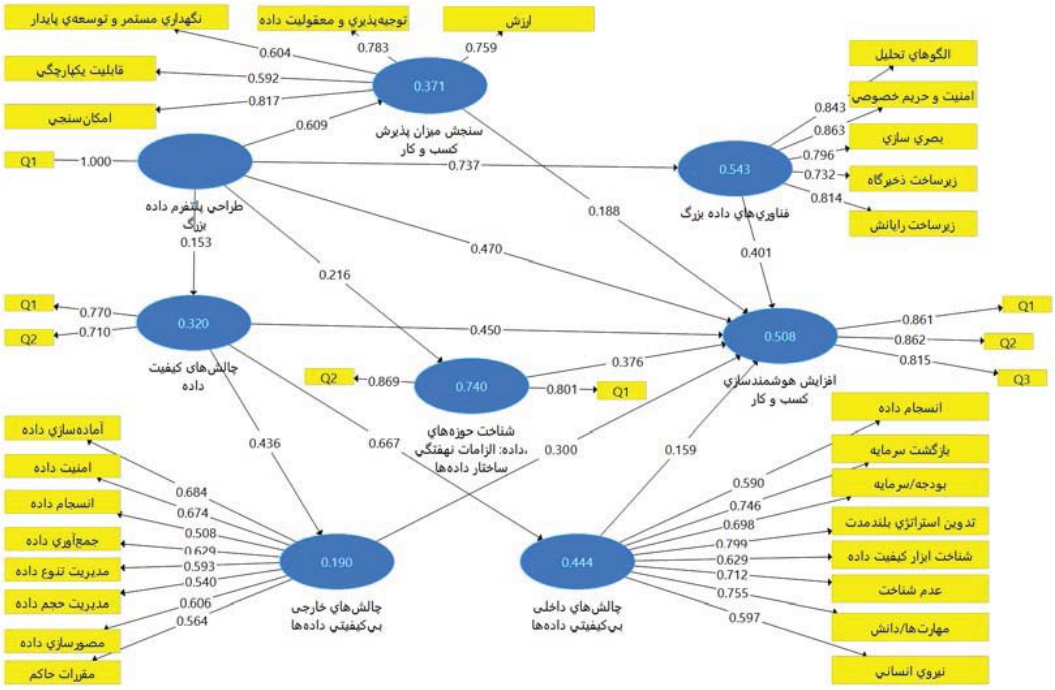
ب-۲) معیار R^2: ضریب تعیین (R^2) ارتباط بین مقدار واریانس شرح داده شده یک متغیر نهفته را با مقدار کل واریانس آن سنجش می‌نماید. مقدار این ضریب نیز از ۰ تا ۱ متغیر است که مقادیر بزرگ‌تر، مطلوب‌تر است [۲۰]، مقادیر نزدیک به ۰/۶۷ مطلوب، نزدیک به ۰/۳۳ را معمولی و نزدیک به ۰/۱۹۰ را ضعیف ارزیابی می‌نمایند [۲۴]. در ادامه و در جدول (۷) مقادیر ضریب تعیین که صرفاً برای سازه‌های سطح دوم مدل وجود دارد بیان شده است.

جدول ۷- مقادیر ضریب تعیین مدل

متغیرها	R^2	R^2 تعدیل‌شده
افزایش هوشمندسازی کسب‌وکار	۰/۵۰۸	۰/۵۰۰
سنجش میزان پذیرش کسب‌وکار	۰/۳۷۱	۰/۳۷۰
شناخت حوزه‌های داده: الزامات نهفتگی، ساختار داده‌ها	۰/۷۴۰	۰/۷۳۹
فناوری‌های داده بزرگ	۰/۵۴۳	۰/۵۴۲
چالش‌های خارجی بی‌کیفیتی داده‌ها	۰/۱۹۰	۰/۱۸۹
چالش‌های داخلی بی‌کیفیتی داده‌ها	۰/۴۴۴	۰/۴۴۳
چالش‌های کیفیت داده	۰/۳۲۰	۰/۳۱۹

جدول ۸- معیار Q^2 مدل

متغیرها	Q^2	Redundancy
افزایش هوشمندسازی کسب‌وکار	۰/۰۵۴	۰/۱۲۵
سنجش میزان پذیرش کسب‌وکار	۰/۱۲۴	۰/۱۲۱
شناخت حوزه‌های داده: الزامات نهفتگی، ساختار داده‌ها	۰/۲۰۰	۰/۲۰۱
فناوری‌های داده بزرگ	۰/۱۴۴	۰/۴۰۶
چالش‌های خارجی بی‌کیفیتی داده‌ها	۰/۱۳۲	۰/۱۶۰
چالش‌های داخلی بی‌کیفیتی داده‌ها	۰/۰۸۸	۰/۳۹۱
چالش‌های کیفیت داده	۰/۰۹۵	۰/۲۱۲



نمودار ۱- ضرایب مسیر و بارهای عاملی مدل

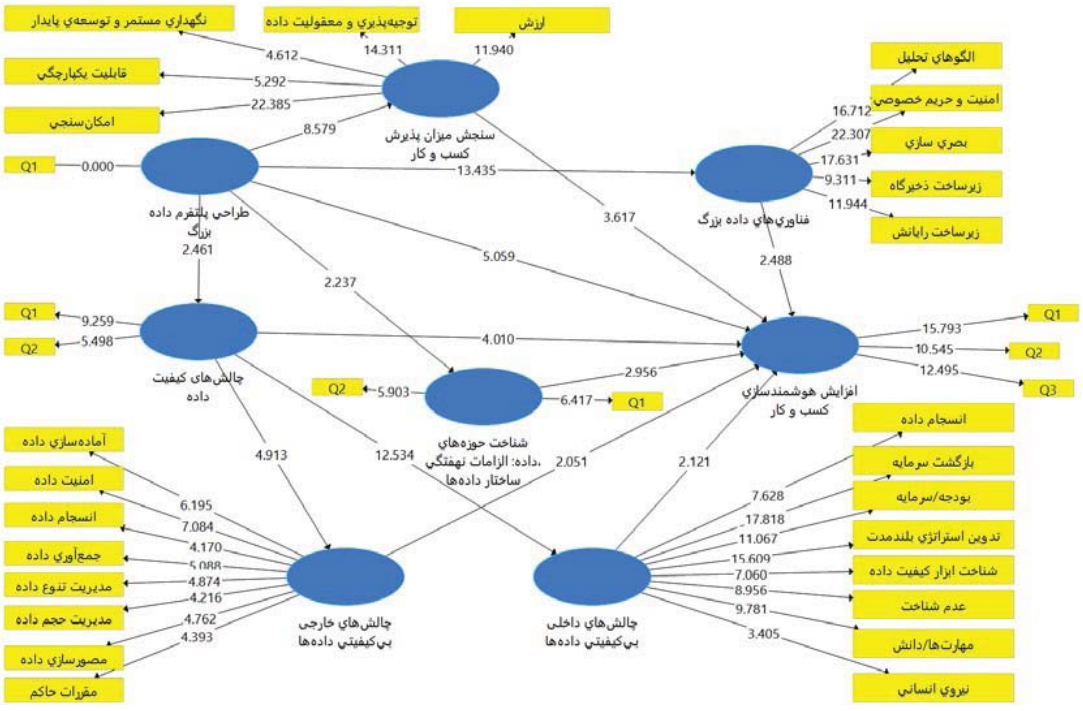
ب-۳) معیار Q^2 (Stone-Geisser Criterion): این سنجه یک معرف تناسب پیش‌بین مدل است و مقدار Q^2 بیشتر از صفر برای یک متغیر نشان‌دهنده تناسب پیش‌بین مدل برای آن متغیر است در مقابل مقادیر صفر و پایین‌تر فقدان تناسب پیش‌بین را نشان می‌دهند. در ادامه و در جدول (۸) مقادیر Q^2 که صرفاً همچون مقادیر R^2 برای سازه‌های سطح دوم مدل وجود دارد بیان شده است.

ب-۴) معیار Redundancy یا شاخص افزونگی: کیفیت مدل ساختاری را نشان می‌دهد و صرفاً برای متغیرهای سطح دوم مدل می‌باشد و باید اعداد مثبت باشند که در مدل پژوهش همه مثبت می‌باشند و در جدول (۸) بیان شده‌اند.

در نمودار (۲) ضرایب مسیر و بارهای عاملی مدل تحقیق مشخص شده است. این نمودار از خروجی آزمون PLS algorithm می‌باشد. اندازه ضریب مسیر نشان‌دهنده قدرت و قوت رابطه بین دو متغیر نهفته است. برخی محققین بر این باور هستند که ضریب مسیر بزرگ‌تر از ۰/۵۰ ضریب مسیری تأثیرگذار در مدل است [۲۶].

جدول ۹- ماتریس اهمیت – عملکرد (IPMA) سازه افزایش هوشمندسازی کسب‌وکار

متغیرهای مکنون	اهمیت	عملکرد
طراحی پلتفرم داده بزرگ	۰/۷۹	۰/۸۰
فناوری‌های داده بزرگ	۰/۷۵	۰/۷۶
سنجش میزان پذیرش کسب‌وکار	۰/۷۴	۰/۶۰
چالش‌های کیفیت داده	۰/۴۹	۰/۵۴
چالش‌های خارجی بی‌کیفیتی داده‌ها	۰/۴۹	۰/۵۰
چالش‌های داخلی بی‌کیفیتی داده‌ها	۰/۴۳	۰/۳۹
شناخت حوزه‌های داده: الزامات نهفتگی، ساختار داده‌ها	۰/۳۰	۰/۴۲



تحلیل ماتریس اهمیت – عملکرد (IPMA)

خصیصه کلیدی رویکرد مدل‌سازی معادلات ساختاری کمترین مربعات جزئی (PLS-SEM) استخراج مقادیر متغیرهای مکنون است. تحلیل ماتریس اهمیت –عملکرد (IPMA) در بسط یافته‌های نتایج اساسی PLS-SEM با استفاده از مقادیر متغیرهای مکنون مفید است. این توسعه مبتنی برآوردهای PLS-SEM از روابط مدل مسیری است و یک بعد دیگر به تحلیل اضافه می‌کند که مقادیر متوسط متغیرهای مکنون را مدنظر قرار می‌دهد. برای یک متغیر مکنون درون‌زای معین که یک سازه کلیدی هدف را در تحلیل نشان می‌دهد IPMA اثرات کل مدل ساختاری (اهمیت) و مقادیر متوسط متغیرهای مکنون (عملکرد) را مقابله می‌دهد تا حوزه‌های پراهمیت را مشخص کند [۲۷].

اجرای ماتریس اهمیت – عملکرد (IPMA) نیازمند شناسایی یک سازه هدف است که ما در این پژوهش متغیر افزایش هوشمندسازی کسب‌وکار را به‌عنوان سازه هدف انتخاب کردیم، به‌عبارت‌دیگر اهمیت متغیرهای مکنون برای یک سازه هدف که به وسیله ماتریس اهمیت – عملکرد (IPMA) تحلیل می‌شود از اثرات کل این متغیرها به وجود می‌آید. در ادامه و در جدول (۹) نتایج ماتریس اهمیت – عملکرد (IPMA) سازه افزایش هوشمندسازی کسب‌وکار بیان شده است.

در جدول (۹) می‌توان اثرات کل (اهمیت) و مقادیر شاخص (عملکرد) برای متغیر افزایش هوشمندسازی کسب‌وکار را در مقیاسی از صفر تا ۱۰۰ درصد دید. با توجه به این جدول طراحی پلتفرم داده بزرگ دارای بالاترین عملکرد سازه نسبت به دیگر سازه‌ها است ضمن اینکه از درجه اهمیت بالایی نیز برخوردار است.

نحوه بررسی روابط مدل

نمودار ۲ پیشرو مربوط به آزمون Bootstrapping می‌باشد که مقادیر T را مشخص می‌کند. همان‌طور که پیش‌ازاین نیز ذکر شد با مقایسه مقدار t محاسبه شده برای هر مسیر می‌توان به بررسی تأیید و یا عدم‌تأیید روابط و اثرات موجود در مدل پژوهش پرداخت. بدین‌سان اگر مقدار آماره t بزرگ‌تر از قدر مطلق ۱/۹۶ گردد، در سطح اطمینان ۹۵٪ رابطه و اثر مورد نظر قابل‌تأیید می‌باشد [۲۷].

جدول ۱۰- شاخص‌های لازم در بررسی روابط و تأثیرات موجود در مدل

مسیرهای مدل	نمونه اصلی	میانگین نمونه	انحراف معیار	آماره t	ضریب مسیر	p-value
سنجش میزان پذیرش کسب‌وکار -> افزایش هوشمندسازی کسب‌وکار	۰/۱۸۸	۰/۲۱۵	۰/۰۵۹	۳/۶۱۷	۰/۱۸۸	۰/۰۰۲
شناخت حوزه‌های داده: الزامات نهفتگی، ساختار داده‌ها -> افزایش هوشمندسازی کسب‌وکار	۰/۰۷۶	۰/۰۹۲	۰/۵۵۹	۲/۹۵۶	۰/۳۷۶	۰/۰۰۳
طراحی پلتفرم داده بزرگ-> افزایش هوشمندسازی کسب‌وکار	۰/۰۷۴	۰/۰۴۸	۰/۸۱۹	۵/۰۵۹	۰/۴۷۰	۰/۰۰۰
طراحی پلتفرم داده بزرگ-> سنجش میزان پذیرش کسب‌وکار	۰/۶۰۹	۰/۱۵۵	۰/۱۹۳	۸/۵۷۹	۰/۶۰۹	۰/۰۰۰
طراحی پلتفرم داده بزرگ-> شناخت حوزه‌های داده: الزامات نهفتگی، ساختار داده‌ها	۰/۲۱۶	۰/۱۸۰	۰/۱۶۳	۲/۲۳۷	۰/۲۱۶	۰/۰۰۶
طراحی پلتفرم داده بزرگ-> فناوری‌های داده بزرگ	۰/۷۳۷	۰/۱۲۲	۰/۲۰۷	۱۳/۴۳۵	۰/۷۳۷	۰/۰۰۰
طراحی پلتفرم داده بزرگ-> چالش‌های کیفیت داده	۰/۱۵۳	۰/۱۵۲	۰/۱۳۰	۲/۴۶۱	۰/۱۵۳	۰/۰۲۳
فناوری‌های داده بزرگ -> افزایش هوشمندسازی کسب‌وکار	۰/۴۰۱	۰/۱۰۰	۰/۳۲۳	۲/۴۸۸	۰/۴۰۱	۰/۰۳۳
چالش‌های کیفیت داده-> افزایش هوشمندسازی کسب‌وکار	۰/۰۳۰	۰/۱۶۱	۰/۳۰۰	۴/۰۱۰	۰/۴۵۰	۰/۰۰۰
چالش‌های داخلی بی‌کیفیتی داده‌ها -> افزایش هوشمندسازی کسب‌وکار	۰/۱۵۹	۰/۱۶۰	۰/۱۰۰	۲/۱۲۱	۰/۱۵۹	۰/۰۱۲
چالش‌های خارجی بی‌کیفیتی داده‌ها -> افزایش هوشمندسازی کسب‌وکار	۰/۰۴۵	۰/۵۶	۰/۰۹۵	۲/۰۵۱	۰/۳۰۰	۰/۰۲۶
چالش‌های کیفیت داده -> چالش‌های خارجی بی‌کیفیتی داده‌ها	۰/۴۳۶	۰/۳۶۰	۰/۰۶۹	۴/۹۱۳	۰/۴۳۶	۰/۰۰۰
چالش‌های کیفیت داده -> چالش‌های داخلی بی‌کیفیتی داده‌ها	۰/۶۶۷	۰/۵۲۰	۰/۱۱۰	۱۲/۵۳۴	۰/۶۶۷	۰/۰۰۰

با توجه به نمودارهای خروجی ۲ و ۳ نرم‌افزار در مدل‌سازی پلتفرم داده بزرگ و لزوم آن در کیفیت داده و هوشمندی کسب‌وکار در ادامه و در جدول (۱۰) کلیه روابط مدل با شاخص شدت و قدرت اثر (ضریب مسیر) و معناداری (T-values و P-value) بیان می‌شود:

فرایند پژوهش

در این پژوهش در ابتدا به مدل‌سازی و تعریف روابط میان متغیرهای «طراحی پلتفرم داده بـزرگ»، « شناخت حوزه‌های داده در کسب‌وکار»، « فـناوری‌هـای داده بـزرگ»، «سنجش میزان پذیرش کسب‌وکار»، « افزایش هوشمندسازی کسب‌وکار»، «چالش‌های کیفیت داده»، «چالش‌های خـارجی بی‌کیفیتی داده‌ها» و «چالش‌های داخلی بی‌کیفیتی داده‌ها» پرداخته شده است. ازآنجایی‌که نـوع تحقیق مدل‌سازی و اکتشافی است از رویکرد مدل‌سازی معـادلات ساختاری کمترین مربعات جزئی (PLS-SEM) استفاده شـد. در شکل ۱ مـدل پیشنهادی این پژوهش به‌منظور بیان ارتباط بسیار نزدیک طراحی پلتفرم داده بـزرگ با چالش‌های بی‌کیفیتی داده‌هـا و افزایش هوشمندسازی کسب‌وکار نمایش داده شده است. در طراحی پلتفرم داده

بزرگ پیشنهادی با توجه به مبانی نظری ابتدا و در مرحله اول حوزه داده‌های هر کسب‌وکار شامل الزامات نهفتگی و ساختار آن بررسی و شناسایی می‌شوند. سپس در مرحله دوم میزان قابلیت پذیـرش داده بـزرگ از سوی کسب‌وکار با توجه به معیارهایی مانند امکان‌سنجی، توجیه‌پذیری و معقولیت، ارزش، قابلیت و یکپارچگی، نگه‌داری مستمر و توسعه پایدار موردسنجش قرار می‌گیرد. در مرحله پایانی با توجه به دسته‌بندی معرفی شده از فناوری‌های داده بزرگ، نتایج حاصل از دو مرحله قبلی و هوشمند ساختن کسب‌وکار، نـرم‌افـزارهـای مناسب بررسی و پیاده‌سازی می‌شوند.

فناوری‌های داده بزرگ در فرآیند طراحی پلتفرم داده بزرگ در این پژوهش شامل زیرساخت‌های رایانش و ذخیره‌گاه، الگوهای تحلیل، امنیت و حریم خصوصی داده‌ها و بصری‌سازی می‌باشد که تمامی چالش‌های خارجی بی‌کیفیتی داده با برنامه‌ریزی و استقرار پلتفرم داده بزرگ مرتفع و ابزار کیفیت داده شناسایی خواهند شد. همچنین زیرساخت‌های ذخیره‌گاه و انتخاب الگوهای تحلیل از فناوری‌های داده بزرگ لازمه هوشمندسازی کسب‌وکار است که افزایش کیفیت داده‌ها، هوشمندی کسب‌وکار را افزایش داده و منجر به تصمیم‌گیری سریع‌تر و بهتر در کسب‌وکار موردنظر خواهد شد. ریسک‌های کیفیت داده مجموعه‌ای از خطاهای انسانی، فقدان منابع داخلی، استراتژی نامناسب، عدم‌کفایت تکنولوژی مرتبط در حال استفاده، فقدان تکنولوژی مرتبط، فقدان مهارت موردنـیاز بـرای استـفاده صحیح از تـکـنولوژی مـوجـود، فقدان ارتباطات داخلی بین واحدها، ناکافی بودن حمایت مدیریت ارشد، بودجه ناکافی، عدم وجود مدیریت داده (حجم، تنوع و سرعت تولید)، عدم وجود زیرساخت‌های امنیتی داده، فقدان مقررات مناسب (به دلیل وجود ساختارهای بوروکراتیک) بیان شده است. لذا در صورت عدم رفع ریسک‌های کیفیت داده امکان هوشمندسازی کسب‌وکار امکان‌پذیر نخواهد بود بلکه نتیجه‌ای معکوس خواهد داشت.

به این منظور سازه‌های «طراحی پلتفرم داده بزرگ» با یک سؤال (سنجه)، « شناخت حوزه‌های داده در کسب‌وکار» با دو سؤال (سنجه)، « فناوری‌های داده بزرگ» با پنج سؤال (سنجه)، «سنجش میـزان پذیـرش کسب‌وکار» با شش سؤال (سنجه)،

« افزایش هوشمندسازی کسب‌وکار» با سه سؤال (سنجه)، «چالش‌های کیفیت داده» با دو سؤال (سنجه)، «چالش‌های خارجی بی‌کیفیتی داده‌ها» با هشت سؤال (سنجه) و «چالش‌های داخلی بی‌کیفیتی داده‌ها» با هشت سؤال (سنجه) به‌عنوان سازه‌ها در مدل کمترین مجذور مربعات تعریف و وارد نرم‌افزار شده‌اند. بعد از سنجش روایی و پایایی مدل، همان‌طور که در نمودار ۲ نشان داده شده است، مقادیر آزمون t در این نرم‌افزار حاکی از معنادار بودن تمامی روابط تعریف شده در مدل ساخته شده دارد.

بحث و نتیجه‌گیری

با توجه به جدول ۱۰ و نمودارهای ۱ و ۲ کلیه متغیرهای مورد بررسی مدل و سنجه‌های مربوط در سطح اطمینان ۵۹ درصدی معنادار شدند فقط سنجه «حسابرسی» از «چالش‌های داخلی بی‌کیفیتی داده‌ها» به علت پایین بودن میزان بارعاملی خود با توجه به داده‌های پرسشنامه حذف شد – در مدل‌سازی کمترین مربعات جزئی سنجه‌های با بارعاملی زیر ۵۰ درصد حذف می‌شوند – دراین‌بین طراحی پلتفرم داده بزرگ و چالش‌های کیفیت داده نسبت به دیگر متغیرهای مورد بررسی دارای بالاترین میزان اثر بر هوشمندسازی کسب‌وکار با توجه به مقدار ضریب مسیر خود شدند. همچنین با توجه به ماتریس اهمیت – عملکرد در جدول ۸، طراحی پلتفرم داده بزرگ، فناوری‌های داده بزرگ و سنجش میزان پذیرش کسب‌وکار به ترتیب دارای بیشترین درصد اهمیت و عملکرد در مدل می‌باشند. همچنین سنجه‌های «تدوین استراتژی بلندمدت»، «مهارت/دانش» و «بازگشت سرمایه» به ترتیب با توجه به بار عاملی خود در نمودار ۱ (ضرایب مسیر و بارهای عاملی) دارای بالاترین قدرت تبیین متغیر چالش‌های داخلی بی‌کیفیتی داده‌ها را از خود نشان دادند. سنجه‌های «آماده‌سازی داده»، «امنیت داده» و «جمع‌آوری داده» به ترتیب دارای بالاترین قدرت تبیین متغیر چالش‌های خارجی بی‌کیفیتی داده‌ها هستند و باز با توجه به همان نمودار ۱ امنیت حریم خصوصی، الگوهای تحلیل و زیرساخت رایانش به ترتیب دارای بالاترین قدرت تبیین متغیر فناوری داده بزرگ هستند و امکان‌سنجی، توجیه‌پذیری و معقولیت داده و ارزش به ترتیب دارای بالاترین قدرت تبیین متغیر سنجش میزان پذیرش کسب‌وکار هستند. حال با توجه به این نتایج نظر نگارندگان آن است که با توجه به نقش مؤثر طراحی پلتفرم داده بزرگ در کسب‌وکارها در فرهنگ‌سازی چالش‌های داخلی از یک‌سو و چالش‌های خارجی بی‌کیفیتی داده‌ها از سوی دیگر تا حد زیادی مرتفع می‌شوند و هوشمندسازی کسب‌وکار را سبب می‌شود.

پی‌نوشت

1 Return of investment(ROI)
2 External Data
3 Data- Centric
4 . Latency Requiremments
5 .Structure
6 Cronbach Alpha
7 Composite Reliability (CR)
8 Communality

منابع

[1]. Kučera, J., Chlapek, D., & Nečaský, M.,2013. Open government data catalogs: Current approaches and quality perspective. Technology-Enabled Innovation for Democracy, Government and Governance, Lecture Notes in Computer Science. 8061. (pp. 152–166). Berlin Heidelberg: Springer. http://dx.doi.org/10.1007/978-3-642-40160-2_13.

[2]. S. W. Tee, P.L. Bowen, P. Doyle, F.H. Rohde, "Factors influencing organizations to improve data quality in their information systems, « Accounting & Finance, vol. 47, pp. 335-355, 2007.

[3]. C. Batini, C. Cappiello, C. Francalanci, A. Maurino, "Methodologies for data quality assessment and improvement, « ACM Computing Surveys (CSUR), vol. 41, p. 16, 2009.

[4]. F. G. Alizamini, M.M. Pedram, M. Alishahi, K. Badie, "Data quality improvement using fuzzy association rules, « 2010, pp. V1-468-V1- 472.

[5]. Pipino, L.L., Lee, Y.W., &Wang, R.Y., Apr 2002. Data quality assessment. Communications of the ACM, 45(4), 211–218. http://dx.doi.org/10.1145/505248.506010.

[6]. The 2016 global data management benchmark report,2016. Experian data quality.

[7]. Patterson, T., 2015, « The Use of Information Technology in Risk Management", Complex Solutions Executive IBM Corporation.

[۸]. Castro, D., Korte, T.,2013, "Data Innovation 101: An Introduction to the Technologies and Policies Supporting Data-Driven Innovation", Center for Data Innovation.

[۹]. Eastwood, M., Vesset, D. and Morris, D.H. (2005), "Delivering value in business intelligence", HP White Paper, available at: http://research.ittoolbox.com/white-papers/lg.asp?grid ¼ 3374 (accessed March 13, 2007).

[۱۰]. Armbrust, M., Armando Fox, Rean Griffith, Anthony D. Joseph, Randy Katz, Andy Konwinski, Gunho Lee, David Patterson, Ariel Rabkin, Ion Stoica, Matei Zaharia, 2010. "A View of Cloud Computing, « Commun. ACM, Vols. vol.53, no. 4, pp. 50-58.

[۱۱]. Combs, M., (2014, 11/19/2014). Study Shows Big Return on Big Data. Available: http://www.veristorm.com/content/study-shows-big-return-big-data.

[۱۲]. McKinsey Global Institute report,'Big Data: the next frontier for innovation, Competition and Productivity"May 2011.

[۱۳]. Ginsberg, J., Mohebbi, H., Patel, R., Brammer, L., Smolinski, M., Brilliant, L., Detecting influenza epidemics using search engine query data. Nature, 457(7232): 1012-1014,2008.

[۱۴]. Rivera, J., (2014, 11/19/2014). Gartner Survey Reveals That 73 Percent of Organizations Have Invested or Plan to Invest in Big Data in the Next Two Years. Available: http://www.gartner.com/newsroom/id/2848718.

[۱۵]. Herschel, R.T. and Jones, N.E. (2005), "Knowledge management and business intelligence: the importance of integration", Journal of Knowledge Management, Vol. 9 No. 4, pp. 45-55.

[۱۶]. روحانی، سعید. حسینی، سمیه. تحلیل‌های عظیم داده، نقشه راه پیاده‌سازی، فناوری و ابزارها. انتشارات نیاز دانش. چاپ اول. ۹۴.

[۱۷]. BIG DATA WORKING GROUP, 2014, "Big Data Taxonomy",cloud security alliance.

[۱۸]. علی اکبری و همکاران، پی‌دی‌اف مقالات همایش داده‌های عظیم، ۱۳۹۳.

[۱۹]. محسنین، شهریار، اسفیدانی، رحیم، ۱۳۹۶، معادلات ساختاری مبتنی بر رویکرد حداقل مربعات جزئی به کمک نرم‌افزار Smart-PLS، نشر کتاب مهربان ، چاپ دوم، تهران.

[۲۰] Cepeda, G., Nitzl, C., and Roldán, J. L. (۲۰۱۷). Mediation Analyses in Partial Least Squares Structural Equation Modeling: Guidelines and Empirical Examples., in Partial Least Squares Path Modeling: Basic Concepts, Methodological Issues and Applications, H. Latan and R. Noonan (eds.), Springer: Cham, pp. ۱۹۵-۱۷۳.

[۲۱] Hair, J. F., Hollingsworth, C. L., Randolph, A. B., and Chong, A. Y. L. An Updated and Expanded Assessment of PLS-SEM in Information Systems Research. Industrial Management & Data Systems, Volume 117(2017), Issue 3, pp. 442-458.

[۲۲]. Hair, J. F., Hult, G. T. M., Ringle, C. M., and Sarstedt, M. (۲۰۱۷). A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM), ۲nd Ed., Sage: Thousand Oaks.

[۲۳] Rigdon, E. E., Sarstedt, M., & Ringle, C. M. (2017). On Comparing Results from CB-SEM and PLS-SEM. Five Perspectives and Five Recommendations. Marketing ZFP, 39(3), 4-16.

[۲۴]. آذر، عادل، غلام‌زاده، رسول، ۱۳۹۵، مدل‌سازی معادلات ساختاری کمترین مربعات جزئی PLS-SEM، نشر نگاه دانش، تهران

[۲۵]. داوری، علی و رضازاده، آرش (۱۳۹۷)، مدل‌سازی معادلات ساختاری با نرم‌افزار PLS، چاپ دوم، انتشارات جهاد دانشگاهی، تهران.

[۲۶]. حنفی زاده، پیام و زارع رواسان، احد (۱۳۹۱)، روش تحلیل ساختارهای چند سطحی با استفاده از نرم‌افزار smart pls، انتشارات ترمه، تهران.

[۲۷]. Hair, F., Marko, J.S., Lucas, H., and Volker, G.K., 2014. Partial least squares structural equation modeling (PLS-SEM) An emerging tool in business research. European Business Review, 26 (1), 106-121.